NCME

FROM THE PRESIDENT:
THE MEANING OF EDUCATIONAL MEASUREMENT SCALES

Mark D. Reckase, Michigan State University

This is the last newsletter article that I will be writing as president of NCME. This brings a certain amount of sadness because I have enjoyed working with all of the Board members and the NCME office staff. It is not that I am going very far away from the organization. I still get to be officially the past president for a year. After that, while I may still be past president, it will be more difficult to get me to do things. As parting official words, let me invite you all to become active within NCME. It is a great organization and working with colleagues in the field is very rewarding. I know that everyone is very busy, but this is your organization and it only meets the needs of the membership if the members are willing to help make that happen.
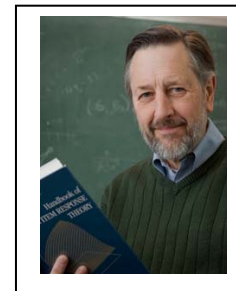
I hope to see you all in San Diego. I know that budgets are getting tight, but the meeting promises to be an excellent one. Be sure to take advantage of all it has to offer.

I am going to take advantage of this last newsletter article to make a few points about the way we estimate the achievement of examinees based on their responses to test questions. If you have strong reactions to my remarks, contact the newsletter editor. He will probably be happy to have more submissions to the newsletter and you can present your own perspective.

The Tyranny of the Number-Correct Score Scale

The number-correct scoring method has been around for a long time. It has the beauty of simplicity. What could be easier than counting up the number of correct responses? But over the last few years, I have been reflecting on the meaning and usefulness of the number-correct score or more generally on all scores that are the sum of item scores. The reflection on this topic is stimulated by the many times that I hear recommendations to convert IRT θ-estimates into estimated number-correct scores because the number-correct scores are supposed to be easier to understand. I am not sure that is true. I would like you to do a short thought experiment. Imagine that you have a room full of people and you ask them to count how many pieces of paper currency they have with them. Then, you order the people according to that count and state that the person with the highest count is the person with the greatest amount of money with them. Will this process give an accurate ranking of people on the amount of money that they have?

I hope that your answer is no. Obviously, the pieces of paper currency have different values and the count does not take that into account. If there is not a lot of variation in the values of the currency that the people in the room hold, the rank ordering will be pretty accurate. But if there is a lot of variation, the person with one $100 bill will be ranked much lower than the person with twenty $1's. Counting correct responses to test items has the same problem. Why is it that it makes sense to give an equal score for the correct response to every dichotomously scored test item? The items on a test are very likely different in difficulty and discrimination. It would seem that getting a good estimation of a person's achievement needs to take that into account. Item response theory models do that, but simple number-correct scoring models do not. Maybe the rough approximation given by the number-correct score was okay when people were not paying attention to every tiny change in test scores, but in the era of high stakes testing it seems that we need to move in the direction of getting the most accurate measure of location on a scale that we can. We should educate people to properly interpret a location on a scale rather than arguing that we need to use number-correct scores because they are the only thing that people in general can understand.

Logical Units for Achievement Test Scores

I wish that I had disposable wealth, so that I could offer a prize for a great scientific accomplishment. I am thinking about the prize that is being offered for the first car that gets 100 miles per gallon of gas, or the prize for the first human powered airplane to fly over some major body of water. Unfortunately, I don't have a million dollars to put up for such an endeavor, but if I did, I would give it for the first person to develop a logical measurement scale for school based achievement.

Those involved in large scale achievement testing have been doing vertical scaling of achievement tests for years. I am sure that the educational measurement field as a whole is a little skeptical about vertical scaling, but it seems a reasonable thing for policy makers to ask how much students have achieved based on an academic year's worth of education. A problem with current vertical scales is that every testing program develops a new vertical scale using different units and setting the origin of the scale at a different place. This seems to me to be roughly equivalent (very roughly) to the early work to define a common temperature scale. There is no correct scale for measuring temperature, but over time, two scales seem to have won out. We now almost universally use either the Fahrenheit or the Celsius scales for reporting temperature. These scales are based on quite different philosophies, but we use them anyway and people do not think about how they came into existence. What I would like to see happen is that someone develops a scale for educational achievement that is like one of the temperature scales. I would like to give a prize to the person who did.

The Fahrenheit scale was based on the coldest temperature that Fahrenheit could produce using ice and various chemicals. That temperature was set at zero on the scale. The equivalent on an achievement scale would be setting zero to the lowest level of performance that we would expect to observe in a K-to-12 educational system. The temperature of 100 on the Fahrenheit scale is close to the body temperature of humans. We might set 100 on the achievement scale at the amount of achievement expected of the typical high school student upon getting a diploma. As with temperature, these are idealized values. The same could be done for achievement. Then the unit of measurement would be related to the amount of achievement expected over the usual span of education.

The Celsius scale was based on the properties of water, because water is such a critical part of life. The point at which water freezes under controlled conditions is given the value 0 and the point that water boils under controlled conditions is given the value of 100. It happens that these two points are much further apart than the points selected by Fahrenheit so the units of the measurement scale are bigger. The parallel in educational measurement would be to determine some critical educational point, such as being literate and call that 0 and another critical point, like being considered well educated and call that 100. Of course, the person who devises the scale gets to decide these two points.

If we could devise this kind of scale, then we could think about the amount of growth in achievement that is likely with one year of instruction. Is it 1 unit or 10 units? The meaning of that number will depend on the size of the units on the scale. Now we do not have much good data on what is the typical amount of gain in achievement over a year because are test scales are so varied in meaning. If this problem can be resolved, I would happily award the prize, if I had the money.

I hope these comments stimulate some thoughts about educational measurement. I also hope to see you all in San Diego.

HELLO FROM THE NEW EDITOR
By, Thanos Patelis, The College Board

Dear Friends and Colleagues: I am honored and thrilled to get this opportunity to work with the NCME officers, directors, meeting chairs, the NCME Newsletter Advisory Board and all of you in continuing the exceptional efforts and work that Scott Bishop and the past members of the NCME Newsletter Advisory Board provided us. So, foremost, I want to offer much appreciation to Scott Bishop and the past members of the Advisory Board. Next, I hope to introduce two things in this and future issues: (1) a standing column by graduate students and (2) a spotlight on the people that make this a great organization. So, in this issue, I am pleased to introduce Carol Barry from James Madison University, who is offering us her first column. Carol will continue to provide her columns in the next three issues highlighting issues for and about graduate students. In the next issue, I hope to unveil some articles that highlight the people of our organization and look to you all for suggestions. Finally, in order to continue the utility of our newsletter, I look to you for submissions that would be of interest to NCME members. So, please drop me an email with suggestions. Sincerely and at your service, your friendly neighborhood newsletter editor, Thanos.

