

PEARSON



# TEST, MEASUREMENT, & RESEARCH SERVICES

*Quarterly Newsletter*

VOL 2 | NO 2 | 2009

## INSIDE

Editor's Note  
Announcements  
Advisory Board  
Pearson Sessions at the 2009  
CCSSO National Conference  
on Student Assessment  
Seminars & Training  
Other Presentations  
2009 Fellows  
TrueScores

## EDITOR'S NOTE

*by Jason L. Meyers*

**WELCOME TO THE FIFTH INSTALLMENT OF THE PEARSON TEST, MEASUREMENT, AND RESEARCH SERVICES (TMRS) NEWSLETTER, A QUARTERLY PUBLICATION AIMED AT PUBLICIZING THE ONGOING RESEARCH EFFORTS OF OUR GROUP TO THOSE BOTH WITHIN AND OUTSIDE OF OUR PEARSON COMMUNITY.**

I'd like to thank those of you who have sent me your contributions for the newsletter and commend you on your outstanding research; this effort would not be possible without you. I'd also like to acknowledge the excellent work of the Advisory Board who helped shape the direction of this issue. I recognize and appreciate that everyone must take time out of their busy schedule to make this newsletter a reality.

In this issue, I am pleased to announce the release of a set of Pearson guidelines for making computer-based testing more accessible to students, to introduce the 2009 Pearson Fellows, and to provide a recent blog by Dr. Jon Twing. In addition, this newsletter contains a detailed listing of all CCSSO presentations involving Pearson TMRS staff. This listing accentuates the NCME and AERA program presented in the last newsletter and further illustrates the wealth of research activities conducted by our group. By my count, we had 18 individual presentations by Pearson staff at this conference. Finally, in this issue we also detail the journal articles being published and the seminars and training sessions conducted by our staff during the second quarter of 2009.

We aim for widespread dissemination of this newsletter. If you or someone you know would like to be added to our distribution list, or if you require a printed version of the newsletter, please contact me directly. I also welcome questions, comments, and suggestions in a continual effort to improve the newsletter.

Back issues can be downloaded from [www.pearsonedmeasurement.com/research/newsletter.htm](http://www.pearsonedmeasurement.com/research/newsletter.htm). We hope you enjoy this issue and look forward to hearing from you.

## ANNOUNCEMENTS

### *Pearson Unveils Guidelines to Make Computer-Based Student Testing More Accessible*

#### **RESEARCH IDENTIFIES TACTICS TO ENSURE EQUITY FOR ALL STUDENTS**

Pearson, the global leader in education and education technology, recently announced the release of its Universal Design for Computer-Based Testing Guidelines, aimed at making computer-based testing more usable and accessible for all students, particularly those with disabilities.

"With computer-based testing becoming more prevalent at every turn, Pearson understands the importance of ensuring all students have an equal opportunity to perform well when being evaluated," said Denny Way, Senior Vice President, Psychometric & Research Services for the Assessment and Information Group of Pearson.

Working with the Center for Applied Special Technology (CAST), a nonprofit organization devoted to expanding learning opportunities for all, Pearson conducted a year-long study as the basis for its computer-based testing guidelines. Pearson believes the guidelines will help test developers better understand students' thought processes that occur during test-taking and will improve the design and accessibility of computer-delivered items and assessments.

*Continued on page 2*

**Dr. Jason L. Meyers, Editor**  
Assessment & Information, Pearson  
512.989.5300 ext 224004  
[jason.meyers@pearson.com](mailto:jason.meyers@pearson.com)

## NEWSLETTER

### ADVISORY BOARD

#### Paul Nichols

Vice President of Psychometric & Research Services

#### Mary Kino

Principal Research Scientist

#### Bob Dolan

Senior Research Scientist

#### Chingwei (David) Shin

Senior Research Scientist

#### Edward Wolfe

Senior Research Scientist

#### Kelly Burling

Research Scientist

#### Yuehmei (May) Chien

Associate Research Scientist

#### Jason L. Meyers

Editor, Research Scientist

*The Pearson Test, Measurement, & Research Services Newsletter is published quarterly. The newsletter is not copyrighted; readers are invited to copy any articles that have not been previously copyrighted. Credit should be given in accordance with accepted publishing standards.*

