

P-9-18-07
App. 9-18-07
Info. 10-9-07

07-12

Undergraduate Distance Education Review Form
(Required for all courses taught by distance education for more than one-third of teaching contact hours.)

Existing and Special Topics Course

Course: EDSP 477 Assessment of Student Learning

Instructor(s) of Record: Dr. Gurmala Rattan

Phone: 724 357-3787 _____

Email: gurmala@iup.edu

Step One: Proposer

A. Provide a brief narrative rationale for each of the items, A1- A5.

1. How is/are the instructor(s) qualified in the distance education delivery method as well as the discipline?
2. How will each objective in the course be met using distance education technologies?
3. How will instructor-student and student-student, if applicable, interaction take place?
4. How will student achievement be evaluated?
5. How will academic honesty for tests and assignments be addressed?

B. Submit to the department or its curriculum committee the responses to items A1-A5, the current official syllabus of record, along with the instructor developed online version of the syllabus, and the sample lesson. This lesson should clearly demonstrate how the distance education instructional format adequately assists students to meet a course objective(s) using online or distance technology. It should relate to one concrete topic area indicated on the syllabus.

Step Two: Departmental/Dean Approval

Recommendation: Positive (The objectives of this course can be met via distance education)

Negative

Edward M. Zeverson 9/11/07
Signature of Department Designee Date

Received
SEP 12 2007

Endorsed: May Ann Ralath 9-12-07
Signature of College Dean Date

Liberal Studies

Forward form and supporting materials to Liberal Studies Office for consideration by the University-wide Undergraduate Curriculum Committee. Dual-level courses also require review by the University-wide Graduate Committee for graduate-level section.

Step Three: University-wide Undergraduate Curriculum Committee Approval

Recommendation: Positive (The objectives of this course can be met via distance education)
 Negative

Gail Sechrist 9-18-07
Signature of Committee Co-Chair Date

Forward form and supporting materials to the Provost within 30 calendar days after received by committee.

Step Four: Provost Approval

Approved as distance education course Rejected as distance education course

Chad Samuel 9/25/07
Signature of Provost Date

Forward form and supporting materials to Associate Provost.

EDSP 477 – Assessment of Student Learning – Distance Education Course Proposal

Gurmal Rattan, Ph.D.

A 1. How is/are the instructor(s) qualified in the distance education delivery method as well as the discipline?

The instructor has received several College of Education and Educational Technology mini grants dating back to 2000-2001 to develop course materials using Webct technology. In addition, the instructor also received 3 additional mini-grants in 2003-2004 to assist three other faculty members in the department to develop their course materials using Webct. Currently, all course work/materials for this instructor has utilized Webct since 2001. Additionally, the instructor has also taught GSR 615_DE Elements of Research as a distance education course during the Spring 2007 semester. The course was well received and student comments are available for review.

A 2. How will each objective in the course be met using distance education technologies?

The objectives for the course will be met in a similar fashion to the regular course. Students will be required to read the assigned materials in the textbook and complete the related workbook activities. A discussion group format will be used to enable students to originate questions/comments via posts and responds to other students' posts as discussed in the next section. In addition, power point notes of the lecture will also be available on Webct. There will also be quizzes and exams that the student will take via Webct to assess their mastery of the course materials. Students will also be required to complete class assignments (e.g., developing Short-answer, True-False, Matching, Multiple-Choice, Essay questions) and submit their responses/projects via Webct.

A 3. How will instructor-student and student-student, if applicable, interaction take place?

The main feature to promote interactions between the students will be the Discussion tool of Webct. In teaching the GSR 615_DE, it was effectively used to promote student interactions. Students will be required to post one original messages per chapter and follow up on other student's responses to this post. In addition, each student will be required to respond to 1-2 postings from other students. The instructor will add to the above contributions to clarify, expand, or prompt students with other questions. The students will be separated into manageable groups to ensure a healthy participation by each student. The email feature will also be used to address personal questions students may have. The instructor will also be available by phone during office hours as needed.

A 4. How will student achievement be evaluated?

Student achievement will be assessed by using Webct based chapter quizzes, exams, and their participation in the Discussion groups. In addition, students will be submitting their test construction project (KARS) on Webct and/or the P drive. The workbook will also be submitted for grading.

A 5. How will academic honesty for tests and assignments be addressed?

Students will be expected to uphold the university academic integrity policy for this course. To aid in this, Webct based quizzes and exams will be time restricted and only accessible to students at predetermined times.

EDSP 477 Syllabus of Record

Syllabus of Record

EDSP 477 Assessment of Student Learning: Design & Interpretation of Educational Measures

I. Course Description: This course is designed to acquaint students with major methods and techniques of evaluation used to assess and report growth, development, and academic achievement of learners in preschool, elementary, and secondary schools, including interpretation of standardized test information.

II. Course Objectives: General course objectives for the student include:

1. Awareness of the role of assessment in teaching
2. Understanding of the role of reliability and validity in both classroom and standardized tests
3. Understanding of the various methods of assessment and their alignment with learning goals
4. Skill building in the development of various teacher-made tests and evaluative procedures
5. Capability in the use of assessment before, during, and after instruction
6. Awareness of the needs of special populations, such as those with disabilities, multicultural populations and those not proficient in English, as they relate to assessment
7. Interpretation and utilization of data/scores provided by standardized tests through the use of elementary statistics
8. Awareness of trends, issues, and relevant research in assessment related to educational reform
9. Awareness of the use of technology in assessment

Course Objective	Course Assessment Measuring Objective
1	Test 1, Test Critique, Test Construction Project*
2	Exam 1, Class Activities,
3	Test Construction Project*, Exam II
4	Test Construction Project*, Class Activities
5	Test Construction Project*, Exam II
6	Class Activities, Exam II
7 & 8	Exam II
9	Website Activity

* Indicates Key Assessment

COURSE SCHEDULE*

Date Reading	Topic	Assigned
Week 1	Course Overview/Role of Assessment in Education	

Use of Assessment Throughout the Instructional Process

Sample Activities**: Print the word *Test* on the board. Have students write their immediate reactions & discuss the emotional elements of testing/Have students write a brief paragraph describing their worst & best test experience /Present lifelike student scenarios. Ask students what they need to know to better serve the student's educational needs Then have them choose what type of assessment method might help them get that information./Have students read a controversial article on assessment & discuss/Have students review test publisher catalogs to learn the variety of tests, as well as the costs, & business aspect of testing