DEADLINE EXTENDED!!

Call for Nominations: Editor, *Educational Measurement: Issues and Practice*
 The three-year term as editor for *Educational Measurement: Issues and Practice* is ending for the current editor Susan Brookhart. The NCME Publications Committee is seeking nominations for the next editor. Nominees should have extensive experience reviewing and publishing research related to educational measurement. Please send all nominations to Stephen Sireci at sireci@acad.umass.edu. All nominations should include the vita of the nominee. Thank you for helping us select the next editor of one of the most important journals in the field of educational measurement. The deadline for nominations has been extended until Sunday, March 22, 2009.

A GRAD STUDENT'S GUIDE TO NCME: MAKING THE MOST OF YOUR EXPERIENCE

Carol Barry, James Madison University

We find ourselves now one month away from the NCME conference. The next few weeks will likely be spent preparing presentations, printing posters, painfully whittling down our talks to fit within that 10- to 12-minute time limit, and finalizing travel plans. To help us begin mapping out our week in San Diego, this issue of the NCME Newsletter is filled with highlights from this year's program as well as suggestions for fun things to do in the area. As graduate students, however, we may still be uncertain about how to get the most out of our NCME experience. What should I attend? How do I meet people? This article is intended as a guide for grad students, both new and old, to answer some of the lingering questions about what to do at NCME.

Navigating the Program

Sometime before leaving for the conference, take thirty minutes to sit down with the program and map out a schedule of things you want to attend (e.g., sessions, training sessions, social events) and where they are located. In doing so, keep in mind that you should attend things that *you* are interested in. It may be tempting to simply pick sessions that others in your program are attending. Realize that it is ok to venture out on your own. If there are authors whose work you admire, try to attend their sessions; you may have the opportunity to speak to them afterwards. Finally, although there are usually five timeslots scheduled for each day, it might be helpful to limit yourself to about three sessions per day. This helps keep you fresh, keeps your mind awake, and gives you some time to take care of yourself.

Network, Network, Network

The opportunity to network is one of the best reasons for going to conferences, and the NCME conference is a great place to meet people. Although it can be intimidating, don't be afraid to introduce yourself to someone whose work you may have read. The wonderful thing about being a graduate student is that many professionals are flattered to know they've had an impact on your development and are happy to speak with you. If you're a bit shy, ask your advisor, mentor, or a faculty member from your program to introduce you. Make connections with graduate students from other schools as well; you never know what might turn into a collaborative relationship. The main thing to realize is that this conference is full of people who will someday be your peers and colleagues. Start getting to know them early. To facilitate these meetings, it's especially helpful to attend every social event you can. The NCME and AERA Division D Joint Welcome Reception (formerly the NCME No-Host Reception), the NCME Breakfast, and the Fitness Walk/Run are great places to start. I also encourage you to step outside your comfort zone at these events. Don't group together with students from your program, but instead use this opportunity to make *new* contacts.

Attend Graduate Student Events

There are two main advantages for attending events that are directed to graduate students. First, these events tend to include some of the most influential professionals in the field and focus on some of the most important current topics. Furthermore, these sessions tend to be geared toward a graduate student audience, which means that we have the opportunity to learn about cutting edge issues from those who know them best. Second, attending these events is another wonderful way to meet graduate students from other programs. Again, these are individuals who are your peers in the field. Get to know them. Learn about your common interests. They may one day become your colleagues or co-authors. At the very least, you will begin to see familiar faces throughout the conference.

Attend Pre-Conference Training Sessions

If you can afford it, it is a great idea to attend a pre-conference session. Although these do cost extra, these sessions tend to be smaller and longer than a regular paper session. This gives you the opportunity to interact with the presenter(s) and really focus on the topic. Additionally, these sessions provide you the opportunity to learn about techniques or topics that you don't get in your coursework.

Behave Professionally

It may be helpful to view conferences as 24-hour job interviews. This isn't meant to scare you, but just as a reminder to behave in a professional manner at all times. Simply put, you never know what potential employer or coworker may be in the room with you. First impressions go a long way, and you want to be sure to make a good one.

Take Some Time for Yourself

Attending the NCME conference can be stressful, between presenting, attending events, and networking. Try to take some time for yourself while in San Diego. Take advantage of the wonderful (and warm!) location. Get outside, take a walk, clear your head, stick your toes in the ocean, or treat yourself to a fun meal. Setting aside some time for yourself each day, even if it is only 15 minutes, will help you reflect on the things that you have learned and the people you have met, stay focused, and get the most out of being at this conference.


NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION
2009 ANNUAL MEETING: PROGRAM HIGHLIGHTS

April 12 – 16 in San Diego, CA
(Locations are at the Hard Rock Hotel unless otherwise indicated)

Much appreciation is extended to our annual meeting chairs, Daniel Bolt, University of Wisconsin-Madison, and Kathleen Gialluca, Pearson, for putting together an exciting program in San Diego. Please make sure you visit www.ncme.org to register, secure hotel rooms, and see the specifics of the program. Here's an overview below:
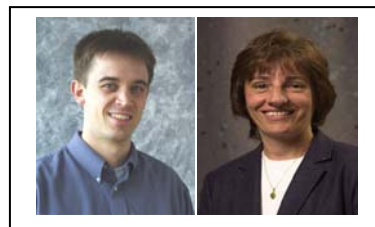
Presidential Address

What I Think I Know
Mark Reckase
Day: Wed.; Time: 8:00 – 10:15 a.m. (Breakfast & Business Meeting); Location: San Diego Convention Center, Ballroom 20D

Career Award Recipient Address

Scores and Scales for Educational Tests
Moderator: David Frisbie
Presenter: Michael Kolen
Discussant: Michael Kane
Day: Wed.; Time: 10:35 a.m. – 12:05 p.m.; Location: Legends 2

Invited Symposia

Issues in the Use of Automated Essay Scoring in High Stakes Assessments
Organizer/Moderator: Brent Bridgeman
Participants: David Williamson, Tim Davey, Brent Bridgeman, Catherine Trapani, Karen Lochbaum, John De Jong, Yigal Attali
Discussants: Mark Shermis, Brian Clauser
Day: Tues., Time: 8:15 – 10:15 a.m.; Location: Legends 3

Measuring and Evaluating Growth in Student Achievement: A Conversation about Technical and Conceptual Issues
Organizer/Moderator: Derek Briggs
Participants: Dale Ballou, Lou Mariano, Daniel McCaffrey, J.R. Lockwood, Damian Betebenner, Derek Briggs, Mark Wilson
Discussants: Michael Kolen, Richard Patz, Frank Rijmen
Day: Tues.; Time: 8:15 – 10:15 a.m.; Location: Legends 2