Send information to:

**Dr. Jason L. Meyers**

Phone: 512.989.5300

Ext 224004

[jason.meyers@pearson.com](mailto:jason.meyers@pearson.com)

400 Center Ridge Drive

Austin, Texas 78753

## ANNOUNCEMENTS (CONT.)

*Continued from page 1*

“These guidelines, when applied with fidelity, would move large-scale assessment a long way in just the right direction—toward more flexible and accurate measures of the academic progress of all students,” says David Rose, Chief Scientist of Cognition and Learning at CAST.

“Our study defines guidelines that will enable test developers to level the playing field and more accurately test the knowledge and skill sets of all students with a variety of abilities,” said Bob Dolan, Senior Research Scientist at Pearson and primary author of the study. “The guidelines have proven useful to Pearson’s test developers in working with our customers to implement innovative online assessments.”

During the study, Pearson and CAST considered existing guidelines for computer accessibility based on the frameworks of universal design and Universal Design for Learning, as well as other cognitive and educational research. In addition, they examined a variety of traditional and innovative computer-based testing item components, such as text, images, animations, and interactive elements, each of which requires students to use various processes to interact with and respond to test questions.

“Regardless of the knowledge and skills computer-based test developers intend to measure, every student, whether with a disability or without, has a diverse range of physical, sensory, and cognitive abilities that either enables or restricts his or her ability to respond to each test item accurately,” Dolan added.

Pearson’s Universal Design for Computer-Based Testing Guidelines examines the specific student challenges related to each test question construct and pinpoints question design solutions that can make test questions more accessible to all students. The study touts the value of digital technology and its ability to incorporate multiple representations, such as text, video, and audio, into computer-based testing.

In addition, the study provides test developers with a comprehensive checklist that addresses each knowledge and skill set by processing category and provides specific tactics to minimize the measurement of construct-irrelevant abilities.

To view Pearson’s Universal Design for Computer-Based Testing Guidelines, visit [www.pearsonedmeasurement.com/cast/index.html](http://www.pearsonedmeasurement.com/cast/index.html).

### About CAST

*Founded in 1984 as the Center for Applied Special Technology, CAST has earned international recognition for its development of innovative learning resources and for articulating the principles of Universal Design for Learning (UDL), a framework for creating inclusive educational environments. For more information, go to [www.cast.org](http://www.cast.org).*

## CSS Expands Management Structure

John Loughran is pleased to announce an expanded Content Specialists’ management structure. The following managers will be supporting each of the content areas:

<b>Suzanne Trevvett</b>	English Language Arts Content Development
<b>Lee Brinkerhoff</b>	English Language Arts Content Support (reports to Suzanne Trevvett)
<b>Linda Fralick</b>	Mathematics Content Development
<b>Bebe Mireles</b>	Mathematics Content Support (reports to Linda Fralick)
<b>Chris Rozunick</b>	Science Content Support
<b>Deb Hill</b>	Social Studies and Alternate Assessments Content Support
<b>Roger Frantz</b>	English Language Learners Content Support

## Awards

The following individuals were recognized for their outstanding contributions at the quarterly Rewards and Recognition presentation during the May session of Learning at Lunch.