Week 2	<p>The Role of Assessment (cont'd) Validity/Reliability Sample Activities: Have students locate an article in the popular press relating to testing & assessment & have students write a reflection paper/List several aspects of validity & reliability & have students note whether they apply to validity, reliability, or both</p>	L & M Ch 1, 2, 4
Week 3	<p>Validity/Reliability/Instructional Objectives Sample Activities: Have students choose one thing they are good at & list what someone would have to learn to be good at that activity. Ask them how they would determine if the person had learned it or not/Given a pair of objectives, have students choose which is stated in performance terms/Have students list learning outcomes for 3 general instructional objectives</p>	Ch 5
Week 4	<p>Quiz #1/ Instructional Objectives Components of Instrument Construction</p>	Ch 3
Week 5 Responsibility	<p>Ethics & Procedures in Testing Planning Classroom Instruments and Assessments Sample Activities: Using life-like scenarios, have an "ethics committee" of students reflect on how they might be handled and what laws, best practices, & ethical issues are involved. Provide Code of Responsibilities in Measurement for use in the activity</p>	Code of Ch 6
Week 6	<p>Oral Questioning Techniques Creating Pencil & Paper Assessments Website Assignment Due Sample Activities: Pose several questions and have students evaluate the quality of each/Have students list 2 effective & 2 ineffective questions they hear in their own classes on a given day</p>	Ch 7, 8
Week 7	<p>Written Tests-Constructing Selected Response & Free Response Items Sample Activities: Have a table of model tests constructed by past classes/Have students create test items based on a text chapter. Then have the different teams evaluate each other's items</p>	Ch 9, 10

Week 8	Exam 1	
Week 9 12	Performance-Based Assessment/Portfolios Sample Activities: Use essay responses from a prior 477 class and have students evaluate with a rubric	Ch 11,
Week 10	Performance-Based Assessment/Portfolios (cont'd) Test Construction Project Work Time TEST CRITIQUES DUE	Ch 14
Week 11	Quiz #2/Assembling & Administration Curriculum-Based Assessment, for example, DIBELS Suggested Activity: Have students view these procedures by accessing the AIMSWeb and DIBELS websites in class	
Week 12 17	TEST CONSTRUCTION PROJECT DUE Achievement & Aptitude Tests Suggested Activities: Bring selected tests to class for student review	Ch 16,
Week 13 18, 19	Evaluating and Interpreting Test Performance, for example, PSSA Grading & Reporting Graphing of Pre and Post Assessment Data Sample Activities: Have students do scatter plots on the board for a variety of hypothetical tests in different content areas at different schools & interpret meaning/Obtain PSSA reports from a school district & have them interpret as a class/Stage mock parent conferences where teachers (students) provide standardized test results to parents	Ch 15,
Week 14	Special Populations/Current Trends & Issues Sample Activities: Show samples of real tests that have been modified for students with special needs/ Have students review school district reports to the community and newspaper articles for themes regarding the role of assessment/ Have a speaker from a school district or IU to present on methods of modification & assistive technology	
Week 15	Exam II (Final Exam)	

* Topic dates are approximate

**Sample activities are suggestions only and will vary by instructor

III. Sample Course Requirements & Evaluation Methods (May vary with instructor)

- ◆ 2 exams and 2 quizzes covering reading and class material (Each exam is worth 45 points. Each quiz is worth 20 points.)

Exams and quizzes will include multiple choice, short answer, and essay questions.

Responses will be evaluated with respect to accuracy and amount of substantive content;

evidence of understanding of material covered in lectures, reading and class discussion; and depth of awareness regarding issues related to assessment in education.

- ◆ Construction of a teacher-made test in the student's content area of choice (worth 90 pts.)

This test must include a variety of item types as well as a performance assessment and a description of how the test would be modified for a child with a specific disability. A table of specifications should also be provided. Students should do readability studies on their instrument to insure the appropriate grade level. Procedures for doing this will be provided in class.

Further instruction to be provided in class.

- ◆ Critique of 1 test. It can be a test you have been administered, a test from a class in which you are currently observing or teaching, or a test from a packaged curriculum. (worth 60 pts.)

Further instruction to be provided in class.

- ◆ The location of 1 Web site related to issues of assessment for classroom teachers. You will be given a list of questions to answer regarding this website. (worth 15 pts.)
- ◆ Participation in the classroom learning community including involvement in cooperative learning activities, class discussion, and in-class exercises.

Class Policies

Assignments will be due as scheduled. See University policies regarding attendance, grading, plagiarism, and withdrawal from a course

Required Reading (Alternatives will be reviewed by the Department chair or curriculum committee.)

Linn, R., & Miller, M. D. (2005). *Measurement and assessment in teaching* (9th ed.). Columbus: Prentice Hall.

Linn, R., & Miller, M. D. (2005). *Student exercise manual for measurement and assessment in teaching* (9th ed.). Columbus: Prentice Hall.

Reading Packet (available from Pro-Packet)

NCME (1995). *Code of professional responsibilities in educational measurement*. Washington, DC: Author.

Pennsylvania Department of Education 2005 Accommodation Guidelines. Harrisburg: Author.

Sample Grading Scale

A - 90%-100% of total pts possible
B - 80%-89% of total pts possible

C - 70%-79% of total pts possible
D - 60%-69% of total pts possible
F - less than 60% of total pts possible

V. Resource Requirements

No special materials or equipment required other than readings and test to critique.

V. Bibliography

- Airasian, P. (2005). *Classroom assessment: Concepts and applications*. New York: McGraw-Hill.
- American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Anderson, L. W. (2003). *Classroom assessment: Enhancing the quality of teacher decision making*. Mahwah, NJ: Lawrence Erlbaum.
- Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13-18.
- Arter, J., & Busick, K. U. (2001). *Practice with student involved classroom assessment*. Upper Saddle River, NJ: Merrill Prentice Hall. This is a workbook with many examples and ideas for creative assessment approaches.
- Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests. *Educational Research*, 44(3), 279-288.
- Burger, S. E., & Burger, D. L. (1994). Determining the validity of performance-based assessment. *Educational Measurement: Issues and Practice*, 31(1), 9-15.
- Chase, C. I. (1999). *Contemporary assessment for educators*. New York: Longman.
- Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning, *Educational Leadership* 60(1), 40-44.
- Frisbie, D.A., Miranda, D. U., & Baker, K.K. (1993). An evaluation of elementary textbook tests as classroom assessment tools. *Applied Measurement in Education*, 6(1), 21-36.
- Gallagher, J. D. (2000). *Classroom assessment for teachers*. Upper Saddle River, NJ: Merrill.
- Gronlund, N. (2000). *How to write and use instructional objectives*. Upper Saddle River, NJ: Prentice Hall. Provides assistance in writing clear objectives essential to the instructional and assessment process.
- Guskey, T. R. (2003). How classroom assessments improve learning. *Educational Leadership*, 60(5), 6-11.
- Herrnstein, R.J. & Murraray, C. (1994) *The bell curve*. New York: Free Press.
A controversial look at the use of assessment, intellectual ability, and achievement in American society. Guaranteed to stimulate your thinking!
- Hoy, C., & Gregg, N. (1995). *Assessment: The special educator's role*. Pacific Grove, CA: Brooks/Cole.

