



FROM THE PRESIDENT

By David A. Frisbie, University of Iowa

Over this past year, I have used this Newsletter column to keep NCME members informed about the many significant organizational changes that have transpired. I won't recount the history and reasons for the changes below but will point forward to a year of transition when we begin to adapt to the changes.

Journal Publications. Our transition from AERA as publisher of our journals to Blackwell Publishing Co. has already begun. By now you should have received the first issue of JEM and EM:IP published for us by Blackwell. During this next year (our year runs from April to April), you will receive information about how you can access our current journals electronically and how, in the near future, you will be able to search previous issues electronically. I'll have more to say about our publications in Montreal, and our website will be used to communicate with everyone about these changes as they occur.

Management Services. We now have finalized a contract with The Rees Group, located in Middleton, Wisconsin, to provide central office services for NCME. Although the contract begins July 1, 2005, transition work will begin on June 1 so that many of the operational aspects of our work will be in place shortly after July 1. Bruce Wheeler, the Rees staff member who will be our next Executive Director, will attend our annual meeting in Montreal as an observer to learn more about NCME and our annual meeting. Beginning in June, look for additional information on our website about contacting The Rees Group, rather than AERA, regarding NCME questions.

Future Annual Meetings. We are working with the AERA staff on establishing a contract for AERA to provide selected annual meeting services to NCME so that the two organizations can continue meeting jointly in the future. Although we don't have a formal agreement yet, both organizations are proceeding with planning for 2006 as though an agreement will be reached this summer.

Montreal Annual Meeting. By now you should have received a copy of the annual meeting program, which also appears on the NCME website. Enrollments for the training sessions, to be held Sunday and Monday before the NCME program sessions begin, are high enough, as I write, that all proposed sessions are likely to be offered. If you have not enrolled in advance for training, it will be possible to register on site as long as space exists in the session you want. (Check the NCME website to learn how to register for training after you have registered for the annual meeting.)

Membership and Finances. The Membership Committee has been active in verifying our membership roster and in encouraging former members to rejoin. In addition, our financial condition continues to improve; we are on track to finish the year in the black and can begin to consider new initiatives or increased financial support for some of our committees' work. Full reports on membership and finances will be presented at the NCME Breakfast.

The Future of NCME. The organizational changes we have begun this year are far from complete, but they will provide for a firm administrative foundation for supporting whatever professional or academic paths our membership decides to take. It is a perfect time for introspection—reviewing our stated mission, establishing new goals and affirming old ones, examining our primary activities and wondering about new ones, and looking at who we are as well as who we think we might want to become. During the next year, we should do some self-assessment, some needs analysis, and maybe even some strategic planning. (We could call it something else, but we still should do it!) My sense is that our members' interests have shifted somewhat: what many now need from NCME membership is not consistent with what we have traditionally tried to offer. To the extent that my observations are accurate, it would be good to determine which constituencies we should serve and how best to serve them.

I'm grateful to many members who generously contributed their time this year to help NCME in so many ways. I've been extremely fortunate to have Suzanne Lane and Jim Impara on the Executive Committee with me to share the work, decisions, and occasional turmoil that this year has presented. Members of our Board of Directors have been diligent in fulfilling their responsibilities. Julie Noble and Xiaohong Gao have been superb in creating a program for our annual meeting that we'll all find meaningful. I thank them all for their unselfishness in serving NCME and for helping me personally this year.

DEVELOPING ACCESSIBLE READING ASSESSMENTS

By Martha L. Thurlow, National Center on Educational Outcomes
and Cara Cahalan-Laitusis, Educational Testing Service

In October, 2004, the U.S. Department of Education's Office of Special Education Programs (OSEP) funded two research and development projects to produce reading assessments accessible for students with disabilities that affect reading. This funding reflects a concern that there are still many issues facing students with disabilities when they participate in reading tests in the United States today. These issues include variability in accommodations that are allowed across states and how scores are reported.

The projects are named PARA and DARA. The **Partnership for Accessible Reading Assessments (PARA)** is a consortium of the National Center on Educational Outcomes (NCEO), the Center for Research on Evaluation, Standards, and Student Testing (CRESST), and Westat, Inc. **Designing Accessible Reading Assessments (DARA)** is housed at Educational Testing Service.

The two OSEP-funded projects have a common set of four goals:

1. Formulate a definition of the construct of "reading proficiency" that provides a basis for research and development of accessible large-scale tests of reading proficiency that yield both valid measures of proficiency against academic standards and meaningful individual reports for the full range of students with disabilities that affect reading.
2. Conduct a program of research on the assessment of reading proficiency to determine the effects of various factors of test development, design, and administration on accessibility, validity, and comparability for students with disabilities that affect reading.
3. Develop research-based principles and guidelines for making large-scale assessments of reading proficiency more accessible for students who have disabilities that affect reading.
4. Develop and field-test instruments or methods for assessing reading proficiency that are suitable for large-scale administration for school accountability purposes; that are accessible to students who have disabilities that affect reading; that maintain validity and comparability of scores; and that can provide valid measures and individual reports for the full range of students with disabilities that affect reading.