<b>Alicia Chrest</b>	Team Player – South Carolina and LEAP
<b>Jerod Flippen</b>	All-Star – TAKS-M and TAKS-Alt
<b>Marguerite Hartill</b>	Team Player – TOP Prompt
<b>Amy Hathorn</b>	Team Player – TOP Prompt and Harcourt Index Art Library Initiative
<b>Lisa Hawk</b>	Team Player – Timecode templates
<b>Jean Henry</b>	Team Player – TOP Prompt
<b>Jadie Kong</b>	Team Player – TELPAS
<b>Karen Leach-Esposito</b>	Team Player – TOP Prompt
<b>Lisa Mahronic</b>	Team Player – SD STEP
<b>Paul Malary</b>	Team Player – Harcourt Index Art Library Initiative
<b>Matthew Malay</b>	Team Player – TOP Prompt
<b>Sue Medberry</b>	Team Player – TOP Prompt
<b>Wes Overton</b>	Team Player – South Carolina
<b>Jan Picciuti</b>	Team Player – TOP Prompt
<b>Stan Retzloff</b>	Team Player – Harcourt Index Art Library Initiative
<b>George Sauerberg</b>	Team Player – South Carolina
<b>Amy Schoon</b>	Team Player – South Carolina
<b>Deanna Thomann</b>	Team Player – South Carolina
<b>Ursula Willing</b>	Team Player – Harcourt Index Art Library Initiative
<b>Wenyi You</b>	Team Player – TAKS

## PROGRAM EXCERPTS OF PEARSON SESSIONS AT THE 2009 CCSSO NATIONAL CONFERENCE ON STUDENT ASSESSMENT

### *Changing the Way We Measure Science: The Promise of High-Fidelity Test Items*

#### SUNDAY, JUNE 21

Advances in technology have dramatically improved the potential for presenting high-fidelity simulations to students. Through simulations we can allow students to participate in otherwise dangerous laboratory situations, observe phenomena that might ordinarily take months or years of data collection during a single class, or view output from equipment that would be prohibitive for most schools to access (e.g., electron microscopes). We can also harness this potential for assessing students' science achievement. Minnesota currently uses science figural items on its state assessment. These items incorporate video, graphics, and simulations to improve assessment authenticity. The NAEP Science Interactive Computer Tasks use similar formats, but take student assessment a step further, including drag-and-drop experiment design items, student construction of complex graphics, background data access via interactive software, etc. This session will focus on the rationale, development, utility, and psychometric and other challenges associated with high-fidelity science test designs.

#### PRESENTERS

**Arthur Thacker** Human Resources Research Organization (HumRRO)  
**Shu-Kang Chen** Educational Testing Service (ETS)  
**Jim Wood** Minnesota Department of Education  
**Kristi Thompson** Pearson

#### DISCUSSANT

**Rob Kirkpatrick** Pearson



Continued on page 4

## PEARSON 2009 CCSSO SESSIONS

Continued from page 3

### *Assessing Writing Online: The Benefits and Challenges*

**SUNDAY, JUNE 21**

The transition to assessing student writing online presents both benefits and challenges to states and their students. This session will discuss the logistical, content, scoring, and political issues states face while implementing the transition to assessing student writing online. Louisiana will provide insight on implementing an online writing assessment, and Minnesota will discuss the challenges it faced while attempting to develop an online writing component. In addition, research and content-based perspectives will be offered from two testing companies, Pearson and Pacific Metrics Corporation.

#### **PRESENTERS**

**Jennifer Isaacs** Pacific Metrics Corporation  
**Denny Way** Pearson  
**Claudia Davis** Louisiana Department of Education  
**Dirk Mattson** Minnesota Department of Education

### *Three States' Experiences in Implementing a Vertical Scale*

**SUNDAY, JUNE 21**

In this session, we describe the methods used by three states to implement vertical scales in reading and mathematics across grades 3 through 8. A vertical scale has become a desirable component of a state's assessment program in recent years because schools that do not meet their AYP requirements in terms of the number of students meeting standards may still be counted as meeting the AYP requirements if they can show that acceptable progress has been made. In Virginia, tests are administered online and on paper, and a vertical scale had to be created that would be applicable to both modes. In Texas, tests are administered in English and Spanish,

so two different vertical scales were developed. Mississippi is implementing a multiyear plan, which includes developing vertical linking items that reflect the progression of content through the curriculum, and monitoring student performance prior to reporting results on the vertical scale.