- Hughes, S. (1993). What is alternative/authentic assessment and how does it impact special education? *Educational Horizons*, 72(3), 28-35.
- Klitz, G. (2002). Involving students in the improvement process through portfolios. *Journal of School Improvement*, 3(1), 54-56.
- Lane, S., Mei, L., Ankenmann, R.D., & Stone, C.A. (1996). Generalizability and validity of a mathematics performance assessment, *Journal of Educational Measurement*, 33(1), 71-92.
- Lescher, M.L. (1995). *Portfolios: Assessing learning in the primary grades*. Washington, D.C: NEA Professional Library.
- Lukhele, R. Thissen, D., & Wainer, H. (1994). On the relative value of multiple choice, constructed response, and examinee selected items on two achievement tests. *Journal of Educational Measurement*, 31(3), 243-250.
- Madhavi, J., Bursuck, W., Havekost, D.M., Epstein, M.H., & Palloway, E.A. (1994). School district testing policies and students with disabilities: A national survey. *School Psychology Review*, 23(4), 694-703.
- McMillan, J.H. (2004). *Classroom assessment: Principles and practice for effective instruction*. Boston: Allyn & Bacon. Good chapter on effective oral questioning and appropriate teacher responses.
- Mehrens, W. A., Popahm, W. J., & Ryan, J. M. (1998). How to prepare students for performance assessments. *Educational Measurement: Issues and Practice*, 17(1), 18-22.
- Montgomery, K. (2001). *Authentic assessment: A guide for elementary teachers*. New York: Longman.
- Naizer, G.L. (1997). Validity and reliability issues of performance-portfolio assessment. *Action in Teacher Education*, 18(4), 1-9.
- Overton, T. (2002). *Assessment in special education: An applied approach*. New York: Merrill.
- Penta, M. Q. (2002). Student portfolios in a standardized world. *Kappa Delta Pi Record*, 38(2), 77-81.
- Popham, W.J. (2005). *Classroom assessment: What teachers need to know*. Boston: Allyn & Bacon. Easy to read and very teacher friendly.
- Sachs, P. (1999) *Standardized minds*. Cambridge, MA: Persuis.
- The President's Advisory Commission on Educational Excellence for Hispanic Americans: *Testing Hispanic students in the United States*. (2000). Washington, DC: Author.
- Salvia, J. & Ysseldyke, J. (2003). *Assessment in special and inclusive education*. Boston: Houghton Mifflin.
- Seidel, S., Walters, J., Kirby, E., Olf, N., Powell, D., Scripp, L., & Veenema, S. (1997). *Portfolio practices: Thinking through the assessment of children's work*. Washington, DC: NEA Professional Library. Contains a good deal of practical advice for laying the groundwork in the development of portfolio use.
- Shinn, M.R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford. An in-depth look at this procedure.
- Siegle, D. (2002). Creating a living portfolio: Documenting student growth with electronic portfolios. *Gifted Child Today*, 23(3), 60-64.

Spinelli, C. G. (2005). *Classroom assessment for students in special and general education*. Upper Saddle River, NJ: Prentice Hall.

Stiggins, R. J. (2005). *Student-involved assessment for learning*. Upper Saddle River, NJ: Merrill Prentice Hall. An emphasis on assessments for classroom use with concentration on the appropriate use of performance-based assessment.

Trice, A. D. (2000). *A handbook of classroom assessment*. New York: Longman.
This is basically a reworking of what is in your text, but sometimes it helps to read it in more than one place and presented a bit differently.

U.S. Department of Education. (1994). *Changing education: Resources for systemic reform*. (PIP 94-1509). Washington, DC: U.S. Government Printing Office. Short articles of interest on many of the topics discussed in this course with special emphasis on issues of importance in the reform movement.

Weber, E. (1999). *Student assessment that works: A practical approach*. Boston: Allyn & Bacon.
Some good ideas on how to collaborate with other educators and work with parents around the assessment issue.

Relevant Websites

www.aimsweb.com

There is a small fee to use this site but you can download materials for curriculum-based assessments.

A site to assist with the readability assessment of your test.
www.med.utah.edu/pated.authors/readability.html

Another site addressing readability levels.
www.school.discovery.com/schrockguide/fry/fry.html

California Reading Comprehension Standards
<http://www.cde.ca.gov/cdepress/Eng-Lang-Dev-Stnd.pdf>

Florida Math Standards
<http://www.firm.edu/doe/curric/prek12/pdf/math6.pdf>

History and Social Science Standards
<http://www.pen.kqw.va.us/VDOE/supertendent/Sols/historysol2001.pdf>

Center for Research on Evaluation, Standards, and Student Testing
<http://www.cse.ucla.edu>

Hands-on science performance assessment task
<http://www.ed.gov/NCES/naep>

A curriculum-based measurement manual for teachers and learn about a variety of academic interventions.
www.interventioncentral.org.

Information about item analysis by computers
<http://www.principalproducts.com>

Assessment System Corporation
<http://www.assess.com>

About the PSSA
http://www.pde.state.pa.us/a_and_t/site/default.asp

Achievement Test Publishers:

CTB/McGraw-Hill <http://www.ctb.com>

Riverside Publishing <http://www.riverpub.com>

Harcourt Educational Measurement <http://www.hemweb.com>

Information and Ideas about Portfolios

Provides guidance on the use of student portfolios in instruction and assessment
<http://www.literacynet.org/icans/chapter02/portfolios.html>

Example of scoring criteria that could be used in a student portfolio for history
<http://www.edheritage.org/forms/portfolio.htm>

Electronic portfolios for pre-student teachers
<http://depts.washington.edu/ncate/exhibitroom/worksamples/studentwork.html>

EDSP 477 Assessment of Student Learning – Distance Education/Online Section

EDSP 477 Assessment of Student Learning – Distance Education/Online Section

Class Time: On-Line
Room: Webct
Instructor: Gurmral Rattan, Ph.D.
Office Stouffer 251
Office Phone (724) 357-3787 or 357-2316 (main office)
E-mail [mailto: gurmalra@iup.edu](mailto:gurmalra@iup.edu)

Text:

Linn, R., & Miller, M.D. (2005). Measurement and assessment in teaching, (9thed.). Columbus: Prentice Hall.

Linn, R., & Miller, M.D. (2005). Student exercise manual for Measurement and assessment in teaching, (9thed.). Columbus: Prentice Hall.

Course Description:

The design and scope of this course is to acquaint the student with the major methods and techniques of evaluation that are used to assess the growth, development, and academic progress of pupils in the elementary and secondary schools. Emphasis is placed upon various objectives, qualities, and characteristics desired in methods of evaluation. It is also intended to provide the student with skills in the variety of teacher-made tests and uses of standardized tests.

The course includes elementary statistics concerning graphs, sampling, frequency distribution, measures of central tendency, dispersion, the normal curve, and correlations.

