



URBAN DISTRICTS TAKE A CHANCE: SIGNIFICANT IMPROVEMENTS AT THE FOURTH GRADE

By Sharon Lewis, Council of the Great City Schools

The Council of Great City Schools approached the National Assessment Governing Board in 2000 with the idea of a trial urban NAEP. Urban districts volunteered for several reasons: (1) they are committed to improving student achievement; (2) they are not afraid of being measured against the highest possible standards; (3) they want to be able to compare their students with other students in large urban cities; (4) they are confident that the reforms that have been in place for the past five to seven years are being to take hold and that real progress is being made.

In 2002 five urban districts – Atlanta, Chicago, Houston, Los Angeles, New York – volunteered to participate in the National Educational Assessment Program (NAEP) in reading and writing. Participation in the Trial Urban District Assessment (TUDA) meant that, for these districts, sample sizes for NAEP would be increased to ensure valid reporting. Although TUDA results would be reported out at a different time, data for these districts would be reported in the same manner as state and national NAEP. In 2003 four additional cities – Boston, Charlotte-Mecklenburg, Cleveland, San Diego – volunteered to participate in NAEP reading and mathematics tests. Because the District of Columbia is a state and a local district, their results are also included in TUDA. Contrary to what many may think, participation in TUDA is restricted primarily by the limited money available from the National Center for Educational Statistics (NCES) to fund the additional assessments. The Council has received requests from many more urban districts eager to participate in TUDA.

On the surface, TUDA results are what one might expect. However, upon closer inspection, there are a few surprises. In general, urban students are performing significantly lower than national averages in reading and math. Conversely, fourth and eighth graders in one district, Charlotte-Mecklenburg, scored as well or significantly higher than students nationally in math and reading on the 2003 TUDA.

Additionally, achievement gaps in urban centers appear to mirror, with few exceptions, those in the nation. On the other hand, for six of the ten urban districts – Atlanta, Charlotte, DC, Houston, New York City and San Diego -- the 2003 TUDA reading results show that their percentages of White fourth graders who scored at or above proficiency levels is the same or higher than White students nationally. African American fourth graders in Charlotte, Houston, and New York City scored at or above proficiency levels at the same or greater rates than African American students nationally. Hispanic students in Charlotte, Cleveland, Houston, and New York City also outpaced their peers nationwide.

It is also important to note that although their scores are lower, urban students appear to be closing the gap with their peers nationally. Grade 4 reading scores for students in large central cities increased significantly from 2002 to 2003 compared to no discernible change for students nationally; grade 8 reading scores dipped slightly for both large central cities and the nation. Large Central City was first

reported by NCES in 2003 and is defined by the US Census as a central city with a general population of at least 250,000. There are approximately 67 cities that meet this criterion.

Results of the NAEP 2003 Trial Urban District Assessment Percent of Students At or Above Proficient Reading Grades 4 and 8

	Grade 4		Grade 8	
	2002	2003	2002	2003
Nation	30	30	31	30
Large Central City	17	20	20	19
Atlanta	12	14	8	11
Boston	NA	16	NA	22
Charlotte	NA	31	NA	30
Chicago	11	14	15	15
Cleveland	NA	09	NA	10
DC	10	10	10	10
Houston	18	18	17	14
Los Angeles	11	11	10	11
New York City	19	22	--	22
San Diego	NA	22	NA	20

NA = not applicable. District did not participate that year.

-- Insufficient number of NYC students participated for scores to be valid.

Results of the NAEP 2003 Percent of Students At or Above Proficient Trial Urban District Assessment Math Grades 4 and 8

	Grade 4	Grade 8
	2003	2003
Nation	31	27
Large Central City	21	17
Atlanta	13	6
Boston	12	17
Charlotte	41	32
Chicago	10	9
Cleveland	10	6
DC	7	6
Houston	18	12
Los Angeles	13	7
New York City	21	20
San Diego	20	18

Additional information on TUDA may be found at www.nces.ed.gov.

ANNUAL MEETING INFORMATION INSIDE!

See pages 5 – 8 for details about the NCME Annual Meeting in San Diego, April 13-15, 2004.

SAN DIEGO

By Steven Baratte, San Diego City Schools Communications Department

In addition to perfect weather year-round, San Diego offers wonderful dining opportunities, cultural resources and recreational activities.

