



## FROM THE PRESIDENT

*Anne R. Fitzpatrick, Educational Testing Service*

Greetings, Everyone!

I hope that those of you who attended the 2007 annual meeting in Chicago enjoyed the meeting. For those of you who could not come to Chicago, I hope that we will see you next year in New York City.



Before I comment on the meeting in Chicago, I want to first extend a warm welcome to the three new members of NCME's Board of Directors. Mark Reckase is now serving as NCME's Vice President, and Larry Rudner and Todd Rogers replaced Board members Linda Cook and Terry Ackerman, whose terms ended this year.

## The 2007 Annual Meeting in Chicago

In all we had 1,404 people register for the meeting in Chicago. Of the 1,404 registrants, 383 were graduate students and 116 identified themselves as new graduate students or new regular members! Last year, a total of 1,223 people registered, of which 309 were graduate students and 59 were new members. I know that the Graduate Student Issues Committee and the Membership Committee worked extremely hard to attract students and new members; these numbers suggest that they were very successful.

Many thanks for developing an excellent meeting program are owed to Mary Pitoniak and Michael Jodoin, the 2007 program co-chairs, and to Cheryl Cardell, chair of the Training and Development Committee. Mary and Michael accepted 153 paper proposals or 52% of the total number submitted, and 27 proposals for coordinated sessions or 69% of the total number submitted. A total of 280 people attended 12 pre-conference training sessions.

The operational aspects of the training sessions and Annual Meeting ran very, very smoothly. Credit for this must be given to Bruce Wheeler, NCME's Executive Director, and other staff in NCME's Central Office. Whether dealing with missing badge holders, lost raincoats, ceiling leaks, temperamental audio-visual equipment, or temperamental NCME members (there were only two!), the Central Office staff handled matters graciously and efficiently. We are very fortunate to have their assistance, as our annual meetings are complex things to run.

We also are grateful to staff in AERA's Central Office for selecting Chicago's Intercontinental Hotel as the site for all of NCME's training and paper sessions. It is a beautiful hotel with spacious meeting rooms and a very responsive staff, which definitely enhanced the quality of our meeting.

Thanks are also owed to all presenters, chairs, moderators, discussants, and reviewers of the online proposals submitted for paper sessions, poster sessions, symposia and training sessions. Where would we be without your contributions?

## Planning for the 2008 Annual Meeting Program

Although many of you are still recovering from the 2007 Annual Meeting, some of us decided to skip the recovery period and plunge directly into planning the 2008 Annual Meeting. Among the busiest are Karen Barton ([karen\\_barton@ctb.com](mailto:karen_barton@ctb.com)) and Tasha Beretvas ([tasha.beretvas@mail.utexas.edu](mailto:tasha.beretvas@mail.utexas.edu)), who have agreed to serve as the 2008 Annual Meeting Program co-chairs, and Alina von Davier ([AVonDavier@ets.org](mailto:AVonDavier@ets.org)), who is the new chair of the Training and Development Committee. If you have ideas about ways to improve the program, please contact Karen or Tasha. If you have suggestions for new training sessions or ways to improve them, please contact Alina.

The online proposal submission system for the next year's annual meeting has just opened, and it will remain open until August 1, 2007. Proposals for papers to be presented and for training sessions can be submitted online during this period by going to NCME's website at [www.NCME.org](http://www.NCME.org).

The 2008 meeting is scheduled for March 25 to 27, 2008, in New York. This early date has posed a problem for us that we will need your help to solve. We expect that we may be unable to resolve scheduling conflicts between NCME's program and AERA's program. These conflicts arise because AERA assigns a person to a session that is scheduled at the same time as an NCME session in which the person also is involved. This year we had difficulty obtaining final scheduling information from AERA, which prevented NCME's program co-chairs from resolving schedule conflicts and issuing a final version of the NCME program until early March, 2007.

In 2008, we obviously can't wait until early March to finalize our program -- and you can't wait until then to finalize your travel arrangements. The only solution that we have been able to come up with is as follows. If you are a primary author who has, say, one paper accepted by NCME and one paper accepted by AERA, we are going to ask that you arrange for someone else to deliver one of the papers, just in case you are double-booked. If you are not double-booked, so much the better, but we need to prepare you for the possibility that you might otherwise have to be in two places at once!