Prerequisite: General psychology, EDSP 102 Educational Psychology

Course Objectives: General course objectives for the student include:

1. Awareness of the role of assessment in teaching
2. Understanding of the role of reliability and validity in both classroom and standardized tests
3. Understanding of the various methods of assessment and their alignment with learning goals
4. Skill building in the development of various teacher-made tests and evaluative procedures
5. Capability in the use of assessment before, during, and after instruction
6. Awareness of the needs of special populations, such as those with disabilities, multicultural populations and those not proficient in English, as they relate to assessment

7. Interpretation and utilization of data/scores provided by standardized tests through the use of elementary statistics
8. Awareness of trends, issues, and relevant research in assessment related to educational reform
9. Awareness of the use of technology in assessment

Course Objective	Course Assessment Measuring Objective
1	Test 1, Test Critique**, Test Construction Project*
2	Exam 1, Class Activities,
3	Test Construction Project*, Exam II
4	Test Construction Project*, Class Activities
5	Test Construction Project*, Exam II
6	Class Activities, Exam II
7 & 8	Exam II, Exam III
9	Website Activity

* Indicates Key Assessment

** For dual level course students

Course Requirements:

1. Students will be responsible for reading the assigned text.
2. The Class Activities portion for this class consists of participating in a Discussion group - see below for details.
3. Three examinations and chapter quizzes will also be administered to assess student mastery of course materials. **No make-up quizzes or exams will be permitted if it is not taken during the designated date and time unless prior approval has been obtained from the instructor.**
4. Each student will be required to construct a Classroom Test in his or her content area. Test questions for the Test Construction project will be comprised of: short answers, true-false, matching, multiple-choice, and essay. The student will need to post an electronic copy in Microsoft Word to Webct. Complete directions for this assignment are posted on Webct.

5. Students will be required to complete the Workbook assigned for this class – see below for details.

6. Students will be able to access their test scores using the My Grades icon on Webct. Scores will be posted for all the chapter quizzes, Test Construction Project, Workbook, and all Exams in addition to the final Course Grade. It will be the student's responsibility to check these grades for accuracy and inform me if there are any errors. **If I do not hear from you *two weeks* before the final exam (exam 3) is scheduled, I will assume the grades are accurate and base your final course grade on those scores.**

Students are also encouraged to change their passwords on their Webct accounts on a regular basis.

7. Final letter grades will be determined by the distribution of scores obtained by the class. Weighting of each assignment is as follows:

<u>Assignment</u>	<u>Weight</u>
Test 1	19%
Test 2	19%
Test 3	19%
Quizzes	10%
Class Activities	14%
Test Construction	19%

Please Note: No additional work will be assigned to improve your final grade. Only the components listed above will be used to calculate your final course grade. If you are receiving a failing grade (D or F) at any time throughout the semester, it will be your responsibility to contact me for assistance. Do NOT wait until the semester is almost over for help – by then, it may be too late. Student who are receiving a passing grade are also welcome to seek assistance at any time.

Assignment of final letter grades will be as follows:

Grade Percentage Score

A	90-100%
B	80-89%
C	70-79%
D	60-69%
F	< 59%

- Late submissions on the Test Construction will be deducted 1/19 points per day late out of the total amount allotted for this project. For example, if the Test Construction project is late 19 days, you will receive a 0 (zero) for the project. This late penalty also applies to submitting your project on the "P" drive. Late submissions for the Workbook will likewise be deducted 1/10 points per day late out of the total amount allotted for this project.

Please keep in mind that Exam 2 and the Test Construction project are due one day apart.. You have enough advanced notice on both these items that there will be no extension provided on either the exam or project. .

- Students caught cheating, plagiarizing or violating other infractions of the Academic Integrity Policy set by IUP will be dealt with accordingly.
- If you are having trouble accessing Webct, I will only be able to check and see if you are on my Webct course list. For any and all other problems, please contact:
 - IUP Student Computing Help Desk** Monday-Friday 10 a.m. to 6PM.
Phone: 724-357-2198.
 - Web Address: <http://www.iup.edu/ats/sts> E-Mail: student-helpdesk@iup.edu or
 - Brian Carothers 724-357-7726.

Additional Information about Course Requirements

A. Test Construction Project

Complete details about the Test Construction Project are posted on Webct. However, as part of completing the **Test Construction Project**, students will also need to

determine the readability of the test that you have developed to ensure that it is appropriate for the grade level of your students. You will need to complete three readability studies:

Readability Studies:

- a. The first one will be the Smog Readability Formula that can be accessed with the following link: <http://www.med.utah.edu/pated/authors/readability.html>.
- b. The second one will consist of the Fry's Readability Formula that can be accessed by the following link:
<http://school.discovery.com/schrockguide/fry/fry.html>. The instructions for calculating the readability formulas are explained in the above links.
- c. The third formula will consist of the Flesch-Kincaid Readability Formula. This formula is available in the Microsoft Word program and is automatically calculated once the parameters on Word are applied.

Please attach details of the calculations using the above formulas with your Test Construction Project. *Samples of the Test Construction project along with the readability formula are presented on Webct. Students are strongly encouraged to use these samples to model their own projects*

Submission Requirements:

Please submit the above paper as requested below:

- 1) Microsoft Word - only submit in Microsoft Word using 12 pt. font in Times New Roman (do not use another word processing program). Submit to Webct and the "Hand-In" folders in the P drive.

It will be your responsibility to learn about accessing the "P" drive. Please link to: "P" drive at ats or "P" drive at acad". Please use Microsoft's Internet Explorer to access these sites (you may have trouble using Netscape or another browser).

B. Class Activities

The Class Activities will consist of a Discussion group and constructing a Student Homepage

1. **Discussion Group** - The "Discussion" format will be used to clarify any questions you have about the textbook or workbook. There will be approximately 10 students per group with a total of three or four groups.

The requirements for the discussion groups are:

- a.) Each person must post one message (use "compose message") that deals with a topical area for the chapter that is being covered. We will limit the

discussion to the chapter(s) that appears on the tentative schedule listed at the bottom of this syllabus. For example, limit your topics to Chp 2 for the week of Jan 23 and Chp 4 for the week of Jan 30, etc. The person initiating the original message must then respond to comments and feedback from others. The original message should be meaningful questions, comments or issues that help clarify the content covered in the textbook or workbook. Each person should initiate a new message for each chapter. The new message can contain overlapping information with a previous message, as long as it adds incremental or unique information to the discussion.

b.) Each person must participate by replying to at least 1-2 other messages that are initiated by your colleagues. To that end, you must read the entire thread of the discussion before you “reply”. When replying to student or instructor initiated messages, it is important that you add something different, new, and relevant to the discussion. It is ok to agree with someone’s comment, but you should try to expand on it in a meaningful way.