Making Test Score Scales and Reports More Understandable and Useful
Organizer/Moderator: Ronald Hambleton
Participants: John Hattie, Sandip Sinharay, Shelby Haberman, Gautam Puhan, Thanos Patelis, Haifa Matos, Krista Breithaupt, David Chuah, April Zenisky, Stephen Sireci
Discussant: Joseph M. Ryan
Day: Tues.; Time: 4:05 – 6:05 p.m., Location: Legends 2

New Directions in Test Security and Cheating Detection Research
Organizer/Moderator: Alan D. Mead
Participants: David Chuah, Ben-Roy Do, Bradley Brummel, Fritz Drasgow, John Mattar, Aster Tessema, Dennis Maynes, Alan Mead, Gunnar Schrah, Leanne Buehler, Bobby Baker
Discussants: Gerald Melican, Anthony Zara
Day: Tues.; Time: 4:05 – 6:05 p.m.; Location: Legends 6

Revising our Test Standards
Organizer: Rosemary Reshetar
Participants: Linda Cook, Barbara Plake, Brian Gong, Denny Way, Lauress Wise
(Co-sponsored with AERA Division D, scheduled as an AERA session)
Day: Wed.; Time: 2:15 – 3:35 p.m.; Location: Omni San Diego, Salon D

Standard Setting in an Accountability Growth Context: A Process or One-Time Event?
Organizer/Moderator: Isaac Bejar
Participants:Michael Kane, Damian Betebenner, Steve Ferrara, Dubravka Svetina, Anne Davidson, Jim Pellegrino, David Abrams
Discussants: Robert Linn, Ronald Hambleton
Day: Wed.; Time: 4:05 – 6:05 p.m.; Location: Legends 3

Alternate Assessment based on Alternate Achievement Standards: Improving Technical Rigor
Organizer: Claudia Flowers
Moderator: Martha Thurlow
Participants: Diane Browder, Scott Marion, Jim Pellegrino, Linda Cook, Marianne Perie, Stanley Rabinowitz
Discussants: Michael Kolen, Suzanne Lane
Day: Wed.; Time: 4:05 – 6:05 p.m.; Location: Legends 5

Bradley Hanson: The Man Behind the Award and His Legacy as a Psychometrician
Organizer/Moderator: Jimmy de la Torre
Participants: Deborah Harris, Gary Skaggs, Won-Chan Lee, Jianbin Fu, Xiaohong Gao, Anton Beguin
Discussant: Richard Patz
Day: Thurs.; Time: 10:35 a.m. – 12:05 p.m.; Location: Legends 2

Committee-Sponsored Symposia

Accurate Assessment of Student Achievement: Today's Challenges and Solutions
Organizer: Dubravka Svetina
Moderator: Kimberly A. Swygert
Participants: Jamal Abedi, Chad Buckendahl, Christy Hovanetz-Lassila, Robert Lissitz, John Tanner
Sponsored by the Graduate Student Issues Committee
Day: Tues.; Time: 4:05 – 6:05 p.m.; Location: Legends 3

Large Scale Assessment and Accommodating Students with Disabilities: Past, Present & Future
Organizer/Moderator: Sara Bolt
Participants: Martha Thurlow, Barbara Plake, Cara Cahalan Laitusis, Sami Kitmitto, Victor Bandeira de Mello, Jerry Tindal
Sponsored by the Diversity Issues in Testing Committee
Day: Wed., Time: 10:35 a.m. – 12:05 p.m.; Location: Legends 3

NCLB at Year 8 in the Assessment of English Language Learners:Taking Stock of the Assessment and Accountability Systems
Organizer/Moderator: Phil Morse
Participants: Jamal Abedi, David Francis, Rebecca Kopriva
Discussants: Gregory Cizek, Robert Linquanti
Sponsored by the National Association of Test Directors
Day: Wed.; Time: 4:05 – 6:05 p.m.; Location: Legends 2

Graduate Student Poster Session
Sponsored by the Graduate Student Issues Committee

This 12th annual poster session of NCME's Graduate Student Issues Committee provides an opportunity for graduate students to share their work and receive feedback from professionals and their peers.
Day: Wed., Time: 4:05 – 6:05 p.m.; Location: The Edge

NCME and AERA Division D Joint Welcome Reception for Current and New Members

Members of NCME and AERA Division D are invited to attend the NCME and AERA Division D Joint Welcome Reception for Current and New Members. Please note that this reception is a new event for the 2009 meeting.
Day: Wed., Time: 6:15 – 7:30 p.m.; Location: San Diego Convention Center, Ballroom 20D

NCME Fitness Run/Walk
Organizers: Brian F. French and Jill van den Heuvel

Run a 5K or walk a 2.5K course along the waterfront. Pre-registration is required. Check the NCME website for details.
Day: Thur.; Time: 5:45 a.m. - 7:30 a.m. Location: Meet in the lobby of the Hard Rock Hotel at 5:45 a.m.

Pre-Conference Training Sessions – April 12 – 13, 2009

- The 2009 NCME pre-conference training sessions will be held at the Hard Rock Hotel in San Diego, California on Sunday, April 12 and Monday, April 13, 2009.

- Advance registration for the training sessions is strongly encouraged. The only way to register in advance for the training sessions is to use NCME's on-line registration system. To do this, please go to http://www.ncme.org.

- Registration on-site will be available only for those training sessions that have not been filled through advance registration.

- Refunds of registration fees for the training sessions cannot be made after March 2, 2009.

- Please note that Internet connectivity will not be available at the conference and that, where applicable, participants should download the software required prior to the training sessions.

Sunday, April 12, 2009

AA: Using R for Everyday Research
Presenter(s): Brian Habing, University of South Carolina; Jessalyn Smith, University of South Carolina
Fee: $40; Time: 8:00 a.m. - 12:00 Noon; Location: Legends 2

The free statistics package R has become a favorite of statisticians over the past decade – and it offers a large number of benefits to quantitative researchers in all areas of educational research. With you working along through each step on your own laptop computer, this training course will cover some of the most useful aspects of R for any researcher, including: making fully customized graphs (including color, axes, and labels); manipulating data sets in an intuitive way to quickly get the precise subset of subjects and variables that you want; and performing statistical analyses with a single command. The course will end with basic examples of how R can be used to simulate data sets (with an example perfect for classroom use) and how it can be easily customized to perform functions that aren't built in.

This course is designed for those who have had a two-course sequence in quantitative methods but have no previous experience with R. Participants must bring their own (windows compatible) laptop computer; all required software will be provided.