Gaslamp Quarter

The Gaslamp Quarter is located in the heart of Downtown San Diego, just steps away from the Convention Center. It's eight blocks long (from Broadway to Harbor Drive) and two blocks wide (from Fourth to Sixth Avenues). The Gaslamp Quarter offers premier dining, shopping and entertainment district, where you'll find a truly eclectic blend of food, fun and culture.

The Gaslamp Quarter Association
614 Fifth Avenue, Suite E
San Diego, CA 92101
(619) 233-5227
www.gaslamp.org.

Balboa Park

The Park covers 1,200 acres and is minutes north from Downtown San Diego. There are more than 85 cultural and recreational organizations, including fifteen museums and various performing arts groups, like the Marie Hitchcock Puppet Theatre and the world-famous Old Globe Theatre, which presents at least 14 productions and 550 performances a year.

Balboa Park Visitors Center
1549 El Prado, Suite #1
San Diego, CA 92101
(619) 239-0512
www.balboapark.org

Old Town

Dating back to the time of the early Spaniards in California, Old Town reflects the rich and colorful history of early California. Shows, festivals, dining, artisans, and a wealth of shops can all be found here. Take a walk through Bazaar Del Mundo, which is composed of small shops, and Mexican cuisine that captures the taste of Mexico. Don't forget to stop and enjoy a world-famous Margarita.

Old Town Chamber of Commerce
2461 San Diego Ave #202
San Diego CA 92110
(619) 291-4903
www.oldtownsandiego.org

Mission and Pacific Beaches

Mission Beach is in the center of a long stretch of white sand that is locally known as "the Strand," and extends for more than two miles. There is a host of amenities and restaurants in the area, making this the happening place year-round. The boardwalk consists of a narrow concrete ribbon of cement that runs the length of the strand, allowing ocean-side cycling, jogging, roller blading or just walking.

Mission Beach and Boardwalk
3190 Mission Blvd
Mission Beach

Just north of Mission Beach is Pacific Beach, with eclectic shops and eateries. If you are looking for casual nightlife, look no further. Pacific Beach has numerous bars and clubs located on Garnet Avenue, just a block away from the sandy beach.

Pacific Beach
www.pacificbeach.org

La Jolla

La Jolla is located 15 minutes from downtown San Diego. It has wonderful beaches, cultural activities and fine restaurants. While La Jolla is known to be one of the most affluent communities in the

United States, it has a down to earth sense due to the beautiful natural scenery and the helpfulness of its residents.

www.lajolla.com

Shopping

Horton Plaza

Horton Plaza is located next to the Historic Gaslamp Quarter in the center of Downtown San Diego. The outdoor plaza offers more than 140 specialty shops and restaurants. There is plenty to see and shop, from the major chains like Macy's and Nordstrom to many small and unique shops that make up this wonderful mall.

Horton Plaza
324 Horton Plaza
San Diego, CA 92101
(619) 239- 8180

Fashion Valley Mall

Fashion Valley Mall is located approximately 10 minutes away from Downtown San Diego and features more than 200 stores and restaurants and an 18-screen AMC theatre.

Fashion Valley Mall
7007 Friars Road
San Diego, CA 92108
(619) 688-9113

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.

Carlsbad Outlets

Carlsbad Company Stores is located approximately 30 minutes north of downtown San Diego. There are some of the best names in designer fashions, home furnishings, footwear, jewelry, toys and accessories, all at up to 70% off everyday.

Carlsbad Company Stores

5600 Paseo del Norte

Carlsbad, CA 92008

(888) 790-SHOP

Other Attractions

San Diego Zoo

The world famous San Diego Zoo is located just north of Downtown San Diego in Balboa Park. You can stroll around the 100-acre grounds, discovering colorful and exotic species of animals displayed in spacious natural habitats.

San Diego Zoo

2920 Zoo Drive

San Diego, CA 92101

(619) 234-3153

www.sandiegozoo.org

San Diego Wild Animal Park

Although approximately 45 minutes north of Downtown San Diego, the Wild Animal Park is well worth the trip. With a 32-acre Heart of Africa walking safari, Wgasa Bush Line Railway, Lorikeet Landing, Nairobi Village, and Condor Ridge, a visit to the San Diego Wild Animal Park is like a safari to many of the world's most exotic places.

