

## **Performance Assessment Design & Analysis**

EPSY 8224 Spring 2011

Thursdays 4:40pm – 7:20pm

Instructor: Leah McGuire

Office: 157 Education Sciences

Phone: 612-624-7622

Email: [lwmcguir@umn.edu](mailto:lwmcguir@umn.edu)

Office Hours:

Instructor: Michael C. Rodriguez

Office: 163 Education Sciences

Phone: 612-624-4324

Email: [mcrdz@umn.edu](mailto:mcrdz@umn.edu)

Office Hours: T/Th 3:00-4:00

Classroom: Education Sciences Room 10

Web: <http://www.edmeasurement.net/performance>

### ***The Course:***

Performance assessment is a tool that has been in use for centuries. Evidence suggests that as early as 2200 BC, Chinese emperors set up civil service exams that were performance based trials. It wasn't until the development of multiple-choice items during WWI and the use of the SAT in the 1920s that objective assessments became the standard for use in large-scale settings in the USA. However, classroom teachers have continued to employ performance-based assessments in both formative and summative evaluation of students. In the last decade or so, performance activities have resurfaced in large-scale settings, particularly in statewide testing programs. The psychometric (technical) qualities of the results have not met current professional standards for making high-stakes decisions about individuals. There has been limited success in the use of performance assessments in some large-scale programs, but the evidence has not been overwhelming in support of these activities over objective measures (e.g., multiple choice tests).

There are many potential uses for performance assessment and we have learned a great deal about their benefits and costs. The literature is rich in this area and the skills needed to develop quality assessments have been demonstrated.

The course will present the issues related to the conceptualization, design, implementation, and analysis of performance assessments as employed in both small-scale (e.g., classrooms) and large-scale (e.g., statewide and national testing programs) settings. The design stage alone is wide-ranging in terms of building test specifications, training task developers, designing tasks and scoring criteria, reviewing tasks for sensitivity and bias, piloting tasks, analyzing pilot results, and finally selecting tasks. We will examine several existing performance assessment systems, including the Advanced Placement assessments, the NBPTS teacher certification portfolio system and school leaders certification assessments, performance based assessment systems for students with severe cognitive/physical impairments, and others.

### ***Primary Course Objectives:***

This course will provide students with the background, lessons learned, and practical exposure to techniques for the development and implementation of performance assessments at various levels for various purposes, with the skills to critique such uses. Students will engage in sophisticated readings, complete small projects designing performance assessment tasks and rubrics under a typical set of guidelines, and critique real applications of performance assessments.

During the data analyses sessions, GENOVA software will be used for demonstration and the free downloadable software will be provided online at the course website. The course will also employ several interactive web sites designed to provide students with the opportunity to build their familiarity with every aspect of performance assessment at several levels, including early elementary education, K-12, higher education, and professional assessments.

Specific objectives for each session will be presented at each session. Generally, the student will be able to

- ◆ Develop tasks and rubrics for measuring academic and professional skills;
- ◆ Evaluate performance assessment tasks, rubrics, and related materials;
- ◆ Interpret performance assessment results from a variety of sources; and
- ◆ Apply principles of measurement for responsible test use.

### ***Requirements:***

Students will be expected to read the assigned materials prior to class to engage in class discussions. Assignments are provided for reflection and to gain experience in the design and evaluation of performance assessments in a variety of settings.

1. Complete performance assessment/portfolio blueprint
  - Due week 4 (10% of grade)
2. Identify existing assessment and critique
  - Due week 9 (20%)
3. Data-based assignment
  - Due week 12 (20%)
4. Final project
  - Due finals week, by May 13 (50%)

There will be one multi-part project. This project will include several stages of the development of a performance assessment in the setting of the student's choice, including specification of (a) assessment purpose, (b) domain description, (c) identification of the aspects of performance that are relevant and applicable to assessment, (d) detailed specifications for task development or sample tasks, (e) sample rubrics, and (f) initial analysis and evaluation plans. This project will equal ½ of the total grade. Each phase of this project is brief and in whole will constitute a complete assessment plan. Details will be provided in class.

We will take as much time as is necessary to complete understanding of the material before we move on. There is time during each session to allow for extended time on a topic. The material is

cumulative to a certain extent so it is imperative that you keep up with the workload. You are encouraged to work with a partner, as long as this works for you, and to utilize office hours to complete your understanding of the material from the moment you begin to feel uncertain. It is during these moments of uncertainty that a great deal of learning can occur.

**Attendance.** The course is intended to be very interactive. Attendance in every session is expected, but there are no grade-based consequences for missing class. However, out of professional courtesy, please notify the instructors if you plan to be absent for one or more sessions.