BB: Quality Control in Test Development, Scoring, and Reporting of Test Scores
Presenter(s): Avi Allalouf, National Institute for Testing and Evaluation; Ruth Fortus, National Institute for Testing and Evaluation
Fee: $65; Time: 8:00 a.m. - 12:00 Noon; Location: Legends 3

Testing in educational and psychological measurement involves a number of important stages, each depending greatly upon the previous one: test development, test scoring, test analysis and score reporting. This training session deals with quality control procedures for these stages.

Quality control procedures are required in order to monitor the testing process and to keep the number of mistakes to a minimum. Mistakes in scoring, for example, can lead to legal action against the testing agency or the educational institution; a high incidence of mistakes in items will have an adverse impact on test reliability and validity.

Professional practitioners should be aware of possible mistakes that can occur during test development, test scoring, test analysis and the reporting of scores. They should act in accordance with up-to-date standards and have a broad knowledge of

quality control practices, as these are critical in the never-ending fight against errors. This session is intended to increase accuracy in test measurement.

In the session, mistakes that might occur at each stage will be presented, followed by examples and quality control procedures for avoiding, detecting or correcting these mistakes.. Many of the quality control procedures discussed are also relevant for internet-delivered and internet-scored testing.

The session will also touch on models that deal with the causes, prediction and reduction of human error.

The workshop will be potentially useful for people who are involved in:
- test development
- test administration
- scoring tests
- item and test analysis (including test norming and equating)
- maintaining test security
- reporting test results and providing feedback to people who have been tested
- policy-making and legislation

The workshop will consist of short modules, each accompanied by real-world examples. Participants will be given hands-on practice in detecting various types of errors. The workshop content is based upon experience gained by the presenters from their work at NITE, and upon an ongoing project of developing quality control guidelines for the ITC (International Testing Commission).

CC: Linking and Aligning Scores and Scales
Presenter(s): Neil Dorans, Educational Testing Service; Jinghua Liu, Educational Testing Service; Mary Pommerich, Defense Manpwer Data Center; Michael Walker, Educational Testing Service
Fee: $110; Time: 8:00 a.m. - 12:00 Noon; Location: Legends 1

The communication of linking issues to test score users is a critical component to ensuring the validity of a linkage. This training session seeks to facilitate communication about the appropriate use and interpretation of linked scores by emphasizing the different meanings that can be attached to different linkages, and the necessary requirements to achieve solid linkages. A foundations portion will present a historical perspective on score linking, provide definitions and distinctions between types of linkages, discuss relevant data collection designs, and give an overview of linking methodology and assumptions. A linking scenarios portion will make expanded distinctions between types of linkages and discuss practical issues, using real world examples. Topics will be equating, tests in transition, concordance, vertical scaling, and linking group assessments to individual assessments. A tools portion will discuss indices that can be used to choose an appropriate linkage type and methods that can be used to evaluate linkage quality. A score interpretation portion will focus on the appropriate usage and interpretation of linked scores, comparing and contrasting across the different linking scenarios.

DD: Developing Noncognitive Assessments
Presenter(s): Patrick Kyllonen, Educational Testing Service; Richard Roberts, Educational Testing Service
Fee: $80; Time: 8:00 a.m. - 5:00 p.m.; Location: Legends 4

Noncognitive qualities are increasingly recognized as important determinants and reflections of success in education from K-12 through graduate and professional school. In this training session we will provide background theory and frameworks for developing noncognitive assessments, and provide hands-on experience in developing and evaluating noncognitive assessments. We will review the major personality models and related noncognitive constructs, discuss methods used to measure noncognitive qualities, demonstrate how to find or to write noncognitive items, present the advantages and disadvantages of different approaches to collecting data, and review strategies for dealing with various validity threats, such as the problem of faking on self assessments. We will demonstrate analysis approaches, including exploratory and confirmatory factor-analysis, and review various uses of noncognitive assessments.

The session will consist of a series of lectures interspersed with examples and empirical findings. Q&A will be encouraged throughout. We will cover the following topics:

- Noncognitive construct frameworks, models, and theories (personality, attitudes, values, beliefs, and other constructs)
- Developing assessments from construct definitions and item pools, including the international personality item pool (IPIP)
- Various methods for assessing noncognitive qualities (self-assessments, others' ratings, situational judgment tests, conditional reasoning, implicit association tests)
- Item writing do's and don'ts

7

- The problem of faking on self-assessments (preventing, detecting, & correcting for it)
- Delivery platforms (web and paper-and-pencil)
- Exploratory factor analysis and other data structure exploration methods
- Confirmatory factor analysis
- Advanced methods (IRT, latent class models, unfolding models)
- Special topics (rating scale issues [optimal number of points; presence of neutral point, "do not know"], reverse key items)
- Indirect measures (e.g., from school records)
- Example noncognitive assessments (self-help for community college; institutional reporting for K-12; high stakes for graduate school)

EE: Generalizability Theory and Applications
Presenter(s): Robert Brennan, University of Iowa; Xiaohong Gao, ACT, Inc.; Won-Chan Lee, University of Iowa
Fee: $135; Time: 8:00 a.m. - 5:00 p.m.; Location: Legends 5

Generalizability theory liberalizes and extends classical test theory. In particular, generalizability theory enables an investigator to disentangle multiple sources of error through the application of analysis of variance procedures to assess the dependability of measurements.

The primary goals of this training session are to enable participants to understand the basic principles of generalizability theory, to conduct relatively straightforward generalizability analyses, and to interpret and use the results of such analyses. Mathematical and statistical foundations will be treated only minimally. Major emphasis will be placed upon quickly enabling participants to conduct and interpret relatively straightforward generalizability analyses, then more complicated ones. Examples will include various types of performance assessments.

Prerequisites include knowledge equivalent to one course in educational measurement and familiarity with ANOVA at the level treated in introductory graduate courses in education and psychology. A book written by the director and entitled Generalizability Theory will be distributed to participants and used as a principle reference in the training session. Computer programs for performing generalizability analyses will be discussed and illustrated. (Participants need not bring laptops.)

FF: Item Response Theory: Parameter Estimation Techniques
Presenter(s): Seock-Ho Kim, Univeristy of Georgia
Fee: $135; Time: 8:00 a.m. - 5:00 p.m.; Location: Legends 6

Theory and methods for the educational and psychological measurement of latent variables using item response theory methodology are discussed. The one-parameter logistic or Rasch, the two-parameter logistic, and the Birnbaum's three-parameter models for dichotomously scored item response data will be reviewed from a theoretical viewpoint with an emphasis on the various estimation techniques of the model parameters. Applications of these models to practical measurement situations will be studied using item response theory computer programs. Topics of the course consist of item calibration, scoring, information, and some applications to instrument construction. Models for polytomously scored items are briefly discussed.