As a summary to sections a and b above, you will at a minimum need to initiate one message for each chapter and then reply to 1-2 other messages per chapter that are authored by your colleagues or the instructor. Please keep in mind, however, that you are not limited to these minimum postings, you can post more messages if you have questions or comments.

c.) Please ensure that your posting are current with the chapter being covered for that week – please do not post “outdated” posting unless they are relevant for that particular week and chapter. To that end, you will need to read the chapter in a timely fashion.

d.) Please make a copy of any messages you post (compose or reply). In the past, student’s postings did not appear in the Discussion group due to problems with their home computer. As a result, students lost points since no records were kept of their postings. To that end, copy all of your posting to a separate MS Word file and it available in the event of a computer glitch.

e.) you should check the discussion board at least once a day for any postings. I will monitor the discussion board and respond to questions or comments as necessary.

f.) please, please, please do not send me your questions using the webct mail unless they are of a personal nature, e.g., grades, illnesses, etc. The advantage of having a discussion board is for everyone to benefit from your questions and the responses that the instructor or another student provides. If you have read the textbook and you still have a question, then they are very legitimate and worthy questions. If you have not read the book, then I suggest you not post any messages, questions or replies until you have done so. Remember, if

you have done the readings, there are no dumb questions – somebody out there probably has the same or similar question.

Group Assignment: I will be assigning people to one of three or four groups. Once I have assigned you to a group, an **active link** for that designated group should appear in your Discussion tool. Please click on that link to compose or reply to messages. To fulfill the requirements for this component of the course, you must post to the assigned group.

Point Value: The discussion component will be assigned 170 points, 10 points for each chapter. Everybody will start off with 10 points for each chapter but points will be removed if you are: a.) do not initiate a meaningful message, b.) do not provide meaningful responses to follow-up questions about your message, c.) do not actively participate in replying to messages.

Special Notes:

- Given that the first week of class starts in the middle of the week, everyone will be awarded 10 crisp points for participation for Chp 2. However, please feel free to post for Chapter 2.

2. **Student Homepages** – Everyone should post something to the student home page with either links to their own webpage or some other relevant information about themselves. Since most of you may not know each other, this would be an opportunity to share something about yourself that you feel comfortable and would not be embarrassed in posting for others to see. This component will be worth 10 points if it is satisfactorily completed. My webpage is posted at the following: <http://www.coe.iup.edu/rattan/>

C. **Student Workbook**

The Workbook assignment forms a part of the Class Activities category. Both the **objectives and subjective** components for each chapter assigned in class must be completed in full in order to obtain points allocated for this assignment. **These workbooks should be mailed in to the instructor on the last class day** – additional details will be provided.

The chapter questions will need to be completed thoroughly (i.e., objective and subjective sections) in order to get full points. Grading will be based on the following criteria:

- a. 9-10 points – all questions completed thoroughly

- b. 7-8 points – all questions completed by not thoroughly
- c. 5-6 points – up to 20% of the questions not answered thoroughly
- d. 3-4 points – 21% - 50% of the questions not answered thoroughly
- e. 0-2 points – more than 50% of the questions not answered thoroughly

Point will be deducted if your workbook is turned in late from the date given above. These “late” points’ deductions will be in addition to points deducted for incomplete work as noted on the above scoring criteria.

Please complete the workbook for the top and/or bottom section for each chapter as noted below.

Chapter	Top Section	Bottom Section
2	All	All
3	All	All
4	All	All
5	All	All
6	All	All
7	A thru D	A thru C
8	A thru D	A thru D
9	A thru B	A
10	A, E	A, E
11	A, D, E	A
12	D	none
13	A thru E	B
14	A, C, D	A, D
15	All	All
16	A thru E	A, B, E
17	A, C, E	A, C
18	A, D, E	A, D, E
19	A thru E	A thru E

Assessment of Student Learning: EDSP 477_DE
Tentative Course Schedule

Date	Assigned Reading	Chapter(s)
Jan 18	Course Overview	
Jan 23	Role of Evaluation	Chp 2
Jan 25, 30 Feb 01	Role of Evaluation /Validity	Chps 2, 4
Feb 06,08,13	Validity/Reliability	Chps 4, 5
Feb 15	Test Selection, Administration,	Chp 18
Feb 20	Instructional Objectives	Chp 3
Feb 22	Test #1	
Feb 27	Review Test #1/Planning Classroom Tests, <u>Ethics & Procedures in Testing</u> Test Construction Project - Instructions	Chp 6
Mar 01,06	Planning/Constructing: Simple Forms, Writing Paper & Pencil Assessments	Chps 6,7
Mar 08,20	Constructing: Multiple Choice, Writing Paper & Pencil Assessments	Chps 7, 8
Mar 22	Measuring: Interpretive/Essay, Writing Paper & Pencil Assessments	Chps 9, 10
Mar 27, 29 Apr 03	Essay/ Measuring: Performance	Chps 10,11
Apr 05	Assembling, Administration, Special Populations Construction Project Due	Chp 14
Apr 10	Test #2	
Apr 12	Review Test #2/Interpreting Test Scores	Chp 19
Apr 17	Interpreting Test Scores /Portfolios	Chps 19, 12
Apr 19	Assessment Procedure,	Chp 13
Apr 24	Curriculum Based Assessments /RTI /Achievement	Chps 16
Apr 26	Achievement/Aptitude	Chps 16,17
Apr 01	Grading and Reporting, Current Trends & Issues	Chps 15
May 08	Exam 3, Tues May 8	

V. Bibliography

- Airasian, P. (2005). *Classroom assessment: Concepts and applications*. New York: McGraw-Hill.
- American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Anderson, L. W. (2003). *Classroom assessment: Enhancing the quality of teacher decision making*. Mahwah, NJ: Lawrence Erlbaum.
- Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13-18.
- Arter, J., & Busick, K. U. (2001). *Practice with student involved classroom assessment*. Upper Saddle River, NJ: Merrill Prentice Hall. This is a workbook with many examples and ideas for creative assessment approaches.
- Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests. *Educational Research*, 44(3), 279-288.
- Burger, S. E., & Burger, D. L. (1994). Determining the validity of performance-based assessment. *Educational Measurement: Issues and Practice*, 31(1), 9-15.
- Chase, C. I. (1999). *Contemporary assessment for educators*. New York: Longman.
- Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning, *Educational Leadership* 60(1), 40-44.
- Frisbie, D.A., Miranda, D. U., & Baker, K.K. (1993). An evaluation of elementary textbook tests as classroom assessment tools. *Applied Measurement in Education*, 6(1), 21-36.
- Gallagher, J. D. (2000). *Classroom assessment for teachers*. Upper Saddle River, NJ: Merrill.
- Gronlund, N. (2000). *How to write and use instructional objectives*. Upper Saddle River, NJ: Prentice Hall. Provides assistance in writing clear objectives essential to the instructional and assessment process.
- Guskey, T. R. (2003). How classroom assessments improve learning. *Educational Leadership*, 60(5), 6-11.
- Herrnstein, R.J. & Murrary, C. (1994) *The bell curve*. New York: Free Press.
A controversial look at the use of assessment, intellectual ability, and achievement in American society. Guaranteed to stimulate your thinking!
- Hoy, C., & Gregg, N. (1995). *Assessment: The special educator's role*. Pacific Grove, CA: Brooks/Cole.
- Hughes, S. (1993). What is alternative/authentic assessment and how does it impact special education? *Educational Horizons*, 72(3), 28-35.
- Klitz, G. (2002). Involving students in the improvement process through portfolios. *Journal of School Improvement*, 3(1), 54-56.
- Lane, S., Mei, L., Ankenmann, R.D., & Stone, C.A. (1996). Generalizability and validity of mathematics performance assessment, *Journal of Educational Measurement*, 33(1), 71-92.