In addition to carrying out individual lines of research, as required by funding the projects are working collaboratively on Goals 1, 3, and 4. The **National Accessible Reading Assessment Projects (NARAP)** is the designated name of the collaborative effort, which started with the development of a joint Web site, www.narap.info. The site contains the overall project goals and several other pieces of information (with more to come) about the current and future work of the collaborative. In addition, you will find links to each of the separate projects with PARA at www.readingassessment.info and DARA at www.ets.org/DARA.

The funded projects have also established a joint General Advisory Committee. This Committee is made up of representatives from test developers, reading educators, researchers, disability groups, and a variety of other interested organizations. Its charge is to review and provide advice on the general plans of the project, and serve as a liaison with significant stakeholder groups. NCME is represented on the Committee by David Frisbie.

Immediate next steps for the projects involve working with a Definition Panel to develop a definition of *reading proficiency*. This

definition will be the foundation for the research that the projects conduct, and may result in adjustments to their proposed research plans. Input on the definitions will be obtained from the field through a series of focus group meetings that will be held in connection with major conferences (including NCME's annual conference in 2005).

NCME ANNUAL MEETING HIGHLIGHTS

Begin on page 4

WHAT TO DO IN MONTRÉAL

By Shelley Rohar, McGill University.

A little bit of Europe...in Canada? Yes, that's what you will find in Montréal! While Montréal has conserved the European charm characterizing the first settlements founded on this continent, it is a resolutely modern city turned toward the world. Home to half of Quebec's population, with its 3.4 million inhabitants, Montréal is a place where French and English traditions meet, rich with international cultures, a place that is the very embodiment of the people that live here: colorful and full of life!

Language

Montréal is the second largest French-speaking city in the world (after Paris). However, in Montréal the majority of people speak both French and English so you can always make yourself understood.

(continued on page 3)

NEWSLETTER ADVISORY BOARD

ROBERT ANKENMANN, University of Iowa
JUDY ARTER, Assessment Training Institute
SCOTT BISHOP, Riverside Publishing
MICHELINE CHALHOUB-DEVILLE, University of Iowa
KATIE FISK, Connecticut State Department of Education
JOAN HERMAN, CRESST/UCLA
SHARON LEWIS, Council of the Great City Schools
DUNCAN MACQUARRIE, Harcourt Educational Measurement
WENDY MCCOLSKEY, SERVE
HILLARY MICHAELS, CTB/McGraw-Hill
CAROL S. PARKE, Duquesne University
S.E. PHILLIPS, Consultant
BARBARA PLAKE, Buros Center for Testing, University of Nebraska-Lincoln
DOUGLAS RINDONE, Harcourt Educational Measurement

SUSAN M. BROOKHART, Editor, Duquesne University

Send articles or information for this newsletter to:

Susan M. Brookhart Phone: (406) 442-8257
2502 Gold Rush Avenue Fax: (406) 442-8257
Helena, MT 59601 e-mail:
susanbrookhart@bresnan.net

The NCME 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.

WYOMING'S INSTRUCTIONALLY SUPPORTIVE NCLB TESTS

By W. James Popham, UCLA

Wyoming educators are currently creating a series of NCLB accountability tests deliberately designed to foster improved instruction in that state's classrooms. These new tests, known as the *Proficiency Assessments for Wyoming Students* (PAWS), reflect the aspiration of Wyoming Superintendent of Public Instruction, Trent Blankenship, to create statewide assessments that would satisfy federal accountability mandates, but would also serve as potent catalysts for the improvement of instruction in his state's public schools. The new tests are being created for Wyoming by Harcourt Assessment, Inc.

PAWS represents the nation's first attempt to construct NCLB tests in complete accord with key recommendations of a national advisory group, the Commission on Instructionally Supportive Assessment. In October 2001, prior to the passage of NCLB, that 10-member Commission identified a series of requirements that large-scale accountability tests must satisfy if those tests were to supply accurate accountability evidence, yet also support improved instruction by teachers.

Three of the Commission's most important requirements for genuine instructionally supportive accountability tests were that such tests must (1) measure students' mastery of only a modest number of extraordinarily significant curricular aims, (2) describe the nature of those aims for teachers with great clarity, and (3) provide aggregatable reports of each student's attainment of every curricular aim assessed. PAWS will satisfy all three of those Commission requirements.

After extensive curriculum-related deliberations, Wyoming educators decided that PAWS will assess students' mastery of only 23 significant outcomes, 8 in reading, 4 in writing, and 11 in mathematics. These outcomes will be succinctly and clearly described—and a student's performance on each assessed curricular aim will be provided to teachers, students, and students' parents.

Wyoming districts may choose *Year-End Tests* whereby a single version of all PAWS tests is given in March, or *Year-Round Tests* in which students are tested on approximately half of the PAWS-designated curricular aims in January and April. Moreover, districts will have access to a wide variety of formative tests so that teachers, using gratis online or paper copies, can monitor their students' progress toward all PAWS tested curricular goals.

Additional information about this PAWS can be obtained from Annette Bohling, Wyoming Department of Education's Deputy Superintendent of Educational Quality and Accountability abohli@educ.state.wy.us.