San Diego Wild Animal Park

15500 San Pasqual Valley Road

Escondido, CA 92027

(760) 747-8702

Sea World

Sea World offers an amazing array of marine exhibits, shows and rides that will entertain even the most serious member of the family.

SeaWorld

Sea World Adventure Park

500 SeaWorld Drive

San Diego, CA 92109

(619) 226-3901

www.seaworld.com

LegoLand

With rides, attractions to activate, shows and building projects in all six themed "blocks," LegoLand Park mixes education, adventure and fun in this first park of its kind in the United States.

LegoLand California

One LegoLand Drive

Carlsbad, CA 92008

(760) 918-5346

www.lego.com

A PARADOX: SEPARATE CLASSROOM AND STANDARDIZED ASSESSMENT TO FIND SYNERGY BETWEEN THEM

By Rick Stiggins, Assessment Training Institute

To find the synergy between classroom and standardized assessment, we must distinguish between assessment used *to promote or cause learning* and *to verify that learning has occurred*. In our work at ATI, we follow the lead of the Assessment Reform Group of the UK, by labeling these uses "assessment FOR learning" and "assessment OF learning" respectively. Both are important, but they are different. We achieve synergy by making sure they work in harmony.

Examples of assessments OF learning include standardized tests used for accountability reporting, as well as classroom assessments

teachers use to determine achievement status for report card grading. These are events that happen after learning is supposed to have occurred. They are intended inform others about students. Assessments FOR learning are those we use to diagnose student needs along the journey to standards and to help students see, understand, and feel in control of their own academic success as they grow. They happen during learning and are intended to inform students about themselves. We achieve synergy when we understand that different assessment users need different information in different forms and at different times to do their job. Some uses are meant to help students grow, while others need to prove they did. We achieve synergy when we balance assessments FOR and OF learning to serve those different, but complimentary, purposes.

One gauges achievement status at a particular point in time, while the other seeks to change achievement status over time. Thus, one is driven by the concept of accountability, while the other is driven by the desire for improvement. They are in harmony when we use assessment to bring students to success and then use it to verify their arrival. In this sense, one asks if students have met standards, while the other asks if they are making progress toward—are climbing the scaffolding to—those standards. So both arise from the same standards, but they treat them differently.

We achieve synergy when we rely on a continuous flow of student-involved classroom assessments FOR learning to keep all student striving for success and then turn to periodic assessments OF learning celebrate their achievements. When we find that balance, remarkable gains in student achievement result, especially for low achievers. Yet balance continues to elude us because, while we have a long history of investment in layer upon layer of assessments OF learning, we have virtually no record of investment in classroom assessment FOR learning. Neither teachers nor school leaders have been given the opportunity to learn how to apply principles of assessment FOR learning. We will tap the immense power of assessment as a tool for school improvement only when we deliver those tools into the hands of practitioners. Only then will we discover the synergy between the two.

To receive a free public service video or DVD entitled "Assessment FOR Learning: A Hopeful Vision of the Future," describing the synergy between assessment OF and FOR learning, contact ATI at 800-480-3060.

SUPPLEMENTAL STATE AND LOCAL ACCOUNTABILITY SYSTEMS: WITHIN AND BEYOND NCLB

By Stanley Rabinowitz, WestEd

The No Child Left Behind Act (NCLB) mandates several indicators to determine school effectiveness. NCLB-sanctioned accountability systems judge success based on the percentage of students meeting state-determined performance standards; this percentage increases at set intervals until all students are expected to master rigorous state content standards by 2014. Schools must not only improve overall but must show significant increases for numerous disaggregated student subpopulations.

NCLB's definition of school success is very specific and relatively narrow. Students must demonstrate mastery of English/language arts and mathematics standards (with a bit of science sprinkled in) from grades three through eight and high school. The NCLB model places primary emphasis on the status of performance, rather than on improvement. Finally, all NCLB accountability decisions are conjunctive. Designations of "failure" may result from the performance of a single student subpopulation on a single content area.

The assumptions behind the NCLB accountability model are inherently neither right nor wrong—they simply reflect one way to define success. Other defensible accountability systems have been developed over the past decade using growth or value-added components as the primary means of judging school effectiveness, rather than status. In addition, accountability models have included other content areas (e.g., social studies, arts) and a broader range of noncognitive indicators.