To identify presenters that might potentially be double-booked, we will ask you when you submit your paper proposal online to indicate whether you are also submitting a proposal to AERA. This information will enable us to focus on the individuals with potential conflicts, although I promise that we will not ignore the rest of you! As the fall progresses, we will keep you informed about the program and the scheduling of sessions, and we will do our best to work with AERA to get the information we need in time. We just don't know whether we will be successful. In advance, however, I ask (and beg) for your patience and understanding. We will do the best we can.

## **News and Notes**

### Publications

As many of you may already know, authors should use our new online system to submit manuscripts that they wish to be considered for publication in the *Journal of Educational Measurement (JEM)* or *Educational Measurement: Issues and Practice (EM:IP)*. Details about how to use this system are provided by Jim Carlson, Editor of JEM, and Sue Brookhart, Editor of EM:IP, in this Newsletter.

The content of NCME's website is being updated by NCME's webmaster, John Hofmann ([jhofmann@ncme.org](mailto:jhofmann@ncme.org)). This will include a new "look" for the website, and a Members Only section that will include a Member Directory and the capability to renew your membership online. An *ad hoc* task force chaired by George Englehard has been developing plans for further enhancing the website so that it will provide more services and information about the science and practice of educational measurement to all who are interested.

### Calls for Award Nominations

Chairs of the awards committees have prepared calls for nominations that appear in this issue of the Newsletter. They also will be posted on the website and published in *EM:IP* in the near future. Comments from this year's award winners are included in this Newsletter.

### Current Committee Work and Opportunities

All committee chairs are in the throes of planning their work for this year and filling openings on their committees. A list of the committee chairs is provided at NCME's website ([www.NCME.org](http://www.NCME.org)). If you are interested in joining a committee, please feel free to contact the chair.

### Test Standards: National and International

NCME is very involved with an array of issues related to test standards. In this Newsletter, Dan Eignor describes the work being done by NCME's Standards Committee, NCME's involvement in the development of an international test standard, and the status of planning for the revision of the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999). Although Dan is now serving as NCME's Past President, his previous experience serving as NCME's liaison to the Joint Committee that developed the 1999 Standards has made him an invaluable advisor to us as we work on projects related to test standards.

## NCME Involvement in an Advisory Group for Morgan State University

Earlier this year, NCME was invited by Pamela Scott-Johnson, chairperson of the psychology department at Morgan State University in Baltimore, to join an Advisory Board that will help to guide a new program in psychometrics being developed at the university. The program is the product of a partnership between Morgan State University and the Educational Testing Service, and it will be coordinated by Steve Koffler. It will offer Master's and Ph.D. degree programs in psychometrics. The program is intended to develop a cadre of trained individuals with technical expertise and to increase the number of minorities involved in psychometrics. NCME's Board of Directors enthusiastically accepted Dr. Scott-Johnson's invitation, noting that it offered NCME an opportunity to encourage and support a very worthwhile endeavor. Kurt Geisinger has been appointed to represent NCME on the Advisory Board.

## Revision of NCME Bylaws

The Board of Directors has also been working on a minor revision of NCME's bylaws. We began this process because we could not determine whether the bylaws permitted electronic balloting or required that paper ballots be used. Electronic balloting was deemed attractive since it is more economical and efficient than mailing paper ballots. Also, we had learned that other associations using electronic balloting found increases in voter turnout, clearly a beneficial effect. We have developed wording to make electronic as well as paper balloting permissible, and we have identified some other, largely cosmetic changes in the bylaws that would be useful to make. We will ask NCME's membership to vote on the proposed changes in late fall when they vote on nominees for positions on next year's Board of Directors.

## Strategic Planning

In November, 2006, the Board of Directors participated in a long-range planning session in order to review and revise strategic plans established by NCME 10 years ago. In the planning session the Directors' future vision for the organization was considered in light of NCME's working mission statement, which is *"To advance the science and practice of measurement in education."*

By the end of the session the Board had defined five broad goals that it determined to be important for NCME to achieve in the next 10 years. These goals are

- Having a voice/influence in public policy and industry standards,
- Developing a broader stakeholder audience,
- Being a recognized source of expertise in educational measurement,
- Providing a broad, accessible forum for resources, and
- Providing members with opportunities for professional development and a strong professional identity.

The Board found that the strategic planning process was very helpful and productive. The goals established will inform our thinking and our decision-making and will help to ensure that the organization will grow and change in desirable directions. Over the summer we will work on short range activities that will help NCME achieve its long range goals.

In closing I want to note that NCME continues to be an incredibly dynamic organization. I take great pleasure in having the chance to help with all its endeavors, and I will continue to keep you up to date on NCME's these endeavors through this Newsletter. In the meantime, I send you my best wishes for a wonderful and relaxing summer.