WHAT TO DO IN MONTRÉAL

(continued from page 2)

Getting Around

Montréal has a great transit system, including a safe and clean Metro (subway). This is undoubtedly the most efficient way to navigate the city. The longest east/west street is Sherbrooke street; an excellent reference point on a map. www.stm.info (transit maps)

Must-See's

If you have only a few days to spend, the following is a list of places you won't want to miss. (Visit these websites for detailed

information: www2.ville.montreal.qc.ca, www.montreal.com and www.tourisme-montreal.org)

- **St-Catherine Street/Downtown**

Bring your wallet and walking shoes when you visit Montréal's main commercial street; if you get tired, not to worry, as you'll find a restaurant (or two!) on every corner. Access Montréal's famous "Underground City" for even more shopping. Unique in the world, Montréal's underground network is a twenty-mile maze of pedestrian passageways linking eight shopping centers and hundreds of restaurants, boutiques and specialty shops. Many fashionable streets are found in this vicinity, each with its own character: Crescent, Bishop and De La Montagne are close together and offer mostly restaurants and bars. A very popular nighttime place. Metro stops: Guy, Peel, McGill.

- **Old Montréal and the Old Port**

Cobblestone streets, old European-like charm, restaurants galore, museums and shops, remarkable 18th and 19th-century architecture and a wide variety of educational and fun-filled activities. Not quiet, but definitely fun! There's a lot to be seen so if you want to give your feet a rest, consider a calèche (horse-drawn carriage). Metro stops: Place D'Armes, Champ de Mars.

- **China Town**

Giant red and gold arches mark the entrance to China Town, located a short walk south of downtown. China Town is delightful to walk through. There is a hidden temple to discover, many restaurants, and stores that sell Oriental herbs and medicines. Metro stops: Metro Place D'Armes.

- **St-Laurent Boulevard**

If you want to "see & be seen" this is the spot for you! Plenty of trendy avant-garde restaurants and bars/clubs. St. Laurent Boulevard, also called "The Main", divides the city between east and west. Any address with "Est" is necessarily to the east of St-Laurent. If you're in the mood for a taste of Montréal, be sure to stop by Schwartz's for their famous smoked meat on rye. Metro stop: St-Laurent.

- **The Plateau Mont-Royal**

A visit to Montréal is not complete without at least a little walk in The Plateau and a stop at a café or a local neighborhood restaurant. Le Plateau, although a more residential area, is bordered by St-Denis street which is a very lively and artsy area, often compared to the most "IN" neighbourhoods worldwide: bars, restaurants, café's, unique boutiques and designers, theater and charming ambience. Metro stops: Berri-UQAM, Sherbrooke, Mont-Royal.

- **Mont-Royal Park**

If you want the city at your feet (all 769 of them to be precise!), simply climb to the Summit of Mont-Royal Park. Its lookout offers a spectacular view of Montréal. The park itself is replete with green space and year-round activity. On Sundays, if the days are sufficiently warm and you feel like marching to the beat of a different drum, then check out the "Tam Tam Jam" where locals gather to play music and soak up the sun. Metro stop: Mount-Royal, bus or walk to the Summit.

- **The Olympic Park Area**

Located to the East of downtown lies the Olympic Stadium, or the "Big O" as locals call it. Erected in 1976 to host the Summer Olympic games, the Stadium offers wonderful photo opportunities. Ascend 556 feet on the funiculaire to the top of the Olympic Stadium Tower, the world's tallest inclined tower, with a 50-mile view of Montréal, the Laurentian Mountains and the Saint Lawrence Seaway. On a clear day you can see all the way to Vermont and the Adirondacks. Located nearby you will find the Biodôme de Montréal, Insectarium, and The Botanical Gardens (open year round). Metro stops: Pie-IX, Viau.

- **Ile Notre-Dame & Ile Ste-Hélène**

These two islands form a park called Parc Jean-Drapeau. Here you can find the Montréal Casino, the La Ronde amusement park, the Biosphere (environmental exhibition), the David M. Stewart museum (fascinating history of the New World), the Governor's feast (public

dinner and show with Nouvelle France theme), outdoor shows, walking and hiking trails, bicycling and roller blading. The Island is also home to the Gilles-Villeneuve racing track, where the Canadian Formula 1 Grand Prix is held every year. You can visit the official web site (www.parcjeandrapeau.com) for further information.

Art Galleries & Museums

- **Musée des Beaux Arts** (Montréal Museum of Fine Arts) is the grande dame of the Canadian museum world. Admission to the permanent displays is free. Metro stop: Guy-Concordia -- www.mmfa.qc.ca
- **Musée d'art contemporain** (Museum of Modern Art). Specializing in works dating from 1940 onwards, the museum hosts shows in all media used by contemporary artists. Metro stop: Metro Place-des-Arts -- www.macm.org
- **Canadian Centre for Architecture** The CCA is both a museum that hosts shows of all kinds related to architecture, and a major study centre for the discipline. Metro stop: Guy-Concordia -- www.cca.qc.ca
- **Musée d'archéologie et d'histoire de Montréal-Pointe-à-Callière** This ultra-modern building was constructed on the very site where the city was founded. Don't miss the chance to see the amazing multimedia show or to marvel over the artifacts on display in the archeological crypt beneath the museum. Metro stops: Place D'Armes, Square-Victoria.
- **Centre d'histoire de Montréal** (Montréal Museum of History) A deliberately tactile and multimedia experience of what Montréal has been like at different eras of its history. Metro stops: Place d'Armes or Square-Victoria. - www2.ville.montreal.qc.ca/chm/eng/accueil.shtm

Churches

Montréal is known as the city of a hundred steeples. You won't have to look far to find numerous beautiful churches. The following are a few of the churches most popular with tourists.