Prerequisites include knowledge equivalent to one graduate course in theoretical educational measurement and familiarity with differential and integral calculus treated in undergraduate mathematics courses. A book coauthored by the director with Frank B. Baker entitled Item Response Theory: Parameter Estimation Techniques will be distributed to participants and used as a principle reference in the training session. Computer programs for performing item response theory analyses will be discussed and illustrated. Participants are encouraged to bring their own laptop computers.

The intended audience is principally upper-level graduate students and new measurement professionals who are interested in learning about the various parameter estimation techniques in the context of unidimensional item response theory models.

RR: Managing Simulation Studies with R
Presenter(s): Brian Habing, University of South Carolina; Jessalyn Smith, University of South Carolina
Fee: $40; Time: 1:00 p.m. - 5:00 p.m.; Location: Legends 2

Simulation studies to validate various procedures' effectiveness are a major part of quantitative and psychometric research. The R statistical package can be used to easily run and manage simulation studies, including those that need to call pre-existing programs such as BILOG, MPlus, NOHARM, PARSCALE, POLYEQUATE, and TESTFACT. This course will guide the participants through using R to easily generate and manipulate a wide variety of data sets, create the command and data files required by other software, run the other software, and read in the output for further analysis.

This course assumes that the participants have at least some familiarity with R – programming experience is not assumed. Participants should bring their own (windows compatible) laptop computer and any executables that they need to integrate into their own simulation studies. Copies of R, NOHARM, and POLYEQUATE will be provided.

GG: A nonlinear mixed models approach to IRT
Presenter(s): Paul De Boeck, KU Leuven, Frank Rijmen, Educational Testing Service; Francis Tuerlinckx, KU Leuven; Mark Wilson, UC Berkeley
Fee: $65; Time: 1:00 p.m. - 5:00 p.m.; Location: Legends 3

The central message of the introduction is that it is beneficial to see IRT models as extensions of generalized linear regression models that seek to model facets of the measurement situation: These facets are most typically persons and items, but the set may be extended to incorporate other facets such as raters, and may also be re-labelled to suit particular applications. While the link function and the random component of the regression model remain the same, the most interesting part of the extension concerns the structural part of the model: (1) the kind of predictive function (linear or nonlinear, e.g. bilinear), (2) the effects (weights) of the predictors (fixed effects or random effects).

Starting from some well-known IRT models, other and less well-known models will be framed in this approach, based on a volume published by Springer: "Explanatory Item Response Models: A generalized linear and nonlinear approach" (De Boeck & Wilson, 2004). We will illustrate how the models can be estimated with the SAS procedure NLMIXED.

The workshop will consist of two parts. In the first part, the explanatory item response framework will be presented, and it will be explained how the framework fits within the family of generalized linear and nonlinear mixed models. Specific attention will be devoted to the distinction between descriptive and explanatory item response models, and the distinction between fixed and random effects. It will be shown how well-known item response models fit within this framework. In addition, the framework naturally leads to new item response models, such as models with both random item and random person effects.

In the second part, an in-depth account will be given of multidimensional item response models, and models for polytomous data. Again, both families of models can be conceptualized as generalized linear and nonlinear mixed models, and doing so naturally leads to model extensions that may be of interest to the applied researcher. In this part, some attention will be devoted to model estimation as well. We will also emphasize random item concepts and models

Throughout, the models are illustrated with datasets on anger and verbal aggression.

Monday, April 13, 2009

HH: Skills Diagnosis with Latent Variable Models
Presenter(s): Jimmy de la Torre, Rutgers University; Robert Henson, University of North Carolina at Greensboro; Jonathan Templin, University of Georgia
Fee: $65; Time: 8:00 a.m. - 12:00 Noon; Location: Legends 1

The primary aim of skills diagnosis is to develop and analyze tests in ways that reveal information with more diagnostic value, when compared with traditional approaches. In the methods for skills diagnosis that we consider mastery of particular skills or states of knowledge can be represented by a list of binary latent variables, indicating mastery of each of a finite set of skills under diagnosis. The main objective of skills diagnosis is to classify examinees according to this list of skills. In this training session, several popular modeling and classification approaches will be discussed. Three conjunctive latent class models known as the DINA, NIDA, and Fusion models will be introduced, and software for fitting these models with Mplus will be demonstrated. The training session is meant to provide practical guidelines for implementing skills diagnosis, and considers essential topics such as construction of fixed-length tests, identifying the attributes measured by items, and model-data fit.

The intended audience for this training session includes anyone interested in cognitive or skills diagnosis who has some familiarity with item response theory or classical test theory. No previous knowledge of latent class models or cognitive diagnosis is required. The material will be useful for faculty and students specializing in educational testing, as well as testing professionals working in government or private testing organizations.

The objective of this training session is to provide a short course in some of the most common methods of latent variable modeling that are being applied in cognitive and skills diagnosis. The emphasis is on education as well as training with a particular piece of software. By the end of this session, participants should have a basic understanding of general latent class models, conjunctive latent class models tailored to cognitive diagnosis, methods for constructing exams, and evaluation of goodness of fit. There will also be a discussion of identifying skills on an exam, and construction of exams when diagnosis is the primary objective.

II: Vertical Scaling Methodologies, Applications, and Research
Presenter(s): Michael Kolen, University of Iowa; Ye Tong, Pearson
Fee: $65; Time: 8:00 a.m. - 12:00 Noon; Location: Legends 2

The potential need for constructing a vertical scale arises whenever a testing program has multiple grade levels and wishes to have a common scale to compare test scores across these grade levels. Vertical scaling uses statistical process to place test scores that measure similar content domain but at different educational levels onto a common scale. The goals of the session are for attendees to be able to understand the principles of vertical scaling, to conduct vertical scaling and to interpret the results of vertical scaling in reasonable ways. Vertical scaling will be contrasted with related equating and linking processes. Traditional and IRT vertical linking methodologies will be described and practical issues will be discussed.

The focus is on developing a conceptual understanding of vertical scaling through numerical examples and discussion of practical issues. Importance and challenges related to vertical scaling will be included. The text for the session is a chapter in the second edition of Kolen and Brennan's (2004) Test Equating, Scaling, and Linking. Methods and Practices (Second Edition). The session is designed for upper level graduate students, new Ph.D.'s, testing professionals with operational or oversight responsibility for vertical scaling, and others with interest in learning about vertical scaling methods and practices. Participants should have at least two graduate course in measurement and two graduate courses in statistics.