- Lescher, M.L. (1995). *Portfolios: Assessing learning in the primary grades*. Washington, D.C: NEA Professional Library.
- Lukhele, R. Thissen, D., & Wainer, H. (1994). On the relative value of multiple choice, constructed response, and examinee selected items on two achievement tests. *Journal of Educational Measurement*, 31(3), 243-250.
- Madhavi, J., Bursuck, W., Havekost, D.M., Epstein, M.H., & Palloway, E.A. (1994). School district testing policies and students with disabilities: A national survey. *School Psychology Review*, 23(4), 694-703.
- McMillan, J.H. (2004). *Classroom assessment: Principles and practice for effective instruction*. Boston: Allyn & Bacon. Good chapter on effective oral questioning and appropriate teacher responses.
- Mehrens, W. A., Popahm, W. J., & Ryan, J. M. (1998). How to prepare students for performance assessments. *Educational Measurement: Issues and Practice*, 17(1), 18-22.
- Montgomery, K. (2001). *Authentic assessment: A guide for elementary teachers*. New York: Longman.
- Naizer, G.L. (1997). Validity and reliability issues of performance-portfolio assessment. *Action in Teacher Education*, 18(4), 1-9.
- Overton, T. (2002). *Assessment in special education: An applied approach*. New York: Merrill.
- Penta, M. Q. (2002). Student portfolios in a standardized world. *Kappa Delta Pi Record*, 38(2), 77-81.
- Popham, W.J. (2005). *Classroom assessment: What teachers need to know*. Boston: Allyn & Bacon. Easy to read and very teacher friendly.
- Sachs, P. (1999) *Standardized minds*. Cambridge, MA: Persuis.
- The President's Advisory Commission on Educational Excellence for Hispanic Americans: *Testing Hispanic students in the United States*. (2000). Washington, DC: Author.
- Salvia, J. & Ysseldyke, J. (2003). *Assessment in special and inclusive education*. Boston: Houghton Mifflin.
- Seidel, S., Walters, J., Kirby, E., Olf, N., Powell, D., Scripp, L., & Veenema, S. (1997). *Portfolio practices: Thinking through the assessment of children's work*. Washington, DC: NEA Professional Library. Contains a good deal of practical advice for laying the groundwork in the development of portfolio use.
- Shinn, M.R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford. An in-depth look at this procedure.
- Siegle, D. (2002). Creating a living portfolio: Documenting student growth with electronic portfolios. *Gifted Child Today*, 23(3), 60-64.
- Spinelli, C. G. (2005). *Classroom assessment for students in special and general education*. Upper Saddle River, NJ: Prentice Hall.
- Stiggins, R. J. (2005). *Student-involved assessment for learning*. Upper Saddle River, NJ: Merrill Prentice Hall. An emphasis on assessments for classroom use with concentration on the appropriate use of performance-based assessment.
- Trice, A. D. (2000). *A handbook of classroom assessment*. New York: Longman. This is basically a reworking of what is in your text, but sometimes it helps to read it in more than one place and presented a bit differently.

U.S. Department of Education. (1994). *Changing education: Resources for systemic reform*. (PIP 94-1509). Washington, DC: U.S. Government Printing Office. Short articles of interest on many of the topics discussed in this course with special emphasis on issues of importance in the reform movement.

Weber, E. (1999). *Student assessment that works: A practical approach*. Boston: Allyn & Bacon. Some good ideas on how to collaborate with other educators and work with parents around the assessment issue.

Relevant Websites

www.aimsweb.com

There is a small fee to use this site but you can download materials for curriculum-based assessments.

A site to assist with the readability assessment of your text.
www.med.utah.edu/pated.authors/readability.html

Another site addressing readability levels.
www.school.discovery.com/schrockguide/fry/fry.html

California Reading Comprehension Standards
<http://www.cde.ca.gov/cdepress/Eng-Lang-Dev-Stnd.pdf>

Florida Math Standards
<http://www.firn.edu/doe/curric/prek12/pdf/math6.pdf>

History and Social Science Standards
<http://www.pen.kqw.va.us/VDOE/supertendent/Sols/historysol2001.pdf>

Center for Research on Evaluation, Standards, and Student Testing
<http://www.cse.ucla.edu>

Hands-on science performance assessment task
<http://www.ed.gov/NCES/naep>

A curriculum-based measurement manual for teachers and learn about a variety of academic interventions.
www.interventioncentral.org.

Information about item analysis by computers
<http://www.principalproducts.com>

Assessment System Corporation
<http://www.assess.com>

About the PSSA
http://www.pde.state.pa.us/a_and_t/site/default.asp

Achievement Test Publishers:

CTB/McGraw-Hill <http://www.ctb.com>
Riverside Publishing <http://www.riverpub.com>
Harcourt Educational Measurement <http://www.hemweb.com>

Information and Ideas about Portfolios

Provides guidance on the use of student portfolios in instruction and assessment
<http://www.literacynet.org/icans/chapter02/portfolios.html>

Example of scoring criteria that could be used in a student portfolio for history
<http://www.edheritage.org/forms/portfolio.htm>

Electronic portfolios for pre-student teachers
<http://depts.washington.edu/ncate/exhibitroom/worksamples/studentwork.html>

Outline of Sample Lesson: Linn et al., Chp 5 – Reliability and Other Desired Characteristics

Includes:

1. Webct power point – Instructor notes
2. Webct power point – Introduction to Reliability and Validity by Dennison S. Bholra, Ph.D. – a graphic bullseye presentation of reliability and validity
3. Student workbook activities
4. Discussion group – part of class activities
5. Webct quiz

Sample Lesson – Chp 5, Sequence of Student Activities

The lesson for chapter 5 is related to the contents in the text and is associated with objectives 2 and 7 in the syllabus. The content in the chapter focuses on the characteristics and properties of reliability such as: definition, methods to estimating reliability coefficients, standard error of measurement, its relationship to validity, how to interpret reliability coefficients in the context of test scores, etc.. The sequence for this lesson will consist of the following activities:

1. Read the chapter in the text “Reliability and other desired characteristics”
2. Access the instructor power points from Webct (Class Notes)
3. Access the supplementary power points by Dennison S. Bhola, Ph.D. from Webct (Class Notes)
4. Complete the student workbook for this chapter as described above (all objective and subjective items for Chp 5)
5. Participate in the Discussion activity by originating 1 post and replying to 1-2 posts from other students) (Discussion)
6. Complete the quiz for this chapter that can be found on Webct (Quizzes/Exams)

Attached are the power point presentations as listed above along with general cover page from Webct.