#### **PRESENTERS**

**Julie Miles** Pearson  
**Ahmet Turhan** Pearson  
**Kay Um** Pearson

### *Accommodations in a Computer-Based Testing Environment*

**SUNDAY, JUNE 21**

States increasingly are delivering or considering delivery of assessments via computer. Some states are working toward (or have fulfilled) a dual administration model while others are exploring using only computer-based testing in the future. Use of the computer opens the door to technological solutions for accommodations. However, the use of technology for accommodations also raises a number of questions, such as ease of use and score comparability. This session will discuss the research and development of computer-based accommodations, such as text-to-speech and onscreen magnification, and discuss the policy and practical issues/potential solutions surrounding their development and use in a secure testing environment. This will be an open presentation, encouraging participation from audience members. Handouts will be available.

#### **PRESENTERS**

**John Poggio** Center for Educational Testing and Evaluation,  
University of Kansas  
**Bob Dolan** Pearson  
**Shelly Loving-Ryder** Virginia Department of Education  
**Todd Nielsen** Iowa Testing Programs

#### **DISCUSSANT**

**Sue Rigney** U.S. Department of Education

#### **MODERATOR**

**Todd Nielsen** Iowa Testing Programs



© 2009 Jupiterimages Corporation

## Ensuring Technical Quality of Formative Assessments

**MONDAY, JUNE 22**

Reference to an assessment as formative is shorthand for the formative use of assessment data—whether coming from standardized tests, teacher observations, or intelligent tutoring systems—with the explicit goal of providing focused interventions to improve student learning. The technical quality of assessments indicates the extent to which interpretations and decisions derived from assessment results are reasonable and appropriate. However, familiar technical requirements such as validity and reliability have been developed with a focus on summative assessment and have not considered a coordinated system of instruction and assessment. For example, reliability has traditionally indicated the degree of consistency of test scores over replications, a definition that would inadequately describe a formative assessment that inconsistently prescribes appropriate targeted instruction over time and across conditions.

This session will discuss different approaches to defining new indicators of technical quality appropriate for ensuring the effectiveness of formative assessment systems.

### PRESENTERS

**Bob Dolan** Pearson  
**Meg Litts** Onamia Public Schools, MN  
**John Poggio** University of Kansas  
**Jerry Tindal** University of Oregon

### DISCUSSANT

**Tim Peters** New Jersey Department of Education

## Application of Validity Studies for College Readiness: the American Diploma Project Algebra II End-of-Course Exam

**MONDAY, JUNE 22**

Too many students graduate from high school unprepared for college—nearly one-fourth of first-year college students must take remedial courses in mathematics. An intended use of the American

Diploma Project (ADP) Algebra II End-of-Course Exam is to serve as an indicator of readiness for first-year college credit-bearing courses—to ensure that students receive the preparation they need while they are still in high school. This session will discuss how the multistate ADP consortium gathered validity evidence for this use of the exam through studies involving: 1) content judgments of college instructors; 2) comparisons of college students' ADP Algebra II End-of-Course Exam performance with subsequent grades in college level courses; and 3) empirical relationships between exam performance and ACT or SAT scores. Representatives from two ADP states, Achieve, and the exam vendor will share their experiences with collecting the data and how results of the exam will be used.

### PRESENTERS

**Julie Miles** Pearson  
**Robert Triscari** Achieve  
**Stan Heffner** Ohio Department of Education  
**Kent Hinton** Hawaii Department of Education

## Keeping All Those Balls in the Air: Challenges and Approaches for Linking Test Scores Across Years in Multiple Format Environments

**MONDAY, JUNE 22**

In this era of requiring annual improvements in student test scores, valid test score linking is one of the most important components of a state testing program. To make things more challenging, many states are implementing computer-based testing programs, but few if any are able to completely shift to CBT platforms. Therefore, in addition to the usual challenges with test score linking, states are trying to ensure that comparable inferences are drawn from the same scale scores no matter the format of the test. This session brings together equating contractors and testing directors from two states that have successfully addressed these challenges to share their lessons learned and offer recommendations to other state leaders.