- **St. Joseph's Oratory.** The dome is the world's second largest, after St. Peter's in Rome. Religious visitors sometimes climb the steps (often on their knees) in the middle, praying at every step.
- **Notre Dame Basilica in Old Montréal.** A masterpiece of neo-Gothic architecture built between 1824 and 1829. It is very colourful and offers a sound and light show in the evenings. It houses a notable Casavant organ and its big bell, le Gros Bourdon, is the largest bell on the continent.
- One church that is etched in the memory of every Quebecer over the age of 45 is **Mary Queen of the World**. In the 1950's the Cardinal said the rosary from that church, and it was played over the radio all over the province. Families would stop everything and kneel by the radio every night at 7.
- **Saint Patrick's Basilica** reminds us of the large Irish population in Montréal. Built in 1847, the tall, elegant Gothic structure has been extensively restored in recent years.

Eat 'Til Your Heart's Content

Montréal's vast cultural diversity can perhaps only be rivaled by its wide variety of gastronomical fare. Home to 3000+ restaurants, whether you're in the mood for hot-dogs and "poutine" (unsure? ask a Montréaler!) or haute-cuisine, you'll find it here. For a comprehensive listing of restaurants visit www.toutmontreal.com/english/eguide/restaurants/restaurants.html#reas

NCME 2005 ANNUAL MEETING SELECTED PROGRAM HIGHLIGHTS MONTREAL, QUEBEC, CANADA, APRIL 12-14, 2005 By Julie P. Noble and Xiaohong Gao, Program Co-Chairs

Presidential Address

Measurement 101: Some Fundamentals Revisited
David Frisbie, University of Iowa/Iowa Testing Programs

Career Award Address

Making Sense of Causal Inference in Program Evaluation and Policy Research
Paul W. Holland, 2004 NCME Career Award Recipient
Moderator: Rebecca Zwick, University of California-Santa Barbara;
Discussant: Steve Raudenbush, University of Michigan

The Use of Writing Assessment Results in High-Stakes Admissions Testing Programs – Invited Symposium

Organizer: Catherine Welch, ACT; Moderator: Tim Burden, ACT
Larry Rudner, Graduate Management Admissions Council,
Catherine Welch, ACT
An evaluation of Intellimetric™ essay scoring system using responses to the GMAT® Analytical Writing Assessment
Ellen Julian, Association of American Medical College
MCAT writing sample today and speculation on possible tomorrows
Peter Pashley, Law School Admissions Council
Assessing writing for law school admissions: From current writing sample to plans for a scored writing assessment
Wayne Camara, College Board
Development, use and impact of a writing test for college admissions
Catherine Welch, ACT
The design, development, use and impact of an optional writing test
Discussant: Mark Shermis, Florida International University

Future Directions for Performance Assessment – Invited Symposium

Organizer: Suzanne Lane, University of Pittsburgh; Moderator: Carol Parke, Duquesne University
Suzanne Lane, University of Pittsburgh
Status and future directions for performance assessments in education
Richard Lehrer, Vanderbilt University
Future directions for designing cognitively rich and developmentally appropriate performance assessments
Randy Bennett, Educational Testing Service
Current issues and future directions in using computers for performance assessment
Richard Patz, Stanford University
Advances in psychometric methods for performance assessment
Discussant: Robert Linn, University of Colorado-Boulder

Next: What Should Be Retained, Adjusted, or Scrapped in the Current Federal Education Policy? – A Panel Discussion – Invited Panel

Organizer and Moderator: Ellen Forte Fast, edCount,
James Popham, UCLA
Patty Sullivan, Council of Chief State School Officers
Brian Gong, National Center for the Improvement of Educational Assessment
Judith Levinson, Evanston School District, NATD
Ellen Forte Fast, edCount
Discussant: Ellen Forte Fast, edCount

William H. Angoff: The Man Behind the Method – Invited Symposium

Organizer: Stephen G. Sireci, University of Massachusetts-Amherst;
Moderator: Henry Braun, Educational Testing Service

Stephen G. Sireci, University of Massachusetts-Amherst
No modifications necessary: Some reflections on Dr. Angoff

John Fremer, Caveon Test Security
The unpublished paper

Linda Cook, Educational Testing Service
Establishing comparable scores on tests given in different languages: Methodological contributions by William Angoff

Howard Everson, College Board
Doing psychometrics wearing white gloves and other lessons from Bill Angoff

Kurt Geisinger, University of St. Thomas
Bill Angoff and the so-called Angoff technique of standard setting

Daniel Eignor, Educational Testing Service
Bill Angoff and non-IRT test equating

Approaches to Accountability and Modernization in Assessment Practices Across Europe – Invited Symposium

Organizer/Moderator: Chris Whetton, National Foundation for Educational Research

Per-Erik Lyrén, Umeå University, Sweden
Approaches to accountability and modernisation in assessment practices in Sweden

Deborah Chetcuti, Grace Grima, University of Malta
The Maltese SEC (Secondary Education Certificate) Examinations: A system of differentiated papers

Gunter Maris, CITO International, Arnhem, The Netherlands,
Galina Kovaleva, Federal Institute of Pedagogical Measurement, Moscow, Russia

Central examinations in Russia and the Netherlands: A comparison

Henryk Szaleniec, Regional Examination Board, Krakow, Poland

External examinations in Poland - Extended response questions assessment.