While NCLB-sanctioned accountability systems permit additional content areas and classes of indicators into the determination of Adequate Yearly Progress (AYP), the conjunctive nature of NCLB accountability decisions limits states' willingness to incorporate them, even though they may be highly valued important measures of school and student success. In addition, as the number of schools and districts identified as in need of improvement increases over time based on the narrow NCLB definition of success, policy makers are seeking supplemental school classification systems to expand the public's perspectives as to what defines an effective school. The goal is **not** to overrule or undermine the AYP decision; rather, it is to provide additional information the public can use to judge the success of public schools.

States and school districts that are moving to augment NCLB requirements with supplemental models and indicators of school success either incorporate local models directly into the NCLB AYP decision or use alternative models side by side with NCLB AYP. These supplemental systems include information in four broad categories: alternate models of school and student success, additional content areas to define success, additional assessment formats to assess these content areas, and additional classes of indicators beyond the NCLB primary use of test data.

The framework below consists of a set of general questions with supporting justifications for states, districts, and schools to use to determine if and how to supplement the NCLB accountability model. Following that is a list of potential barriers to supplementing an NCLB accountability model at the state or local level.

Supplemental Accountability Framework

- I. Why would a state/district/school supplement the NCLB accountability system?
 - *Values* – are there indicators valued by the local community not included in the NCLB model?
 - *More comprehensive views of success* – is a successful school defined more broadly than student performance on mandated E/LA, mathematics, and science tests?
 - *Alternate views of success* – does the state/district/school want to recognize different types of achievement (e.g., significant improvement)?
- II. What supplemental indicators may be considered?
 - *Alternate models of success* – growth models, value added approaches
 - *Additional content areas* – traditional (e.g., social studies) and innovative (workplace readiness)
 - *Additional assessment formats* – locally developed, performance assessments, portfolios
 - *Additional classes of indicators* – behavioral, school climate, AP enrollment, locally developed and defined
- III. How should a state/district/school decide whether to add supplemental indicators?
 - *Values* – does the combined accountability system (NCLB and supplemental) more fully reflect community views of a successful school?

- *Reliability* – does the combined accountability system (NCLB and supplemental) increase the reliability of the school accountability decision?
- *Validity* – does the combined accountability system (NCLB and supplemental) increase the validity of the school accountability system?
- *Value vs. burden* – does the value in supplementing the NCLB accountability system outweigh the burden in collecting, analyzing, reporting, and communicating the additional information?

Barriers to Implementation of a Supplemental Accountability System

States/districts/schools considering supplementing their NCLB accountability systems must be prepared to overcome the following barriers:

- *Single Statewide System* – NCLB requires states to develop one system to evaluate school performance. This expectation should not be interpreted to prohibit supplementation. Policy makers may either attempt to blend the systems into one single AYP decision or use the additional information side by side with the AYP designation. In either approach, federal sanctions will be based on the AYP decision, though they may be mitigated in the arena of public opinion by an alternate supplemental performance label.
- *Confusion of Multiple Labels* – The situation certainly exists (and is often meaningful and desirable) where AYP and supplemental designations may differ. Policy makers will need to invest in educating the public about the differing assumptions behind each designation and explaining why awareness of both is important.
- *Time and Resources* – Development and implementation of a supplemental accountability system is not free. Staff must be designated to lead all components; resources must be found to develop new indicators, and report and explain the results. Hopefully, the value of the more comprehensive picture of school success, based on a fuller range of indicators that more fully reflect local values, will justify the burden of this extra investment.

NEWSLETTER TO GO ALL-ELECTRONIC

In October, 2003, the NCME Board voted to make the Newsletter available in electronic format only, starting with Vol. 12, No. 2 (June, 2004). Issues of the Newsletter will be posted on the NCME website (www.ncme.org). Notice of new issues will be sent via the NCME Listserv. Information about subscribing to the listserv is also on the website.