*Continued on page 6*

## SEMINARS & TRAINING

On June 16, 2009 in San Antonio, Alvaro Arce-Ferrer presented a training session titled **“Item Parameter Drift Detection Methods: A Source for Optimism or a Cause for Concern.”** This training session was directed to practitioners working with one- and three-parameter item response models. The training offered a theory component to introduce essentials of item parameter drift techniques (IPD) and discussed areas for improvement. The training also provided hands-on experiences with IPD techniques with real and simulated sets.

During the week of June 22, 2009, Psychometric and Research Services staff presented a **General Item Analysis Training** to Content Support Services staff in Austin, Texas; San Antonio, Texas; and Iowa City, Iowa.

During the week of June 29, Psychometric and Research Services, in conjunction with staff from Content Support Services, facilitated a **Mock Data Review Training** that was delivered to Content Support Services staff in Austin, Texas; San Antonio, Texas; and Iowa City, Iowa.

## PEARSON 2009 CCSSO SESSIONS

Continued from page 5

### PRESENTERS

**Scott Marion** National Center for the Improvement of Educational Assessment

**Matt Trippe** HumRRO

**Deborah Swensen** Utah State Office of Education

**Tony Thompson** Pearson

**George Henley** Minnesota Department of Education

**Dirk Mattson** Minnesota Department of Education

### DISCUSSANT

**Rich Hill** National Center for the Improvement of Educational Assessment

## *Revising the Standards for Educational and Psychological Testing*

### MONDAY, JUNE 22

AERA, NCME, and the APA have launched an effort to revise the Standards for Educational and Psychological Testing. Several members of the committee drafting this revision will describe issues being addressed and the process and timeline for completing our work. The charge to this committee specifies areas of focus including: (a) the increased use of technology in testing, (b) the increased use of tests for educational accountability, (c) access for all examinee populations, and (d) issues associated with workplace testing. The committee will also review the scope and formatting of the Standards. A state testing director will be invited to serve as discussant, describing implications of the Test Standards for state assessments. A significant part of the session will be devoted to questions and comments from the audience.

### PRESENTERS

**Laress Wise** HumRRO

**Brian Gong** National Center for the Improvement of Educational Assessment

**Linda Cook** Educational Testing Service

**Joan Herman** CRESST

**Denny Way** Pearson

### DISCUSSANT

**John (JT) Lawrence** California Department of Education

## *Lessons Learned and the Road Ahead for the ADP Algebra II Consortium*

### TUESDAY, JUNE 23

Students across 12 states took the American Diploma Project Algebra II End-of-Course Exam for the first time in Spring 2008. In August 2008, the first annual report of the results and findings was released. This session will focus on the lessons learned from the first administration, progress toward the 3 common goals set for this endeavor by participating states and the challenges ahead for the multistate consortium. Among topics to be discussed are the validity studies that are being conducted to help establish the exam as an indicator of student readiness for first-year college credit bearing courses as well as how the states are using the assessment to improve high school Algebra II curriculum and instruction.

### PRESENTERS

**Laura Slover** Achieve, Inc.

**Shilpi Niyogi** Pearson

**Tim Peters** New Jersey Department of Education

**Gayle Potter** Arkansas Department of Education

**Bernie Sandruck** Howard Community College

## *The Role of Technology in Improving Turnaround Time and Quality in Large-Scale Assessments*

**TUESDAY, JUNE 23**

To meet their ever-increasing needs to have faster turnaround of test scores and more defensible scores, states are searching for ways to satisfy their constituents and to meet the demands of high stakes testing. This panel will discuss implementations of programs that leverage technology. One test that is administered frequently and requires immediate turnaround is the ACCUPLACER, which is given to incoming freshmen at community colleges. The College Board uses an online testing environment to administer the tests and an automated intelligence engine to score them. North Carolina uses a distributed scoring model that engages large groups of scorers for its writing assessment and local district scoring of multiple-choice tests for their EOC and EOG tests. Virginia administers all multiple-choice tests online with Rapid Turnaround to score and deliver results. These three models meet the demands of later or on-demand testing and the demands for faster or immediate test results.