Discussants: Eduardo Cascallar, Assessment Group International, Brussels, Belgium; Jannette Elwood, Queen's University, Belfast, Northern Ireland

How are the 1999 Standards Doing in 2005? – A Panel Discussion – Invited Panel

Organizer and Moderator: Daniel Eignor, Educational Testing Service

Michael Zieky, Educational Testing Service
Michael Kolen, University of Iowa/Iowa Testing Programs
Margaret Jorgensen, Harcourt Assessments
Jamal Abedi, UCLA/CRESST
Michael Kane, National Conference of Bar Examiners

Multi-stage and Multiple Fixed Form CBT Design Models, Quality Control, and Support Systems – Invited Symposium

Organizer/Moderator: Richard M. Luecht, University of North Carolina-Greensboro

David B. Swanson, Brian E. Clauser, Gerard F. Dillon, Kathleen Z. Holtzman, National Board of Medical Examiners
Operational testing systems for USMLE Step 3

Gerald Melican, Krista Breithaupt, Craig Mills, AICPA
Multi-stage testing and case studies in a fully-functioning licensing examination

Richard M. Luecht, University of North Carolina Greensboro
Some useful cost-benefit criteria for evaluating computer-

based test delivery models and systems

Dmitry I. Belov, Law School Admission Council, Ronald D. Armstrong, Rutgers University

A monte carlo approach for evaluating and designing multi-stage adaptive tests

Discussants: Stephen G. Sireci, University of Massachusetts-Amherst; April L. Zenisky, University of Massachusetts-Amherst

Current Guidance for Integrity in Testing – Symposium (Sponsored by the National Association of Test Directors)

Organizer/Moderator: Peter Hendrickson, Evergreen Public Schools
Karen E. Banks, Wake County Public School System

A conceptual framework for judging ethical violations and administering sanctions

Gregory J. Cizek, University of North Carolina-Chapel Hill
Personal and systemic influences on integrity in testing

James Impara, Caveon Test Security, G. Gage Kingsbury, Northwest Evaluation Association

Cheating detection within computer adaptive tests

Discussant: Joe O'Reilly, Mesa Public Schools

Trials and Tribulations of Transitioning from Measurement Theory to Practice – Symposium (Sponsored by the Graduate Student Issues Committee)

Organizers/Moderators: Jeffrey B. Hauger, University of Massachusetts, Olesya Falenchuk, University of Toronto.

Gregory J. Cizek, University of North Carolina-Chapel Hill
A look the other way: From measurement practice to theory

Edward H. Haertel, Stanford University
High stakes tests as policy tools: Dilemmas of theory and practice

Stephen G. Sireci, University of Massachusetts-Amherst
Measurement problems revisited

Wim J. van der Linden, University of Twente
Multiple categories, steps, dimensions, and levels in item response theory

Diversity Issues in Measurement and Policy Impacting High School Graduation and University Admissions – Symposium (Sponsored by the Diversity Issues and Testing Committee)

Organizer/Moderator: Concepcion M. Valadez, UCLA

Dan Solórzano, University of California-Los Angeles
Access, measurement, and the educational pipe-line

Rachel F. Morán, University of California-Berkeley
Educational equity and high stakes testing legal issues

Jorge Chapa, Indiana University
Educational policy implications of demographic trends

Discussant: Richard Valencia, University of Texas-Austin

Continuing to Measure Progress of Students and Schools Under the No Child Left Behind Act: Policymakers and Measurement Professionals Working Together – Symposium (Sponsored by the NCME Outreach and Partnership Committee)

Organizer/Moderator : Marianne Perie, Educational Testing Service
Marianne Perie, Educational Testing Service

A summary of what the measurement community is currently providing states in their state reports and how that has changed under NCLB

Mary Yakimowski, Baltimore City Public Schools
Experiences of a large, urban district: What are some characteristics of good testing that can make a positive impact on instruction?

Gerald Tirozzi, National Association of Secondary School Principals

What information are schools currently receiving from the measurement community regarding instruction and how does that compare to what they want to see?

AASA representative

Describing an ideal testing situation: How would schools and states work with measurement professionals to develop the most informative tests?

Discussant: Scott Young, National Conference of State Legislatures

Classroom Assessment Showcase Poster Session (Sponsored by the NCME Classroom Assessment Award Committee and the ATI Foundation)

Organizer: Rick Stiggins, Assessment Training Institute

Approximately 15 Montreal area teachers will present their classroom assessment work at a coordinated poster session. The teachers will be honored for their work at the NCME Breakfast.