**Prerequisites.** This course is intended to be an advanced measurement course. As such, students who have completed an introductory measurement course (e.g., EPSY 5221) and a year-long graduate statistics sequence (e.g., EPSY 8261/8262) would be well prepared. Ideally, students will also have completed an advanced measurement course (e.g., EPSY 8221 or EPSY 8222). However, students with appropriate backgrounds and experience might consider taking the course and should consult with the instructors prior to enrolling. There are no strictly-enforced prerequisites for this course.

### Evaluation of Student Performance

The final grade in the course will be based on performance on assignments and the major project. You will have opportunities to resubmit work that is returned as incomplete. The expectation is that only complete work will be graded. Grades will be based on the standard University guidelines for letter-grade performance based on corresponding performance on assignments.

Letter Grade	Percentage
A	95-100%
A-	90-94.99%
B+	85-89.99%
B	80-84.99%
B-	75-79.99%
C+	70-74.99%

**Evaluation** ([http://www1.umn.edu/usenate/u\\_senate/usenpol.html](http://www1.umn.edu/usenate/u_senate/usenpol.html)):

- A -- achievement that is outstanding relative to the level necessary to meet course requirements.
- B -- achievement that is significantly above the level necessary to meet course requirements.
- C -- achievement that meets the course requirements in every respect.
- D -- achievement that is worthy of credit even though it fails to meet fully the course requirements.
- S -- achievement that is satisfactory, which is equivalent to a C- or better (achievement required for an S is at the discretion of the instructor but may be no lower than a C-).
- F (or N) -- Represents failure (or no credit) and signifies that the work was either (1) completed but at a level of achievement that is not worthy of credit or (2) was not completed and there was no agreement between the instructor and the student that the student would be awarded an I (see also I)
- I -- (Incomplete) Assigned at the discretion of the instructor when, due to extraordinary circumstances, e.g., hospitalization, a student is prevented from completing the work of the course on time. Requires a written agreement between instructor and student.

### *Late Work and Incompletes*

Late work will be accepted only if the instructor is previously notified – no points will be deducted for late work that was previously agreed upon. It is up to you to stay on track. An incomplete (“I”) will be assigned only if agreed to prior to the last week of class. If course work is incomplete and no prior notification has been given, the grade based on points obtained will be awarded. No options will be given at that point to submit incomplete work.

### *Returning Papers, Exams, and Projects*

We will try to return completed work by the class session following submissions. If, at the end of the semester, you would like to receive remaining work through U.S. mail, submit a self-addressed stamped envelope. Otherwise, you may pick up final projects once grades are submitted in the Educational Psychology Department office in the Education Sciences Building with an ID. Papers will be available there until the second month of the fall semester. Uncollected papers will be destroyed just prior to spring semester the following year.

### *How to Access Your Grades*

Go to OneStop for Students (<http://onestop.umn.edu/onestop/>), click on Grades & Transcripts; on the right side under Quick Links, click on Grades/Unofficial transcript.

The Senate affirms the standard (first adopted by the University Senate on February 16, 1922, and reaffirmed 1993) that one semester credit is to represent, for the average University of Minnesota undergraduate student, three hours of academic work per week (including lectures, laboratories, recitations, discussion groups, field work, study, and so on), or approximately 45 hours of work over the course of an enrollment period. Expectations of faculty and students will be made clear. It is expected that the academic work required of graduate and professional students will exceed three hours per credit per week or 45 hours per semester.

### *University of Minnesota Policy on Scholastic Misconduct*

Scholastic misconduct is broadly defined as "any act that violates the rights of another student in academic work or that involves misrepresentation of your own work." **Scholastic dishonesty** includes, (but is not necessarily limited to): cheating on assignments or examinations; plagiarizing, which means misrepresenting as you own work any part of work done by another; submitting the same paper, or substantially similar papers, to meet the requirements of more than one course without the approval and consent of all instructors concerned; depriving another student of necessary course materials; or interfering with another student's work. Scholastic dishonesty in any portion of the academic work for a course shall be grounds for awarding a grade of F or N for the entire course.

### **Citations/Avoiding Plagiarism**

<http://swc.umn.edu/writinglinks.htm>

From the UM's Student Writing Center. Scroll down to find "Citations/Avoiding Plagiarism."

### **Avoiding Plagiarism**

<http://ollie.dcccd.edu/library/module4/M4-VII/plagar.htm>

Details what plagiarism is, illustrates with examples of “what not to do”, and links to other sites that “further define plagiarism.”

### **The Correct Use of Borrowed Information**

<http://www.winthrop.edu/wcenter/WritingProgram/plagiar.htm>

As the title suggests, this helpful site from the Winthrop University Writing Center features various examples of correct in-text citation for research papers. Writers are shown the proper citation formats for paraphrases, direct quotes, summaries, etc.

### **Technology**

Technology is becoming increasingly important in education and we will pursue learning with the aid of technology in several ways. Students will be allowed to submit projects electronically. Students are encouraged to investigate performance assessment related web sites to support their reading and project work. A page of links to additional readings and resources is provided at the class website.