### **PRESENTERS**

**Daisy Vickers** Pearson

**Jim Kroenig** North Carolina Department of Education

**Ed Hardin** College Board

**Shelley Loving-Ryder** Virginia Department of Education

## *Accurate and Time-Saving: Online Assessment of Oral Reading Fluency Using Advanced Speech Processing Technology*

**TUESDAY, JUNE 23**

This panel will describe four studies conducted across eight states investigating the usability and impact of using an online, automated test delivery and scoring system to measure and track students' oral reading fluency (ORF) performance. The ORF system produced words-correct-per-minute (WCPM) scores for oral reading samples from hundreds of 1st through 5th graders. The session includes discussion of: 1) technical and practical challenges involved in large-scale test

delivery, scoring, and data management; 2) the reliability of the automated scoring system, which produces scores that correlate highly (0.95–0.99) with teachers scoring manually; 3) policy-related impacts of reliability data, comparing machine scores with scores from expert test administrators and classroom teachers; 4) innovative methods for scoring other aspects of ORF (e.g., expressiveness, accuracy); 5) teacher feedback on the value of the automated system, including how automated scoring enables reallocation of teacher time from test administration and scoring to instruction.

### **PRESENTERS**

**Ryan Downey** Pearson

**David Rubin** Pearson

**Jack Shaw** National DIBELS trainer/consultant

**DeAnna Pursai** San Jose Unified School District, CA

## *Developing Valid Alternate Assessments with Modified Achievement Standards: Three States' Approaches*

**WEDNESDAY, JUNE 24**

Implicit in the design of an alternate assessment based on modified achievement standards (AA-MAS) is a validity argument that the assessment appropriately and accurately measures the grade-level academic achievement of students in a targeted subpopulation of students with disabilities. In this session, three states at different stages in the development process will present their approaches to developing valid AA-MAS. The states will discuss their test development models and factors influencing their choice of model. Presentations will cover the rationales and processes used for item development, research used to support test development, activities involving stakeholders, and the process of collecting validity evidence. A discussant will respond to the development models presented, focusing on threats to validity, documenting validity evidence, and employing ongoing validity evaluations. The discussant's presentation will be relevant to the test designs used in the three states and to AA-MAS test development more generally.

*Continued on page 8*



## PEARSON 2009 CCSSO SESSIONS

Continued from page 7

### PRESENTERS

**Kelly Burling** Pearson  
**Shelley Loving-Ryder** Virginia Department of Education  
**Cari Wieland** Texas Education Agency  
**Elizabeth Hanna** Pearson  
**Malissa Cook** Oklahoma Department of Education

### DISCUSSANT

**Stuart Kahl** Measured Progress

### MODERATOR

**Kate Nagle** SRI

## *Legislation in the One Corner, Implementation in the Other: And, It's a Knockout*

### WEDNESDAY, JUNE 24

Assessment legislation often gets passed before an implementation plan is fully vetted. In this session, Minnesota, Nebraska, and Texas square off against challenging assessment legislation. They will describe how they bob and weave as policy is thrown their way to develop implementation plans that are reasonable and in the best interest of their students. Presenters will share legislation that turned the local assessment system upside down, such as Texas legislation limiting field testing at a time when 12 end-of-course assessments are to be rolled out. Join this session and listen as these three states explain how they knocked out the legislative challenges early instead of going the full 12 rounds.