Graduate Student Poster Session (Sponsored by the NCME Graduate Student Issues Committee)

Organizers: Daniel Sass, University of Wisconsin-Milwaukee; Samantha Burg, University of North Carolina-Chapel Hill

1. Andrew J. Poggio, University of Iowa
Revisiting the item format question: Can the multiple choice format meet the demand for monitoring higher order skills?
2. Ying Lu & Songbai Lin, University of Massachusetts-Amherst
Assessing fit of item response theory models
3. Haniza Yon & Mark D. Reckase, Michigan State University
A multidimensional linking of science achievement tests
4. Catherine Turcotte, Sarah Plouffe & Robin D. Tierney, University of Ottawa
Relevant teaching practices and the achievement of Ontario's French-language minority students on the 2001 Progress in International Reading Literacy Study (PIRLS)
5. Sarah Plouffe, Robin Tierney, and Catherine Turcotte, University of Ottawa
Impact of classroom teaching and assessment practices and the use of technology on the achievement of French-language minority students in Ontario on the 1999 TIMSS-R.
6. Jodi Herold McIlroy, University of Toronto
Examining variability in standard setting judges' modeled policies using hierarchical generalized linear modeling (HGLM)
7. Tia Corliss, Phoebe Winter, & Rebecca Kopriva, University of Maryland-College Park
Using focus group results as evidence for validity studies
8. Thomas A. Schmitt, Cindy M. Walker, & Daniel A. Sass, University of Wisconsin-Milwaukee
The impact of speededness on ability estimation within a computer adaptive testing environment
9. Yue Zhao, University of Massachusetts-Amherst
Using multidimensional models to improve diagnostic score reporting for large scale tests
10. Hua Wei, Weihua Fan, Yi Cao, & Robert Mislevy, University of Maryland-College Park
A Bayesian estimation approach to investigate access and accommodations in mathematics assessment
11. Jie Lin & W. Todd Rogers, University of Alberta
Validity of simultaneous approaches to the development of equivalent achievement tests in English and French
12. Leslie R. Odom, University of North Texas
Using technology in measurement curricula: An application of Microsoft® Excel for all things reliability
13. Michaela Gelin, University of British Columbia
Operating characteristics of the MIMIC approach for differential item functioning: Joreskog's estimation method with ML and WLS
14. Xuan Tan, & Mark J. Gierl, University of Alberta
Using global and local DIF analyses to assess DIF across language groups
15. Semonti Basu, Carol Myford, & Christine Salisbury, Rice University & University of Illinois-Chicago

Evaluating the effectiveness of rating designs in assessing student work using a many-faceted Rasch model

16. Katherine A. Tibbetts, University of Hawai'i-Mānoa
The non-equivalence of normative representations of test results across versions
17. Lingyun Gao & Changjiang Wang, University of Alberta-Edmonton
Using five procedures to detect DIF for passage-based testlets
18. Peter Baldwin, University of Massachusetts
Considering parameter estimation with small samples using item response theory

**Wednesday, 5:45-7:30 AM Le Centre Sheraton Hotel Lobby
NCME Fitness Run/Walk**

Co-Directors: David O. Anderson, Educational Testing Service; Pierre Léveillé, Cross des Couleurs, Boutique Endurance, Montreal; Claude David, Cross des Couleurs, Boutique Endurance, Montreal

Run 5K or walk a 2.5K course in beautiful Mont Royal Park.

Meet in lobby of Le Centre Sheraton Hotel at 5:45 AM. Buses leave at 6:00 AM for the park, and will return participants to the hotel at 7:30 AM, in time for the NCME Breakfast.

PREREGISTRATION IS REQUIRED.

This event is made possible through the sponsorship of: Applied Measurement Professionals, Inc.; CTB/McGraw-Hill; Educational Testing Service; Measured Progress Inc.; National Evaluation Systems, Inc.; Pearson Educational Measurement; Riverside Publishing

NCME 2005 ANNUAL MEETING

TRAINING SESSIONS

By James A. Wollack, University of Wisconsin-Madison

NCME is happy to offer an exciting slate of training opportunities for this year's annual meeting. The workshops cover a wide variety of topics and intended audiences; many are aimed at professional researchers, some at practitioners, and others at graduate students. All NCME training sessions will occur on either the Sunday or Monday immediately preceding the conference. Fees for workshops range from \$7.50 to \$95. In many cases, textbooks are included with the registration fee.

Admission to training sessions is limited to ticket holders. Advanced registration for training sessions is strongly recommended. All advanced registration for the workshops must be done electronically in one of two ways. People may sign up on-line, as part of the conference registration process (through the AERA website). This is the best option for those still needing to register for the conference. If you have already registered for the conference, however, you will be unable to use the on-line registration system to sign up for workshops. Instead, please send an email to mkozak@expologic.com indicating the session you would like to attend. In-person registration will also be possible at the NCME registration area at the conference. However, space in these workshops is limited and is filling up fast. In-person registration will only be possible provided space in the workshops is still available. Courses are subject to cancellation for insufficient registration.

Brief abstracts of the sessions are given below. More detailed abstracts are provided at the NCME website.

Session 1: The Kernel Method of Observed Score Test Equating

Presenters: Paul W. Holland, ETS; Alina A. von Davier, ETS; Ning Han, ETS

Sunday, April 10, 8:30-5:00

Fee: \$95

The Kernel Method of Test Equating (KE) is a unified approach to test equating based on a flexible family of equipercenile-like equating functions that contains the linear equating function as a special case. We will cover the theory behind KE as well as its application to many common equating designs. Registration includes a copy of *The Kernel Method of Test Equating* (2003), by Dr. von Davier and Dr. Holland.