NCME ANNUAL MEETING TRAINING SESSIONS
SAN DIEGO, CA, APRIL 12-14, 2004
Allan S. Cohen, Training Session Chair

Admission to training sessions is limited to ticket holders. Tickets may be obtained by writing to Training/NCME, 1230 17th Street, NW, Washington, DC 20036-3078. Please enclose payment and a self-addressed stamped envelope. Courses are subject to cancellation for insufficient registration. Some courses require advanced registration. The deadline for ordering tickets is March 15, 2004. Any tickets available at the conference will be sold on a first-come, first-served basis at the registration area in one of the convention hotels.

Session 1: The Kernel Method of Observed Score Test Equating

Presenters: Paul W. Holand, ETS; Alina A. von Davier, ETS; Dorothy T. Thayer, ETS

Monday, April 12, 9:00 to 5:00

Fee:\$95

Session 2: Graphical Models in Educational Assessment

Presenters: Russell G. Almond, ETS; Robert J. Mislevy, University of Maryland; David M. Williamson, ETS; Duanli Yan, ETS

Monday, April 12, 9:00 to 5:00

Fee:\$95

Session 3: Test Equating Methods and Practices

Presenters: Michael J. Kolen, University of Iowa; Robert L. Brennan, University of Iowa

Monday, April 12, 9:00 to 5:00

Fee:\$95

Session 4: Writing Technical Documentation for a Large-Scale Assessment Program

Presenters: Huynh Huynh, University of South Carolina; J. Patrick Meyer, University of South Carolina; Karen Barton, CTB/McGraw-Hill

Wednesday, April 14, 1:00 to 5:00

Fee:\$25

NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION
2004 ANNUAL MEETING SELECTED PROGRAM HIGHLIGHTS
SAN DIEGO, CA, APRIL 13-15, 2004

Carol S. Parke and Susan M. Brookhart, Program Co-Chairs

Presidential Address

Validity of High-Stakes Assessment: Are Students Engaging in Complex Thinking?

Suzanne Lane, NCME President

Career Award Address

Cognitive Psychology and Educational Assessment

Robert J. Mislevy, 2003 NCME Career Award Recipient

Moderator: John Mazzeo, ETS; Discussant: Robert Glaser, University of Pittsburgh

Meshing Measurement with Curriculum and Instruction: Three Venues, Three Strategies – Invited Interactive Symposium

Organizer/Moderator: W. James Popham, University of California, Los Angeles

Duncan MacQuarrie, Harcourt Educational Measurement

Collaboration at the district level: It can happen, but it depends...

Doug Christensen, Nebraska Commissioner of Education

It takes a team: Going beyond cooperation and consensus

Eva L. Baker, CRESST/UCLA

Building social capital for educational reform: A university perspective

Vertically Moderated Standards: Assumptions, Case Studies, and Applications to School Accountability and NCLB Adequate Yearly Progress – Invited Symposium

Organizer: Huynh Huynh, University of South Carolina; Moderator: Karen E. Barton, Research Triangle Institute

Huynh Huynh, Christina Schneider, University of South Carolina

Vertically moderated standards as an alternative to vertical scaling: Assumptions, practices, and an odyssey through NAEP

Daniel M. Lewis, CTB/McGraw-Hill; Carolyn Haugh, Boulder Valley School District

A standard setting odyssey: On a quest for across-grade consistency

Steve Ferrara, Eugene Johnson, Wen-Hung Chen, American Institutes for Research

Vertically moderated standards: Logic, procedures, and likely accuracy of judgmentally articulated performance standards

William D. Schafer, University of Maryland

State perspectives on moderation of standards: Technical recommendations and policy considerations

Chad Buckendahl, Buros Institute; Huynh Huynh, University of South Carolina; Theresa Siskind, Joseph Saunders, SC Dept of Education

From content standards, through technical advisory committee, and to State Board of Education: A case study based on SC PACT 2003 assessments

Huynh Huynh, Dorinda Gallant, Sameano Porchea, University of South Carolina

Vertically moderated standards for SC PACT 1999 assessments: A look back from longitudinal student data

Discussants: Edward H. Haertel, Stanford University; Robert Lissitz, University of Maryland

Career Highlights and Contributions: Robert L. Ebel – Invited Symposium

Organizers: Gregory J. Cizek, University of North Carolina; David A. Frisbie, University of Iowa; Moderator: David A. Frisbie