### PRESENTERS

**Kimberly O'Malley** Pearson  
**Christy Hovanetz Lassila** Minnesota Department of Education  
**Pat Roschewski** Nebraska Department of Education  
**Gloria Zyskowski** Texas Education Agency

### DISCUSSANT

**Roger Trent** Executive Director Emeritus for the Ohio Department of Education

## *Cognitive Interviews Applied to Test and Item Design and Development for AA-MAS (2 percent)*

### WEDNESDAY, JUNE 24

The session focuses on applying cognitive interviews (CI) in the development of alternate assessments judged against modified achievement standards (AA-MAS) and presents principles for using CI with AA-MAS that were formulated during a research symposium and published in a 2009 white paper. The symposium built on work of Designing Accessible Reading Assessments and Partnership for Accessible Reading Assessment, addressing think-aloud methods with students eligible for AA-MAS from recent research. White paper principles and four recent CI AA-MAS studies will be presented. Studies will be discussed in the context of the principles and how results can be used to inform test design and development in AA-MAS. Copies of the white paper will be available, and audience interaction will be encouraged. Participants will be invited to ask questions, offer experiences, and discuss methods for interviewing students with limited communication, gathering reliable data, and applying CI to AA-MAS test development in their own settings.

### PRESENTERS

**Patricia Almond** University of Oregon & Consultant to SRI International  
**Caroline E. Parker** Education Development Center, Inc. (EDC)  
**Chris Johnstone** National Center on Education Outcomes  
**Shelley Loving-Ryder** Virginia Department of Education  
**Jennifer Stegman** Oklahoma Department of Education  
**Kelly Burling** Pearson

### DISCUSSANT

**Phoebe Winter** Metrics Corporation



## OTHER PRESENTATIONS

In March, **John Loughran**, **Deb Hill**, and **Dyanne Ornelas** presented at the Association for Supervision and Curriculum Development (ASCD) annual conference in Orlando, Florida. Their presentation, *Developing High-Level Multiple-Choice Questions*, focused on the differentiation between difficulty and cognition by presenting several examples of cognitively challenging items at various levels of difficulty.

## 2009 FELLOWS & INTERNS

Each summer, Pearson offers a fellowship for doctoral students with outstanding psychometric skills who wish to gain experience in educational testing. Fellows have the opportunity to work closely with a team of Pearson research scientists in order to

- » Gain experience in many of the tasks involved in a K-12 testing contract.
- » Collaborate on a research paper, abstract, or presentation for national dissemination.
- » Better understand ways to collaborate and communicate about educational measurement.

Applications for the fellowship are due in March and awarded in late March or early April. This year, we received over 50 applications for 4 available positions. We are pleased to announce the 2009 Fellowship recipients in this issue. Please join us in giving them a warm Pearson welcome.

### *Casey B. Nixon, Austin*

Casey is a doctoral student in Educational Psychology at the University of Georgia, Athens. She holds an M.A. in Educational Psychology with a concentration in Research, Evaluation, Measurement, and Statistics from the University of Georgia and a B.A. in Psychology with a minor in Anthropology from Texas State University. Her research interests include, but are not limited to, item response theory, MIMC modeling, and structural equation modeling.

### *Yan Liu, San Antonio*

Yan Liu is a doctoral student in Measurement, Evaluation, & Research Methodology at the University of British Columbia in Vancouver, Canada. She holds an M.A. in Measurement, Evaluation, & Research Methodology from the University of British Columbia, a B.A. from Beijing Second Foreign Language University in Beijing, China, and an Associate Degree in Elementary Education from Hefei Teachers School in Hefei, China.

### *Laine Bradshaw, San Antonio*

Laine Bradshaw is a doctoral student in Educational Psychology at the University of Georgia, Athens. She holds B.S. and M.Ed. degrees in Mathematics Education from the University of Georgia.

### *Hong Wang, Iowa City*

Hong Wang is doctoral student in Research Methodology at the University of Pittsburgh. She holds a B.A. in English in International Business from Nan Kai University in Tainjin, China, and an M.A. in English with a concentration in English as a Second Language from the University of Toledo. Her research interests include application of item response theory models to large-scale assessments, test equating, and the application of multilevel models and structural equation models to educational research.