JJ: Development and Use of Innovative item Types in Computer-based Testing
Presenter(s): Kathleen Scalise, University of Oregon; Mark Wilson, University of California, Berkeley
Fee: $65; Time: 8:00 a.m. - 12:00 Noon; Location: Legends 3

One potential limitation for realizing the benefits of computer-based assessment (CBT) in both instructional assessment and large scale testing comes in designing questions and tasks with which computers can effectively interface (i.e., for scoring and score reporting purposes) while still gathering meaningful measurement evidence. This workshop will allow participants to explore introducing some innovative item types into their assessment content. A taxonomy of 28 innovative item types in computer-based assessment will be introduced. These item types have responses that fall somewhere between fully constrained responses (i.e., the conventional multiple-choice question), which can be too limiting to tap much of the potential of new information technologies, and fully constructed responses (i.e. the traditional essay), which can be a challenge for computers to meaningfully analyze. Participants will bring example items to the workshop or be provided with examples, work hands-on to convert to innovative types through a variety of content approaches, investigate and implement automated scoring options for their selected types, and finish the workshop with modeling practices for collection of high quality assessment evidence, in a CBT interface using IRT.

KK: An Introduction to Student Growth Percentiles: Concepts, Estimation and Use
Presenter(s): Damian Betebenner, Center for Assessment; Jinnie Choi, University of California, Berkeley; Hi Shin Shim, Georgia Tech; Dianne Lefly, Colorado Department of Education; Marie Huchton, University of Colorado, Boulder
Fee: $80; Time: 8:00 a.m. - 5:00 p.m.; Location: Legends 4

The proliferation of annual student testing during the last decade has left states and testing organizations with vast amounts of longitudinal assessment data and few sophisticated means to analyze these multiyear data sets. As a consequence, use of growth analyses to inform discussions about student growth and its relationship to education quality has been limited. In this training session, participants will be introduced to student growth percentiles and shown how to use the open source R software package to calculate student growth percentiles and percentile growth trajectories with large (e.g., state-level) longitudinal datasets. Topics covered will include a conceptual overview of student growth percentiles, data preparation, student growth percentile calculation, percentile growth trajectory calculation and their use with growth standard setting. The session will incorporate real-world examples of how the results of such analyses can be used as part of state and federal accountability systems to inform discussions about educational quality.

LL: Applying Hierarchical Models to Causal Inference
Presenter(s): Guanglei Hong, OISE / University of Toronto; Stephen Raudenbush, University of Chicago
Fee: $80; Time: 8:00 a.m. - 5:00 p.m.; Location: Legends 5

In this training session we will introduce recent development of causal inference concepts and methods for evaluating educational policy and program effects in multi-level settings when randomized experiments are infeasible. We teach hierarchical linear and nonlinear models in combination with propensity score-based methods for causal effect estimation. Education examples will be used throughout in lecture, discussion, and hands-on practice. The session is intended for researchers interested in investigating the effectiveness of educational policies, intervention programs, and various educational practices. After presenting the basics of hierarchical models and of causal inference, we use examples to illustrate (1) how to conceptualize, in terms of potential outcomes, the causal effects of educational interventions carried out in a multi-level school system, (2) how to identify and summarize information of selection bias from multiple sources through analyzing logistic

regression models or hierarchical generalized linear models, (3) how to stratify sample data on the basis of the estimated propensity score, (4) how to use hierarchical models to statistically adjust for the selection bias in multi-level data, (5) how to make explicit statistical assumptions, and (6) how to assess the consequences of possible unmeasured confounders. Participants will practice the procedure of causal effect estimation using HLM version 6 along with SPSS 15.0. Participants are expected to bring a laptop computer with SPSS and HLM standard version or trial edition installed. The standard version or the free 15-day trial edition of the HLM 6 software available at http://www.ssicentral.com/hlm/downloads.html.

MM: Bayesian Networks in Educational Assessment
Presenter(s): Duanli Yan, Educational Testing Service; Russell Almond, Educational Testing Service; Robert Mislevy, University of Maryland; David Williamson, Educational Testing Service
Fee: $80; Time: 8:00 a.m. - 5:00 p.m.; Location: Legends 6

The Bayesian paradigm provides a convenient mathematical system for reasoning about evidence. Bayesian networks provide a graphical language for describing complex systems, and reasoning about evidence in complex models. This allows assessment designers to build scoring that have fidelity to cognitive theories about the domain and yet are mathematically tractable and can be refined with observational data. Topics covered in this tutorial are evidence-centered assessment design, basic Bayesian network representations and computations, available software for manipulating Bayesian networks, refining Bayesian networks using data, and example systems using Bayesian networks. It is recommended that participants bring a laptop to run sample exercises using the student version of Netica (http://www.norsys.com/).

NN: Building and Documenting a Valid Assessment System for Students with Disabilities
Presenter(s): Karen Barton and Lara Osleson, CTB/McGraw-Hill
Invited Speaker: Dianne Lefly, Colorado Department of Education
Fee: $65; Time: 1:00 p.m. - 5:00 p.m.; Location: Celebrate

This course is intended for psychometricians, researchers, state Departments of Education personnel, and test development experts who wish to design, build, and document in technical format reliable, valid, defensible assessments, particularly alternate and modified assessments for students with disabilities. Topics range from assessment policy, design, and development to appropriate statistical design and analyses, special studies, and technical documentation. The session will provide the audience with sound psychometric tools and practices to assure alternate (as well as modified and general) assessments can meet high standards of technical adequacy with practical tips and solutions for documenting evidence in a legally defensible manner. In particular, this session will focus on building validity evidence.

Participants will be guided through each step in designing and building a valid and defensible assessment, with approaches to collecting appropriate validity evidence linked to the Standards (AERA, NCME, APA) and Critical Elements). Parallels and distinctions will be made between alternate assessments and both modified and general assessments. Invited speakers will discuss modified and alternate approaches from a state perspective.

OO: Cognitive Assessment: An Introduction to the Rule Space and Q-Matrix Method
Presenter(s): Kikumi Tatsuoka, Columbia University; Anabelle Guerrero, University of Costa Rica; Enis Dogan, American Institutes for Research
Fee: $110; Time: 1:00 p.m. - 5:00 p.m.; Location: Legends 1

This book introduces a new methodology that allows for the analysis of test results that is free from ambiguous interpretations and demonstrates an individual's true state of knowledge. Measuring the underlying knowledge and cognitive skills is not an easy task because it is impossible to directly observe them; therefore, they are named "latent variables". However, the latent variables useful in cognitive diagnosis must be in the 100's and not just one variable like a "q" ability variable in Item Response Theory. To achieve these difficult goals, we need a new methodology that will transform many unobservable knowledge and skills variables (defined as "attributes" throughout in the book) into observable and measurable attributes without losing their original meanings.