EDSP 477 Assessment of Student Learning



Syllabus



Class Notes



Discussion



Quizzes/Exams



TestConstruction



My Grades



Practice Questions



Study Guide



bookmarks
Office Hours



Calendar



mail
Mail



NCATE
(Hidden)

Procedure

a. If we administer HI to $n = 100$ at one week and again to the same group two weeks later, the correlation between the scores is called a reliability coefficient and ranges between 0 to +1.0.

Outcome of Procedure

- If the same person answered all the questions in exactly the same way on both testing sessions:
 1. What would the reliability coefficient be?
 2. Would we be certain that the answers were truthful? Or valid?
 3. Therefore, a test could have high reliability but not necessarily be valid. That is, reliability is necessary but not a sufficient condition for validity.

Definition

▪ Reliability is the extent to which a test can consistently measure whatever it purports to assess.

Example

- We want to develop a test to measure the trait of honesty e.g. Honesty Inventory (HI), therefore we write a series of questions:
 1. Do you ever deliberately distort the truth?
 2. If you have the opportunity, do you ever cheat on a test?
 3. If a clerk makes a mistake and gives you too much change, do you ignore the error?

Chapter Five

Reliability

Gurnal Rattan, Ph.D.

Writing Activity

1. What type of reliability is more important when you want to ensure that you do a good job measuring the entire domain of skill areas. PIs: elaborate.
2. Why would test-retest yield a higher reliability coefficient than equivalent forms?
3. Why would you prefer a test with a smaller SEM. Discuss what advantages this would have when assigning letter grades to students.
4. What is the Spearman-Brown Prophecy formula and when/why would you apply it?
5. Why would you not want to use a split-half method to establish reliability for a speeded test?

Other Characteristics of HI

1. If HI contained a representative sample of honest behavior, then it would have a good content validity.
2. However, if HI did not disarm or prevent examinees from faking responses, it would lack construct validity (i.e. scores would not assess honesty). Therefore, we could not use it to predict future behavior (i.e. predictive validity)
3. Alternatively, for HI to have construct validity, then other sources called external criteria (e.g. aunt, uncle, relative, teacher) could be used to cross-validate the individual's responses.

Types of Reliability Coefficients

Five types of reliability

Name	What is Measured	Description
Reliability (test-retest)	Reliability of scores over time	Give one group the same test at two different times, and correlate the two scores
Equivalence (alternate forms)	Relationship between two versions of a test intended to be equivalent	Give two alternative forms of the test to a single group within a short period of time, and correlate the two scores
Equivalence and stability	Relationship between equivalent versions of a test given at two different times	Give two alternative tests to a group at two different times (delayed administration), and correlate the scores
Internal consistency	The extent to which the items in a test are similar to one another in content	Give tests to one group, and apply split-half (Kuder-Richardson, or Cronbach's alpha to estimate the internal consistency of the test items
Score-retest	The extent to which independent scores of a single score over time agree on the scoring of an open-ended test	Give copies of a set of tests to independent scores of a single score at different times, and correlate or compute the percentage of score agreement

Test-retest Method

1. Here, the reliability is based on correlating the scores on a test with scores on the same test administered at a later time to the same individuals (usually 2 weeks later). These reliability coefficients are termed "coefficients of stability." This procedure of establishing reliability is done by the publisher.

Equivalent Forms, Alternate Forms or Parallel-Forms Reliability

- Here, the reliability coefficient is determined by administering two forms of the test. The test items in each form are different but each measures the same underlying construct.
1. Equivalent form reliability is used when you want to measure growth or change over a short period of time. For example, has a student made progress with a new reading program within the last two months?

Equivalent Forms, Alternate Forms or Parallel-Forms Reliability cont'd

- Two types of Administrations:
1. Two different forms are administered shortly after each other. That is, *Form A* is administered in a.m. and *Form B* in the p.m. in a counter balanced fashion in order to eliminate the "order effect". The resulting reliability coefficient is a measure of "equivalence".
 2. Same as the above but *Form B* is administered about 2 weeks after *Form A*. The resulting reliability coefficient is a measure of "equivalence and stability".

Split-Half Reliability

- Split-half Method- here, the reliability is established by correlating all odd numbered items with even number items.
 - All odd-numbered items (e.g. 1,3,5...)
 - All even-numbered items (e.g. 2,4,6...)

Split-Half Reliability cont'd

1. Remember, the test is only administered once. The resulting correlation is a measure of "internal consistency".
2. This type of reliability is referred to as a Split-half method of reliability.
3. A Spearman Brown Prophecy Formula is used to correct for attenuation (shortened version). Why?
4. Formula: $r = (2 * r_{xx}) / (1 + r_{xx})$ If $r_{xx} = .6$, then
 - 1. $(2 * .6) / (1 + .6) = .75$

Kuder-Richardson Method

- Here, the method of establishing reliability is based upon correlating the proportion of persons passing each item and the standard deviation of the total scores. The resulting correlation is a measure of "internal consistency."

When to Use Split-half and Internal Consistency

- Yes
 1. Only for power tests, i.e. students have sufficient time to complete all items.
- No
 2. When used on speeded tests, results may yield spuriously high reliability scores (or artificially high reliability). Instead, use test-retest or parallel-forms.

Standard Error of Measurement **SE_M**

- Discrepancies between obtained scores and true scores is referred to as a standard error of measurement (SE_M)
 1. That is, if the same test is given repeatedly without any further instruction, variability in scores is called SE_M . The resulting scores will approximate a normal distribution.
 2. The SE_M and the reliability coefficient are inversely correlated. The higher the SE_M , the lower the reliability coefficient and vice versa.

Question

- Under what conditions will the SE_M be equal to zero?
- If the reliability coefficient is zero, what will the SE_M be?

Factors Influencing Reliability Measures

1. Test length: generally, the longer the test, the more reliable it becomes until it reaches a plateau.
2. Dispersion of Scores: the larger the spread of scores, the more variability in scores and therefore, the higher the reliability.