In addition to the Pearson Fellowship program, Pearson offers an extended 6-month internship program through collaboration with James Madison University. This year's internship recipient is Shelley Ragland. Shelley is working out of the Austin office. Please join me in welcoming Shelley to our group.

### *Shelley Ragland, Austin*

Shelley Ragland is a doctoral student in Assessment and Measurement at James Madison University in Harrisonburg, Virginia. She holds a B.S. in Mathematics from New Mexico State University. Her research interests IRT applications to equating, linking, and scaling.



## TRUESCORES

Each issue of the Pearson Test, Measurement, and Research Services (TMRS) Quarterly Newsletter includes a recent entry from the TrueScores blog written by Jon Twing. For more information on TrueScores, please visit [www.truescores.com](http://www.truescores.com).

### *If David Beats Goliath, Just What Role Does the Psychometrician Play?*

by Jon S. Twing

So I admit it—I, too, like a good, sensationalized feel-good piece of literature, one that particularly has some application to what we do for a living. My very good but departed friend, Ed Slawski, used to say I was such a soft touch.

Regardless, I recently read a piece by that world-famous author Malcolm Gladwell (you know the *The Tipping Point*, *Blink*, and *Outliers* guy) that did indeed move me to write this blog. His article appeared in the *New Yorker* (yes, even an old curmudgeon like me actually subscribes to the *New Yorker*) and is titled “How David Beats Goliath.” In this article, Mr. Gladwell parallels how the underdogs seem to win more often than they should because they change the rules on how the game is played. He uses the full-court press in basketball as an example of a strategy that a smaller, less talented team might use to beat a taller, more talented team. He actually cites statistics regarding the success of such a strategy.

Psychometricians, as you may well know, are very methodical people. They like specifications outlining what it is that they do. They like to follow procedures, are often meticulous, and believe in verification, transparency, and replicability. Simply stated, they **follow the rules**. So, the question at hand is: Will psychometricians be the David or the Goliath of the new assessment reform implicit in the new administration and more explicit as a goal of Secretary Duncan?

Certainly, the brave new world of assessments in the post-NCLB era will be unlike what we have seen to date. Reliability as a measure of internal consistency, and validity as a correlation coefficient with existing measures are not likely to be the psychometric quality mantra moving forward. These new assessments are likely to be driven by needs for problem-solving measures, measures of critical thinking,

assessments of our ability to manage large amounts of information (presumably coming from the Internet), and comparability with international benchmarks all implemented and managed in an online and automated way. By definition, these new assessments will **violate the rules** and will be the David that defeats Goliath. I wonder how the greater psychometric community will react. “Just say no” comes to mind but will likely be a woefully inadequate response.

I certainly hope that under my direction, my staff will embrace the need to see the world differently now, even more so, lest we fall behind and are not in a position to support this next generation of assessments. Pearson plans to be the giant killer in this regard, changing the rules and leading the way into this new generation of learning. What is it that you plan to do?

*Several recent research studies illustrate how TMRS staff is embracing the need to see the world differently from an assessment perspective. Below are a few such examples:*

#### BOOK CHAPTER

Way, W. D., Dolan, R. P., & Nichols, P. (2009, in press). Psychometric Challenges and Opportunities in Implementing Formative Assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of Formative Assessment*. London: Routledge.

#### RESEARCH REPORT

Nichols, P. D. (submitted). Toward a reliability theory for systems for learning.

#### JOURNAL ARTICLE

Nichols, P. D., Meyers, J., & Burling, K. (in press). A framework for evaluating and planning assessments intended to improve student achievement. *Educational Measurement: Issues and Practice*.

#### WHITE PAPER

TMRS staff Ellen Strain-Seymour, Bob Dolan, Kelly Burling, and Denny Way developed a white paper titled, *Strategies and Processes for Developing Innovative Items in Large-Scale Assessments*. The paper was distributed at the 2009 CCSSO National Conference on Student Assessment and will be available at [www.pearsonedmeasurement.com](http://www.pearsonedmeasurement.com).