William A. Mehrens, Michigan State University

A biographical sketch of Bob Ebel

Rick Stiggins, Assessment Training Institute

Robert Ebel's contributions to classroom measurement practices

Linda Crocker, University of Florida

Robert Ebel's contributions to measurement theory

Gregory J. Cizek, University of North Carolina

Bob Ebel: Educational statesman

Hierarchical Modeling of Social and Cognitive Assessment Data – Invited Symposium

Organizer/Moderator: Brian W. Junker, Carnegie Mellon University

Sophia Rabe-Hesketh, University of California, Berkeley

Generalized linear latent and mixed models

Cees A. W. Glas, Jean-Paul Fox, University of Twente, Netherlands

Analysis of variance and regression using multilevel IRT

Jeff Douglas, University of Illinois; Jimmy de la Torre, Rutgers University

Model selection in cognitive diagnosis

Matthew S. Johnson, City University of New York

A B-spline model for nonlinear factor analysis

Discussants: Kadriye Ercikan, University of British Columbia; Richard Patz, Stanford University

Evaluating State Accountability Systems: Validity and Reliability in the Context of NCLB – Invited Symposium

Organizer/Moderator: Ellen Forte Fast, edCount, LLC

Tom Deeter, Iowa Department of Education

Iowa's accountability system: What we know, what we've learned, and where to go from here

Juliane Dow, Massachusetts Department of Education

Massachusetts' approach to school and district accountability

Robin Jarvis, Louisiana Department of Education

The validity of Louisiana's multi-tier accountability system

Carina Wong, Pennsylvania Department of Education

The challenges of designing a valid accountability system in Pennsylvania

Discussants: Eva L. Baker, CRESST/UCLA; Scott Marion, Center for Assessment

What Does it Mean for Classroom Assessments to be Valid? Reliable? – Invited Symposium

Organizer: Dylan Wiliam, ETS

Pamela Moss, University of Michigan

Hermeneutics as a guide to validity in classroom assessment

Jay Parkes, University of New Mexico

What does it mean for classroom assessment to be valid and reliable – in classrooms of diverse students?

Jeff Smith, Rutgers University

Reconceptualizing reliability as sufficiency of information

Rick Stiggins, Assessment Training Institute

Using classroom assessment to overcome a legacy of mistaken assessment beliefs

Dylan Wiliam, ETS

Assessment and the regulation of learning

The Achievement Gap: Test Bias or School Structures – Symposium (Sponsored by National Association of Test Directors)

Organizer: Thel Kocher, Blue Valley Schools; Moderator: Judy Pfannenstiel, Research and Training Associates, Inc.

Steve Schellenberg, Saint Paul Public Schools

The historical context

Stephen G. Sireci, University of Massachusetts - Amherst

The role of sensitivity review and differential item functioning analyses in reducing the achievement gap

Margaret Jorgensen, Harcourt Assessment, Inc.

A test publisher perspective

Deb Lindsey, Milwaukee Public Schools

A district perspective

Discussant: Glynn D. Ligon, ESP Solutions Group

Leave No Child Behind?: Do We Have the Data Required to Make a Call? -- Symposium (Sponsored by the NCME Diversity Issues and Testing Committee)

Organizer/Moderator: Sharon Robinson, ETS

Martin Carnoy, Stanford University

Accountability and state NAEP results in math and reading: What can be inferred?

Chrys Dougherty, National Center for Educational Accountability

How better state data can lead to school improvement

Jun Choi, New Jersey Department of Education

NJ SMART – Standards Measurement and Resource for Teaching: Lessons from the frontline of a statewide effort to collect, analyze, and report quality education data

Ashaki Coleman, ETS

National data systems: Are they half empty or half full?

Discussants: Concepción Valadez, University of California, Los Angeles; Michael Nettles, ETS

The Promise and Perils of Innovative Assessment Designs – Symposium (Sponsored by the NCME Graduate Student Issues Committee)

Organizer/Moderator: William P. Skorupski, University of Massachusetts - Amherst

Barbara Plake, University of Nebraska, Lincoln

Standard setting with innovative assessment designs

Robert Mislevy, University of Maryland

The challenge of context

Brian Clauser, National Board of Medical Examiners

Performance assessment in high-stakes licensure examinations

Craig Mills, American Institute of Certified Public Accountants

Making sausage: Developing computer-based tests for operational assessment programs