Reliability Demands and Nature of the Decision

- High reliability is required when the:
 1. Decision is important
 2. Decision is final
 3. Decision is relevant
 4. Decision is irreversible
 5. Decision is not confirmable by other data
 6. Decision concerns individuals
 7. Decision has lasting consequences
 - Example: select or reject college applicants

Reliability Demands and Nature of the Decision cont'd

- Low reliability is tolerable when the:
 1. Decision is of minor importance
 2. Decision making is in early stages
 3. Decision is reversible
 4. Decision is confirmable by other data
 5. Decision concerns groups
 6. Decision has temporary effects.
 - Example: Whether to review a classroom lesson

Five types of reliability

Name	What is Measured	Description
Stability (test-retest)	Stability of scores over time	Give one group the same test at two different times, and correlate the two scores
Equivalence (alternate forms)	Relationship between two versions of a test intended to be equivalent	Give two alternative forms of the test to a single group in close a short period of time, and correlate the two scores
Equivalence and stability	Relationship between equivalent versions of a test given at two different times	Give two alternate tests to a group of two different times (delayed administration), and correlate the scores
Internal consistency	The extent to which the items on a test are similar to one another in content	Give tests to one group, and apply split-half, Kuder-Richardson, or Cronbach's alpha to estimate the internal consistency of the test items
Score-retest	The extent to which independent scores or a single score over time agree on the scoring of an operational test	Give copies of a set of tests to independent scorers or a single scorer at different times, and correlate or compute the percentage of score agreement

Forms of Validity

Form	Method	Purpose
Content validity	Compare content of the test to the domain being measured.	To what extent does the test represent the general domain of interest? Use a table of specifications or ask experts in the field.
Criterion-related validity	Correlate scores from one instrument to scores on a criterion measure, either at the same (concurrent) or different (predictive) time.	To what extent does this test correlate highly with another test (concurrent) or predict some future performance (predictive).
Construct validity	Amass convergent, divergent, and content-related evidence to determine that the presumed construct is what is being measured.	To what extent does this test reflect the construct it is intended to measure?
Consequential validity	Observe and determine whether the test has adverse consequences for test takers or users.	To what extent does the test create harmful consequences for the test taker?

Introduction to Reliability and Validity

Dennison S. Bhola, Ph.D.

Consider the dart boards shown below.



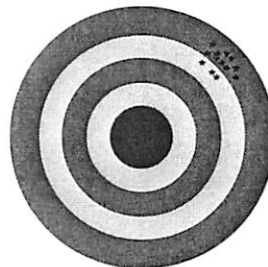
2

The marks you see on the targets were left by darts thrown at them.



3

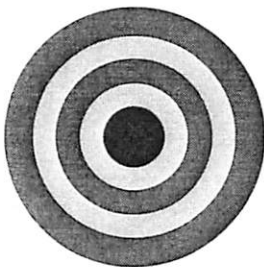
If the goal was to hit the "Bullseye" with each dart...



Then the results were consistent but off-target

4

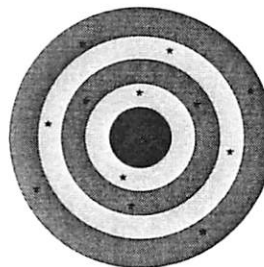
If the goal was to hit the "Bullseye" with each dart...



Then the results were both consistent and accurate

5

If the goal was to hit the "Bullseye" with each dart...



Then the results were neither consistent nor accurate


6

Reliability and Validity

- Reliability has to do with consistency
- Validity has to do with accuracy

7

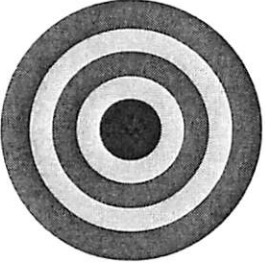
If the goal was to hit the "Bullseye" with each dart...



We have reliability but not validity

8

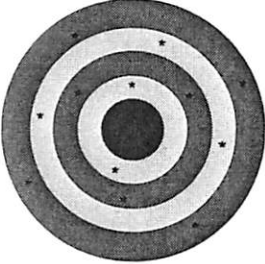
If the goal was to hit the "Bullseye" with each dart...



We have both reliability and validity

9

If the goal was to hit the "Bullseye" with each dart...



We have neither reliability nor validity


10

Reliability and Validity

- To have validity we must first have reliability
 - i.e. reliability is a prerequisite for validity
- Reliability is a necessary but not sufficient condition for validity

11

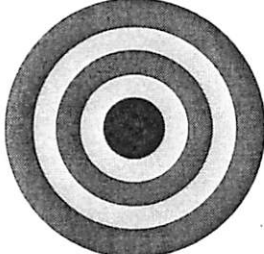
Reliability is a necessary but not sufficient condition for validity



We can have reliability without having validity

12

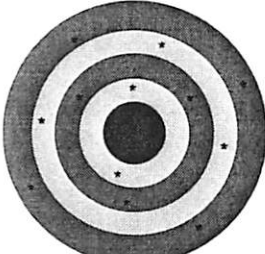
In order to be valid the measurement must be reliable



We cannot have validity without first having reliability

13

Reliability is a prerequisite for validity



Without reliability, we cannot have validity

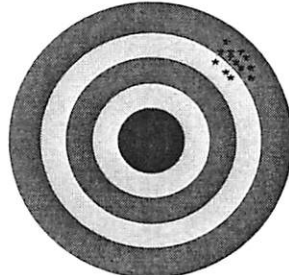
14

Summary

- To have validity we must have reliability
 - Reliability is a prerequisite for validity.
- Reliability is a necessary but not sufficient condition for validity.

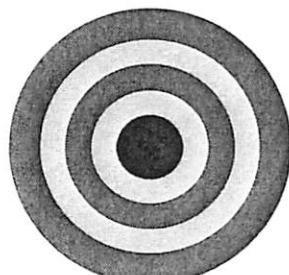
15

Measurement reliable but not valid



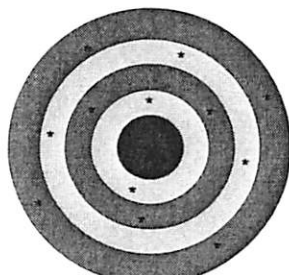
16

Measurement reliable and valid

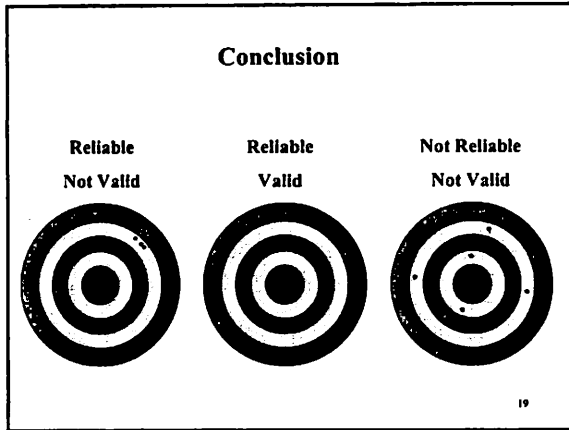


17

Measurement neither reliable nor valid



18



Reliability and Validity

- Reliability has to do with consistency of measurement.
- Validity has to do with accuracy of measurement.

20

Reference

Linn, R.L. & Gronlund, N. E. (2005). Measurement and Assessment in Teaching, 8th Ed. New Jersey: Prentice-Hall, Inc.

See diagram displayed on page 70 of the reference textbook.

Next Session

- How do we develop tests to have validity?
- Are there different types of reliability?
- Are there different types of validity?
- How do I know which is most important for my situation?

22