The purpose of this book is to introduce one such methodology, Rule Space, that has been used since the 1980s and has made it possible to measure these unobservable latent variables and to clearly interpret the results, without losing the original meaning of attributes. The Rule Space Method (RSM) transforms unobservable attributes involved in test items into observable attribute mastery probabilities that are defined as the probability of using each attribute correctly to get the correct answer for given problems. In other words, RSM converts students' item response patterns into the attribute mastery probabilities. The Rule Space Method (RSM), which can determine an individual's strengths and weakness, has been applied to PSAT/NMSQT[©] to generate scoring reports, which inform schools, teachers and parents exactly what the total score of 500 means. Since RSM belongs to an approach of statistical pattern recognition and classification problems popular in engineering areas, this book will be useful to graduate students in a variety of disciplines. This book has ten chapters but in this training session, Emphases are

given to the Q-matrix Theory, Rule Space classifications, the attribute reliability and validity theory. Inquiries about this session should be sent to kumitats@yahoo.com.

PP: Technical Aspects of School Accountability
Presenter(s): Huynh Huynh, University of South Carolina; Robert Kennedy, University of Arkansas for Medical Sciences; Charity Smith, Arkansas Department of Education
Fee: $65; Time: 1:00 p.m. - 5:00 p.m.; Location: Legends 2

The purpose of this training session is to introduce recent technical development regarding school accountability. Technical issues concern with creating school index based on test data, assessing reliability and conditional standard error for the index, setting via school-descriptor and bookmark processes, and assessing reliability and validity of school classifications. Using the Arkansas Act 35 school accountability system as a case study, participants will be guided through the development and operation of the index for school performance (status) and the index of school growth or improvement gain. Handouts given include two technical documents, one for school performance and the other for school growth. Participants are expected to be familiar with basic knowledge of applied statistics and technical aspects of assessment, and a level of awareness of operational and legal issues relating to school accountability.

QQ: Tips for graduate students: Advice for finishing school, obtaining a job, and starting a career
Presenter(s): Deborah Harris, ACT, Inc.; Julio Sanclemente, CTB/McGraw-Hill; Andrew Ho, University of Iowa
Fee: $15; Time: 1:00 p.m. - 5:00 p.m.; Location: Legends 3

This training session has three main components: (1) Finishing up the Ph.D., including finding a dissertation topic, what employers are looking for, how to maximize job possibilities at the conference, and how to maximize experiences while still a student (classes, internships, work experiences, networking, professional associations), (2) Obtaining a job, including how to decide what kind of job you want, how to locate where jobs are available (universities, testing companies, school districts, state departments, professional/licensing organizations, etc.), how to apply for jobs (including targeting cover letters, references, and resumes) and the interview process (questions to ask and questions to be prepared to answer, giving a job talk, following up after the interview), and (3) Beginning a career, including negotiating an offer, job politics, adjusting to the environment, career path, publishing, professional service, being a mentor/finding a mentor, balancing work and life, and what to do if you end up hating your job. Attendees are provided with a rich packet of resource materials covering the above areas, and are actively encouraged to ask questions and provide comments throughout the session. The presenters provide subsequent information after the session via email to attendees on issues/questions that require follow up (in the past this has included items like job possibilities in a particular geographic region or more information on specific grant possibilities for grad students).

SOME RESTAURANTS & THINGS TO DO IN SAN DIEGO
College Board Staff and Friends

Restaurants & Other

Casa Guadalajara
Address: 4105 Taylor St. (in Old Town), San Diego 92110
Cuisine: Mexican
Distance from Hard Rock Hotel: approx. 5 miles (San Diego Trolley or drive/taxi)
Price: Moderate
Description: Located in Old Town's Bazaar del Mundo, this is a great place for both atmosphere and food.

Chive
Address: 558 Fourth Avenue, San Diego 92101
Cuisine: Modern Asian Fusion
Distance from Hard Rock Hotel: approx. 1/2 mile (walk)
Price: Moderate to Expensive with prixe fixe dinners available
Description: Cool place, award-winning cuisine, close to the Hard Rock! http://www.chiverestaurant.com/chive.html

Crown Room at the Hotel del Coronado (for Sunday brunch)
Address: 1500 Orange Ave, Coronado 92118
Cuisine: American
Distance from Hard Rock Hotel: approx. 5 miles (drive/taxi)
Price: Very expensive
Description: If you're going to be in town on a Sunday, the Sunday brunch at the Hotel Del is an institution.

Dussini's Grill
Address: 275 Fifth Avenue, San Diego, CA 92101
Cuisine: Mediterranean Bar Food
Distance from Hard Rock Hotel: 300 feet (walk)
Price: Expensive
Description: Lounge and pool hall with a lot of style. http://www.dussini.com

George's at the Cove
Address: 1250 Prospect St., La Jolla 92037
Cuisine: California Cuisine
Distance from Hard Rock Hotel: approx. 10 miles (drive/taxi)
Price: Expensive
Description: Located at La Jolla Cove, which is worth a visit. The restaurant has wonderful views and food.

Humphrey's
Address: 2241 Shelter Island Dr., San Diego 92106
Cuisine: Seafood
Distance from Hard Rock Hotel: approx. 6 miles (drive/taxi)
Price: Expensive
Description: Views of San Diego Bay and excellent food.

Karl Strauss Brewing Company
Address: 1157 Columbia Street, San Diego 92101
Cuisine: American / brewpub
Distance from Hard Rock Hotel: approx. 1 mile (walk or drive/taxi)
Price: Moderate
Description: A fun microbrewery serving excellent beer and good food. Casual atmosphere.

Lou and Mickey's
Address: 224 Fifth Avenue, San Diego 92101
Cuisine: Classic American, Steak House & Seafood
Distance from Hard Rock Hotel: 240 feet (walk)
Price: Moderate to Expensive
Description: American dishes served such as their famous prime steak, fresh seafood, chicken, and pasta dishes as well as delicious appetizers to nibble on while enjoying a drink in the cocktail lounge. http://www.louandmickeys.com

Old Town Mexican Café
Address: 2489 San Diego Avenue, San Diego 92110
Cuisine: Mexican
Distance from Hard Rock Hotel: approx. 4 miles (drive/taxi)
Price: Inexpensive to Moderate
Description: One of San Diego's favorite Mexican restaurant, located in the heart of historic Old Town, the birthplace of San Diego. Every dish at Old Town Mexican Café is prepared with fresh ingredients from authentic recipes. Try a colorful carnitas plate, a tasty tostada or enchilada, or a succulent chicken straight off the rotisserie. A children's menu and "Heart Smart" selections are available along with tempting desserts, exotic libations, and a delicious Border Breakfast menu. Note: Taco Tuesdays with many specials 4pm-Close. http://www.oldtownmexcafe.com/

Panda Inn
Address: 506 Horton Plaza (in Horton Plaza), San Diego 92101
Cuisine: Chinese
Distance from Hard Rock Hotel: approx. 1 mile (walk or drive/taxi)
Price: Moderate
Description: It's in a shopping mall, but this is one of the best Chinese restaurants in San Diego.