Session 2: Introduction to the Generalized Graded Unfolding Model and the GGUM2004 Software Program

Presenter: *James S. Roberts, University of Maryland*

Sunday, April 10, 8:30–5:00

Fee: \$60

This training session will provide an introduction to unfolding item response theory models, appropriate when individuals respond to questionnaire items on the basis of proximity rather than dominance. The session will emphasize models in the family known as “graded unfolded models,” including the generalized unfolding model (GGUM); the computer program GGUM2004 will also be introduced.

Session 3: Considerations in Setting Performance Standards

Presenters: *Marianne Perie, ETS; Michael Zieky, ETS; Mary Pitoniak, ETS*

Sunday, April 10, 9:00–5:00

Fee: \$65

This session will examine standard setting, and will address how to choose a method, which methods are being used, and how to know if the cut scores set for an assessment yield valid interpretations within the context of a particular testing program. Participants will complete exercises to thoroughly understand the Bookmark and modified Angoff methods. Significant time will be spent studying the validity of standard-setting procedures and the resulting cut scores. Participants will receive a series of articles and sample standard setting materials.

Session 4: Latent Variable Models for Cognitive Diagnosis

Presenters: *Jeffrey Douglas, University of Illinois; Hua-Hua Chang, University of Texas; Louis Roussos, University of Illinois; Andre Rupp, University of Ottawa; Jimmy de la Torre, Rutgers University; Curtis Tatsuoka, George Washington University; Kikumi Tatsuoka, Teachers College, Columbia University*

Sunday, April 10, 9:00–4:15

Fee: \$95

In this training session, several popular modeling and classification approaches to cognitive diagnosis will be discussed, including Rule Space and three conjunctive latent class models known as the DINA, NIDA, and Fusion models. Specific issues to be addressed include model estimation, goodness-of-fit, and computerized adaptive testing (CAT). The training session concludes with a discussion of practical implementation issues and the inferential limitations of the models.

Session 5: Accommodations—Measurement and Policy Perspectives: Elementary Through Postsecondary

Presenters: *Noel Gregg, University of Georgia; Martha Thurlow, University of Minnesota; Nicole Ofiesh, University of Arizona; Cara Cahalan, ETS; Linda Cook, ETS*

Sunday, April 10, 1:00–5:00

Fee: \$35

This session will discuss state-of-the art research on accommodations (e.g., different item functioning [DIF], portfolio assessment, extended time, and alternative text), national response to the accommodation of diverse learners, policy challenges, and critical measurement issues.

Session 6: Principles and Applications of Interrater Reliability

Presenter: *Jamal Abedi, CRESST/UCLA*

Sunday, April 10, 1:00–5:00

Fee: \$20

Different approaches for estimating interrater reliability will be introduced, their limitations will be discussed, and their applications will be illustrated. Also, generalizability of raters will be discussed using a two-facet rater \times task design. A copy of software (ITRS) performing interrater reliability and a two-facet-design generalizability developed by the author will be distributed to the audience.

Session 7: ABCs of Automated Essay Scoring

Presenters: *Mark D. Shermis, Florida International University; Scott Elliot, Vantage Learning, Inc.*

Sunday, April 10, 1:00–5:00

Fee: \$85

This session is designed to take participants in a step-by-step approach to creating and implementing a statistical scoring model for automated essay scoring (AES). The process will be illustrated using the Vantage learning product Intellimetric™ and MyACCESS. Aspects of both programs will be demonstrated to show how AES can be implemented for writing interventions and used as an instructional tool.

Session 8: Multidimensional Item Response Theory

Presenters: *Brian Habing, University of South Carolina; Amy G. Froelich, Iowa State University*

Sunday, April 10, 9:00–4:30

Fee: \$90

This session focuses on developing an intuitive understanding of the concepts and methods of unidimensionality assessment. The session will review the assumptions, introduce multidimensional IRT models, and discuss the common procedure for testing unidimensionality. DIMTEST, DETECT, and HCA-CCPROX methods will be examined in detail, with opportunities to try the procedures on real data. Participants will receive Sijtsma and Molenaar’s (2002) *Introduction to Nonparametric Item Response Theory*, and copies of the software and key articles. Participants are encouraged to bring a laptop computer.

Session 9: Generalizability Theory and Applications

Presenters: *Robert L. Brennan, University of Iowa; Xiahong Gao, ACT, Inc.*

Monday, April 11, 9:00–5:00

Fee: \$90

Generalizability theory enables an investigator to disentangle multiple sources of error measurements, hence is applicable to a broad range of measurement, evaluation, and testing studies that arise in education. This training session will enable participants to understand the basic principles of generalizability theory, to conduct relatively straight-forward generalizability analyses, and to interpret and use the results of such analyses. Mathematical and statistical foundations will be treated only minimally. Participants will receive a copy of Dr. Brennan’s book, *Generalizability Theory*.

Session 10: Fundamentals of Polytomous Response IRT Models and Applications

Presenters: *Ronald K. Hambleton, University of Massachusetts at Amherst; Lisa A. Keller, University of Massachusetts at Amherst*

Monday, April 11, 9:00–4:00

Fee: \$85

In this workshop, we will (1) describe several popular unidimensional polytomous response IRT models (e.g., partial credit, generalized partial credit, and graded response models), their assumptions, and address model parameter interpretations; (2) explain how to estimate model parameters and model fit; (3) describe available software and demonstrate its use (e.g., Parscale), and (4) describe applications to test development and equating using a real data example.