### **Diversity**

The College of Education and Human Development is committed to have every course contribute to our understanding of diversity, including but not limited to: age, creed, disability, ethnicity, gender, global perspectives, international background, language background, learning differences, marital status, multicultural perspectives, national origin, public assistance status, race, religion, sex, sexual orientation, and veteran status. Each of these characteristics plays a role in educational and psychological measurement and research. They are factors that contribute to individual and group differences -- they (may) affect the constructs we set out to measure and the way we interpret and report test results. These issues will be addressed throughout the course and will be used as topics of debate and considerations of selecting specific assessment tasks and procedures to understand the role of group differences in evaluating test results.

It is the University Policy to provide, on a flexible and individualized basis, reasonable accommodations to students who have disabilities that may affect their ability to participate in course activities or to meet course requirements. Students with disabilities are encouraged to contact me when possible to discuss their individual needs for accommodations.

### **University Policies**

#### *Statement on accommodations*

It is University policy to provide, on a flexible and individualized basis, reasonable accommodations to students who have documented disability conditions (e.g., physical, learning, psychiatric, vision, hearing, or systemic) that may affect their ability to participate in course activities or to meet course requirements. Students with disabilities are encouraged to contact Disability Services and their instructors for a confidential discussion of their individual need for academic accommodations. Disability Services is located in Suite 180 McNamara Alumni Center, 200 Oak Street. Staff can be reached by calling 612-626-1333 voice or TTY.

### *Statement on mental health services*

As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance or reduce a student's ability to participate in daily activities. University of Minnesota services are available to assist you with addressing these and other concerns you may be experiencing. You can learn more about the broad range of confidential mental health services available on campus via the Student Mental Health Website at <http://www.mentalhealth.umn.edu>.

### *Statement on classroom conduct*

<http://www1.umn.edu/usenate/policies/classexpectguide.html> and [http://www1.umn.edu/regents/policies/academic/Student\\_Conduct\\_Code.pdf](http://www1.umn.edu/regents/policies/academic/Student_Conduct_Code.pdf)

### *Statement on academic misconduct*

[http://www1.umn.edu/regents/policies/humanresources/Academic\\_Misconduct.pdf](http://www1.umn.edu/regents/policies/humanresources/Academic_Misconduct.pdf)

Scholastic misconduct is broadly defined as "any act that violates the rights of another student in academic work or that involves misrepresentation of your own work." Scholastic dishonesty includes, (but is not necessarily limited to): cheating on assignments or examinations; plagiarizing, which means misrepresenting as you own work any part of work done by another; submitting the same paper, or substantially similar papers, to meet the requirements of more than one course without the approval and consent of all instructors concerned; depriving another student of necessary course materials; or interfering with another student's work.

### *Statement regarding sexual harassment*

<http://www1.umn.edu/regents/policies/humanresources/SexHarassment.pdf>

"Sexual harassment" means unwelcome sexual advances, requests for sexual favors, and/or other verbal or physical conduct of a sexual nature when: (1) submission to such conduct is made either explicitly or implicitly a term or condition of an individual's employment or academic advancement in any University activity or program; (2) submission to or rejection of such conduct by an individual is used as the basis of employment or academic decisions affecting this individual in any University activity or program; or (3) such conduct has the purpose or effect of unreasonably interfering with an individual's work or academic performance or creating an intimidating, hostile, or offensive working or academic environment in any University activity or program. University policy prohibits sexual harassment. Complaints about sexual harassment should be reported to the University Office of Equal Opportunity, 419 Morrill Hall.

### **Support Services**

Add text here. (Go to <http://www1.umn.edu/ohr/teachlearn/syllabus/specialserv.html> to see some possible support services for your students. Also check within your department for resources.)

This publication/material is available in alternative formats upon request. Please see the course website: [www.edmeasurement.net](http://www.edmeasurement.net)

The University of Minnesota is an equal opportunity employer and educator.

*QME Mission Statement*

**To prepare students to become cutting-edge professionals in educational measurement, evaluation, statistics, and statistics education, through excellence in teaching, research, and service; and through investigating and developing research methodology in education.**

*Department of Educational Psychology Mission Statement*

The broad mission of the Department of Educational Psychology is to:

- create and disseminate new knowledge about successful educational practices.
- preserve the established and time honored practices of the past.
- promote the welfare and development of all students from all abilities and backgrounds.

Within this mission, the department's priorities are to:

- prepare graduate students to take leadership in research, teaching, professional practice, and service.
- train new and practicing teachers in the instructional applications of psychological theory.
- foster an appreciation of the role that educational psychology can play in solving educational problems.
- work with schools and individuals to help them achieve their goals.