Rainwater's
Address: 1202 Kettner Boulevard, San Diego 92101
Cuisine: American
Distance from Hard Rock Hotel: approx. 1.1 miles (walk or drive/taxi)
Price: Expensive
Description: Award-winning steak house and seafood. Short distance from the Hard Rock Hotel. http://www.rainwaters.com

Rubio's
Address: 330 W Broadway (near Horton Plaza), San Diego 92101
Cuisine: Mexican fast food
Distance from Hard Rock Hotel: approx. 1 mile (walk or drive/taxi)
Price: Inexpensive
Description: A chain restaurant known for its fish tacos. Fast food, but very "San Diego" – the first restaurant was opened in Mission Beach.

Activity Suggestions

Horton Plaza
Address: 324 Horton Plaza, San Diego 92101
Distance from Hard Rock Hotel: approx. 1 mile (walk or drive/taxi)
Description: It's a shopping mall, but the layout is like being inside an M.C. Escher painting. Even if you don't like to shop, it's worth going just to experience it. If you do like to shop, you'll find plenty there to explore.

Hotel del Coronado
Address: 1500 Orange Ave, Coronado 92118
Distance from Hard Rock Hotel: approx. 5 miles (drive/taxi)
Description: This hotel – known locally as "the Hotel Del" or simply "the Del" – is a San Diego landmark that's worth a visit for its Victorian-era architecture, its history (it was the location for much of the movie "Some Like It Hot") and its wonderful public-access beach. Located on Coronado Island.

La Jolla Cove
Address: 1100 Coast Blvd., La Jolla
Distance from Hard Rock Hotel: approx. 10 miles (drive/taxi)
Description: The entire downtown La Jolla area is a fun place to walk around, shop and explore, but the Cove is its heart.   Not the best beach, but stunningly beautiful, and you might see harbor seals.

Museum of Contemporary Art
Address: 700 Prospect Street La Jolla 92037
Distance from Hard Rock Hotel: approx. 10 miles (drive/taxi)
Description: Modern art museum located in downtown La Jolla.

Museum of Man
Address: 1350 El Prado (in Balboa Park), San Diego 92101
Distance from Hard Rock Hotel: approx. 5 miles (drive/taxi)
Description: Anthropology museum located in Balboa Park.

Old Town
Address: 4133 Taylor St. (Bazaar del Mundo), San Diego 92110
Distance from Hard Rock Hotel: approx. 5 miles (San Diego Trolley or drive/taxi)
Description: A fun area just to wander around, Old Town has historic buildings and the Bazaar del Mundo, a collection of shops and restaurants.

Pacific Beach / Mission Beach / Ocean Beach
Address: 3146 Mission Blvd (Belmont Park), San Diego 92109
Distance from Hard Rock Hotel: approx. 10 miles (drive/taxi)
Description: The entire Pacific Beach / Mission Beach / Ocean Beach area is fun to explore, but if you need a landmark to head to, make it Belmont Park at Mission Beach and go for a ride on the Giant Dipper roller coaster (built in 1925) before you start your beach explorations.

Reuben H. Fleet Science Center
Address: 1875 El Prado (in Balboa Park), San Diego 92101
Distance from Hard Rock Hotel: approx. 5 miles (drive/taxi)
Description: Science museum located in Balboa Park, with great interactive exhibits and an IMAX theater.

San Diego Wild Animal Park
Address: 15500 San Pasqual Valley Rd., Escondido 92027
Distance from Hard Rock Hotel: approx. 30 miles (drive)

Description: Run by the Zoological Society of San Diego (which also runs the San Diego Zoo), the Wild Animal Park allows for a more "safari"-type interaction with the animals and is a worthwhile side trip if you have access to a car and have already been to the zoo.

San Diego Zoo
Address: 2920 Zoo Dr. (in Balboa Park), San Diego
Distance from Hard Rock Hotel: approx. 5 miles (drive/taxi)
Description: If you only have time for one tourist activity in San Diego, this is the one to do. This is a world-famous zoo that deserves its reputation. If you can, take a little time to wander around the rest of Balboa Park while you're there.

News from Our Colleagues about Other Conferences

Division 5 (Evaluation, Measurement, & Statistics) Sessions at the 2009 APA Convention, Toronto

Division 5 has an outstanding program this year in Toronto and we hope that you will plan to attend. With the exception of the Executive Committee Business Meeting the night before the conference, the programming of the sessions sponsored or co-sponsored by Division 5 begins on Thursday August 6, 2009 at 8:00 a.m. and continues through Sunday, August 9 at 2:00 p.m. Preliminary information about the program is available on the Division 5 website: http://www.apa.org/divisions/div5/homepage.html. Workshops and symposia are highlighted below:

- Introduction to Bayesian Analysis (Zhiyong Zhang)
- Meta-Analysis: The State of the Art and Opportunities in Psychological Research (Noel Card)
- Statistical Mediation and Moderation Analysis: How and for Whom Do Interventions Work? (L. Aiken & S. West)
- Factorial Invariance: Why It's Important and How to Test for It (Todd Little)
- Why We Need to Understand Changes in Each Individual Before We Can Understand Changes Among Individuals (Peter Molenaar)
- Using Receiver Operating Characteristic Analysis to Explore the Nature of Recognition Memory (John Wixted)
- A Comprehensive Framework for Multilevel Mediation (Kristopher Preacher)
- Intensive Repeated-Measures Studies: Their Uses and Benefits (Martin Sliwinski)
- Plus addition free supplemental workshops sponsored by the graduate student APA organization (APAGS).
- Ten Years After the Task Force on Statistical Inference Report (Jose Cortina, Chair)
- Quantitative and Qualitative Inquiry – Beyond the Divide (Gwyneth Boodoo & Kenneth Gergen, Chairs:)
- Revising the Standards for Educational and Psychological Testing (Wayne Camara, Chair)
- Evaluating and Optimizing Measurement Comparison Across Longitudinal Studies (Scott Hofer & Kevin Grimm, Chairs)
- Role of Mixed Methods in Psychological Research (Gwyneth Boodoo, Chair)

Finally, Division 5 programming includes invited addresses and award sessions involved Gwyneth Boodoo, Howard Wainer, Lawrence Hubert, W. Holmes Finch, and Libo Li. Please check the details on the website indicated above.