Session 11: Grading

Presenters: *Susan M. Brookhart, Duquesne University; James H. McMillan, Virginia Commonwealth University*

Monday, April 11, 9:00–12:00

Fee: \$45

This training session will emphasize how concepts in educational measurement apply in the context of grading. The goal of this session is to enlarge the repertoire of principles, procedures, explanations, and examples available to measurement specialists in the area of grading. Participants will receive a copy of Dr. Brookhart's book, *Grading*.

Session 12: Examining Validity Issues in State Accountability Systems

Presenters: *Brian Gong, National Center for the Improvement of Educational Assessment; Ellen Forte Fast, edCount; and Art Coleman, Holland & Knight*

Monday, April 11, 8:00–12:00

Fee: \$65

A school or district cannot simply assume that their accountability system is working, even if test scores rise after its implementation. Rather, the quality of the data on which decisions are based, the quality of the decisions, the effectiveness of the sanctions, interventions, and rewards, and the avoidance of negative unintended consequences must all be examined as part of an organized process for supporting the validity of the system. This session will clarify the reasons why this work is critical and provide a framework and examples to guide participants in the design and implementation of their own validation processes. Participants will receive a copy of the CCSSO publication, *A Framework for Examining Validity in State Accountability Systems*.

Session 13: Preparing Tests With Test Security in Mind

Presenters: *James C. Impara, Caveon; David Foster, Caveon*

Monday, April 11, 8:00–12:00

Fee: \$45

As the stakes associated with testing increase, so do the rewards for test fraud. Test fraud takes on many forms: cheating, and stealing test items to use in test prep courses or to sell on the Internet. This session provides insights and strategies to help develop tests that hinder those who would behave in fraudulent ways. Strategies discussed will concentrate on both item development (for multiple-choice, performance-based, and computer-delivered items), and test assembly.

Session 14: A Nonlinear Mixed Models Approach to IRT

Presenters: *Paul De Boeck, K. U. Leuven, Belgium; Mark Wilson, UC Berkeley; Frank Rijmen, K. U. Leuven, Belgium; Francis Tuerlinckx, K. U. Leuven, Belgium*

Monday, April 11, 9:00–2:45

Fee: \$65

This session will begin by discussing the value of viewing IRT models as extensions of linear regression models in the case where the data are repeated categorical observations. Then we will discuss and illustrate how multilevel modeling and structural equation modeling (SEM) for categorical data, and other less well known

models can be expressed from the perspective of nonlinear mixed modeling and vice versa. We will illustrate how the models can be estimated with the SAS procedure NLMIXED and WinBUGS. Last, some new models will be presented to illustrate the flexibility of the approach.

Session 15: Tips for Graduate Students: Advice for Finishing School, Obtaining a Job, and Starting a Career

Presenters: *Deborah J. Harris, ACT, Inc.*

Monday, April 11, 8:00–12:00

Fee: \$7.50

The training session has three main components: (1) Finishing up the Ph.D., including finding a dissertation topic and maximizing experiences while still a student, (2) Getting a job, including finding available jobs, applying for jobs, interviewing; and (3) Strategies for beginning a career, including job politics, adjusting to the environment, career path, publishing, professional service, being/finding a mentor, balancing work and life, and what if I hate my job.

Session 16: Teaching Educational Statistics: A Guided Discussion of 10 Propositions

Presenters: *Ronald C. Serlin, University of Wisconsin; Michael A. Seaman (moderator), University of South Carolina; Patricia Busk, University of San Francisco; Sharon Weinberg, New York University; and Gabriella Belli, Virginia Tech*

Monday, April 11, 2:15–5:15

Fee: \$20

The presenters in this session will interact with one another and with participants in a discussion of 10 propositions that could fundamentally govern instructional decisions in the preparation and teaching of educational statistics courses. This session will not be a how-to session, but will focus on basic principles for teaching educational statistics that are derived from theory, research, and the experiences of the presenters.

Session 17: The History of Educational Testing in the United States

Presenter: *H. D. Hoover, University of Iowa*

Monday, April 11, 2:15–6:15

Fee: \$7.50

This session will focus on the history of achievement and ability testing in the United States, with emphasis on the evolution of large-scale achievement testing during the 20th century and its impact on American schools. Included will be the early development of standardized achievement tests and the reasons for their widespread use, Title I, the minimum competency testing movement, and the use of tests for political purposes, as well as histories of college admission testing, ability testing in schools, and NAEP.

Session 18: Planning and Developing a Testing Program

Presenters: *Stephen Downing, University of Illinois at Chicago; Thomas Haladyna, Arizona State University*

Monday, April 11, 2:15–6:15

Fee: \$40

This training helps participants learn to develop a high-quality testing program that measures educational achievement in the schools or for credentialing. A 12-step process forms the basis for the planning of a testing program. Participants are expected to put into practice concepts, principles, and procedures learned from this training session to create a test plan that can be used to develop a testing program and enable the collection and organization of validity evidence.