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

Organizer: Rick Stiggins, Assessment Training Institute

Approximately 20-25 outstanding California 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: Maureen Ewing, Fordham University; Olesya Falenchuk, University of Toronto

1. Amral Sidiq Ali, University of Toronto
Measuring school improvement
2. Chanh Park, University of Wisconsin-Madison
A Monte Carlo study comparing parametric and nonparametric DIF detection methods for dichotomous items
3. Dan A. Sass, Thomas A. Schmitt, Cindy M. Walker, University of Wisconsin-Milwaukee
An evaluation of BILOG-MG with skewed theta distributions using various estimation procedures: A simulation study
4. Feifei Ye, Ohio State University; Russell Almond, ETS; Robert J. Mislevy, University of Maryland; Duanli Yan, ETS
Sensitivity to prior distributions in calibrating a Bayesian network for language assessment
5. Greg Sadesky, Matthew Gushta, University of Alberta
Applying Rule-Space Methodology to the problem of standard setting
6. Felipe Martinez, University of California, Los Angeles
Comparing classical and item response theories for use in college admissions
7. Kara M. Owens, James Madison University
The relationship between achievement goal orientation and item difficulty selection in a self-adapted test
8. Maria Papapolydorou, Robert J. Mislevy, University of Maryland
A Bayesian estimation approach to investigate test performance for interpreting and understanding graphs
9. Matthew Gushta, University of Alberta
Equivalence across modes of administration: An item-level analysis of computer-and paper-based test versions
10. Peggy K. Jones, Cynthia G. Parshall, John M. Ferron, University of South Florida
Preliminary computer-administered MCAT data versus paper-and-pencil MCAT: A comparability study
11. Pei-Hua Chen, University of Texas at Austin; Se-Kang Kim, Harcourt Educational Measurement
Comparison of three different linking procedures: Concurrent calibration, fixed item parameter, and mean/sigma scaling methods
12. Penny E. Nichol, University of Minnesota
Comparing alternative exam form structures
13. Qianli Ma, Amy B. Hendrickson, University of Maryland
Comparability of Maryland school performance assessment program scores across students with and without accommodations
14. Rebecca Gokiert, Kathryn Ricker, University of Alberta
A comparison of gender DIF on the WISC-III in Canadian and American national standardization samples
15. Shirley Y Y Li, University of Alberta
Assessing the effects of substantive and statistical DIF analyses on common-item nonequivalent group equating design in translated tests
16. Tammiee Dickenson, University of South Carolina
A comparison of ability estimates for locally dependent data using a standard IRT model and a testlet model
17. Weiwei Cui, James Roberts, Han Bao, University of Maryland
Data demands for the generalized graded unfolding model
18. Xia Mao, University of Iowa; Daniel M. Lewis, CTB/McGraw-Hill; Machteld Hoskens CTB/McGraw-Hill
A comparison of vertical scale stability using horizontal versus augmented anchor sets

Measuring 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 Partnerships Committee)

Organizer/Moderator: Ronald J. Dietel, CRESST/ UCLA and NCME Outreach and Partnerships Committee

Senator Dede Alpert, California State Senator

What state policymakers need from the measurement community

Terry Duggan Schwartzbeck, American Association of School Administrators; Geno Flores, Deputy Superintendent for Accountability, California

What school administrators need from the measurement community

Gerald Tirrozi, National Association of Secondary Principals

What state agencies need from the measurement community

Mark Reckase, Michigan State University and NCME

A master plan for how the measurement community can meet the needs of policymakers, administrators, and states

Wednesday, 5:50 – 7:30 AM Marriott Lobby

NCME Fitness Run/Walk

Co-Directors: David O. Anderson, ETS; Steven Taggart, Tagg Running Events

Run or walk a 5K/2.5K course along San Diego Bay in Embarcadero Marina Park. Meet in the lobby of the San Diego Marriott Hotel at 6:00 AM for signup. Group will leave promptly at 6:20 AM for the starting line. Runners/walkers can return by foot to their hotels after completing the course. Pre-registration is required.

This event is made possible through the sponsorship of: Applied Measurement Professionals, Inc.; CTB/McGraw-Hill; ETS; Harcourt Assessment; National Evaluation Systems, Inc.; and Riverside Publishing.