*College of Education & Human Development Mission Statement*

The new College of Education and Human Development is a world leader in discovering, creating, sharing, and applying principles and practices of multiculturalism and multidisciplinary scholarship to advance teaching and learning and to enhance the psychological, physical, and social development of children, youth, and adults across the lifespan in families, organizations, and communities.

## 2011 Performance Assessment Schedule

<b>January</b> <b>20</b>	History of Performance Assessment <ul style="list-style-type: none"> <li>▪ The origins of “assessment”</li> <li>▪ The many contemporary forms of performance assessment</li> </ul>	
<b>27</b>	Current uses of Performance Assessment <ul style="list-style-type: none"> <li>▪ Classroom uses</li> <li>▪ Large-scale uses (low-stakes and high-stakes)</li> </ul>	
<b>February</b> <b>3</b>	Conceptualizing the Assessment <ul style="list-style-type: none"> <li>▪ Purpose and politics</li> <li>▪ Assessment framework and blueprint</li> </ul>	<b>(a)</b> <b>(b)</b>
<b>10</b>	Design of Performance Assessment <ul style="list-style-type: none"> <li>▪ Developing tasks, prompts, scoring criteria and rubrics</li> <li>▪ Training task developers and scorers</li> </ul>	<b>[1]</b>
<b>17</b>	Performance Task Development Workshop	<b>(c)</b>
<b>24</b>	Rubric Development Workshop	<b>(d)</b>
<b>March</b> <b>3</b>	Portfolio Development Workshop	<b>(e)</b>
<b>10</b>	Implementation of Performance Assessment <ul style="list-style-type: none"> <li>▪ Pilot procedures</li> <li>▪ Standardizing administration</li> </ul>	
<b>17</b>	<i>SPRING BREAK</i>	
<b>24</b>	Optimizing Measurement Procedures Psychometrically <ul style="list-style-type: none"> <li>▪ Generalizability &amp; Dependability of Measurements</li> </ul>	<b>[2]</b>
<b>31</b>	Quantitative Analysis of Performance Scores	
<b>April</b> <b>7</b>	<i>NCME – No Class</i>	
<b>14</b>	Data Analysis Options in Practice	<b>[3]</b> <b>(f)</b>
<b>21</b>	Student Presentations	
<b>28</b>	Student Presentations	
<b>May</b> <b>5</b>	Assessing the Quality & Equity of Results <ul style="list-style-type: none"> <li>▪ Employing the <i>Standards</i></li> <li>▪ Evaluating psychometric quality</li> </ul>	



## ***Reading List***

- AERA, APA, NCME (1999). *Standards for Educational and Psychological Testing*. Washington DC: American Educational Research Association.
- American College Testing Program (1997). *Reliability issues with performance assessment: A collection of Papers* (ACT Research Report Series, 97-3). Iowa City, IA: ACT.
- Kane, M. B., & Mitchell, R. (1996). *Implementing performance assessment: Promises, Problems, and Challenges*. Mahwah, NJ: Lawrence Erlbaum Associates.
- LeMahieu, P. G., Gitomer, D. H., & Eresh, J. T. (1995). *Portfolios beyond the classroom: Data quality and qualities* (Center for Performance Assessment MS #94-1). Princeton, NJ: Educational Testing Service.

## **Additional Readings**

- Kozaki, Y. (2010). An alternative decision-making procedure for performance assessments: Using the multifaceted Rasch model to generate cut estimates. *Language Assessment Quarterly*, 7(1), 75-95.
- Niemi, D., Wang, J., Steinberg, D.H., Baker, E.L., Wang, H. (2007). Instructional sensitivity of a complex language arts performance assessment. *Educational Assessment*, 12, 215-237.

## **Generalizability**

- Chen, E., Niemi, D., Wang, J., Wang, H., Mirocha, J. (2007). *Examining the generalizability of direct writing assessment tasks*. CSE Technical Report 718.
- Clauser, B.E., Harik, P., Margolis, M.J. (2006). A multivariate generalizability analysis of data from a performance assessment of physicians' clinical skills. *Journal of Educational Measurement*, 23(3), 173-191.
- Shumate, S.R., Surlis, J., Johnson, R.L., Penny, J. (2007). The effects of the number of scale points and non-normality on the generalizability coefficient: A Monte Carlo study. *Applied Measurement in Education*, 20(4), 357-376.

## **Rater Considerations**

- Harik, P., Clauser, B.E., Grabovsky, I., Nungester, R.J., Swanson, D., Nandakumar, Ratna (2009). An examination of rater drift within a generalizability theory framework. *Journal of Educational Measurement*, 46(1), 43-58.
- Martinez, J.F., Goldschmidt, P., Niemi, D., Baker, E.L., Sylvester, R.M. (2007). Language arts performance assignments: Generalizability studies of local and central ratings. *Educational Assessment*, 12, 267-282.

*Additional online readings and resources will be available at the course website.*