---

## **AERA/APA/NCME STANDARDS**

*Dan Eignor, Educational Testing Service*

The Management Committee for the next revision to the Standards (Suzanne Lane for AERA, Wayne Camara for APA, and Barbara Plake for NCME) presented a session at the annual meeting of AERA/NCME that discussed the initial plans that have been put in place for dealing with the revision efforts. The Management Committee has chosen the co-chairs for the next Joint Committee (the committee that actually will do the revising). They are Barbara Plake, Emeritus Professor and former Director of the Buros Testing Center at the University of Nebraska-Lincoln and Laress Wise, President of the Human Resources Research Organization. Dave Frisbie has agreed to take over



Barbara's responsibilities on the Management Committee so she can devote her attention to her co-chair responsibilities. The Management Committee will be coordinating a process whereby there will be an open solicitation of comments, concerns, or suggestions that have to do with the revision effort from AERA, APA, NCME and other organizations. The period in which these comments and suggestions may be submitted to the Management Committee will be from June to October/November of 2007.

NCME, through the work of the NCME Standards and Test Use Committee, plans on asking NCME members who were either lead persons in preparation of chapters in the 1999 Standards or are authors of related chapters in the new volume of Educational Measurement to provide chapter comments to then be shared with the Management Committee. In addition, members of the Standards and Test Use Committee will provide comments and suggestions.

It is expected that the comments provided will help the Management Committee in deciding which content areas will need representation on the next Joint Committee. The plan is that individual members of the Joint Committee will be selected by May of 2008 and that the work of the Joint Committee will begin in late summer or early fall of 2008.

## **International Standard**

The initial meeting of the committee from the International Organization for Standardization (ISO) working on the International Standard discussed in previous Newsletters met in Berlin in early March of 2007. G. Harris, executive director of the Association of Test Publishers (ATP), served as representative for the ANSI Technical Advisory Group (TAG) at this meeting. The ANSI TAG is made up of representatives from five U.S. organizations, ATP, AERA, APA, NCME, and NOCA (The National Organization for Competency Assurance). ATP has taken the lead in providing the necessary management structure for the TAG. As members of the ANSI TAG, the organizations are able to send a representative or representatives to meetings of the ISO group working on the International Standard where we can give input into the process of developing the International Standard. The TAG will also be able to review and provide input to draft versions of the Standard.

We now have clarification on some of the details of the International Standard, which will actually be a set of standards not unlike those in chapters of the 1999 Standards. The title of the Standard: Procedures and methods to assess people in work and organizational settings. The scope of the standard: This standard contains requirements and recommendations for procedures and methods used to assess people in work and organizational settings. The current plan is to have a committee draft available by December 2007 with a long term target for the International Standard to be in place by July 2010.

As mentioned in previous Newsletters, there are a number of reasons why development of this International Standard is of concern to NCME and other organizations who are members of the ANSI TAG. The major concern has to do with exactly how the existence of an International Standard could impact on the AERA/APA/NCME Standards, and in particular the chapter on employment testing. As an outcome of the March meeting, G. Harris reported that it appeared to him that aligning the ISO Standard with our Standards seems quite doable, which is good news.

The next meeting of the ISO committee working on this Standard is scheduled for October 2007 in Vienna. The organizations making up the TAG will need to decide on an additional representative or representatives to attend this meeting that can focus their attention on technical details.

---

## **NCME JOURNALS NOW USING ONLINE MANUSCRIPT SUBMISSION AND REVIEW SYSTEM**

*James E. Carlson, Educational Testing Service, and Susan M. Brookhart, Brookhart Enterprises LLC*

In April, just prior to the annual meeting of NCME, the incoming editors of *Educational Measurement: Issues and Practice* and *Journal of Educational Measurement* began having manuscripts submitted and reviewed online with the Manuscript Central™ software from ScholarOne. The Publications Committee working with the two editors and our publisher, Blackwell Publishing, had studied systems prior to selecting this one. *EMIP* went online on April 1<sup>st</sup>, and since then has received 14 manuscripts. Since March 21<sup>st</sup> when *JEM* went online, 37 manuscripts have been submitted, many of them from outside the United States. Authors from other countries have especially expressed appreciation for the more efficient and less costly method of submission.

## Using the Systems

Both authors of manuscripts and reviewers use the same method of preparing to use the systems. Manuscript Central is first accessed through a URL, for *EMIP*

<http://mc.manuscriptcentral.com/emip>

and for *JEM*

<http://mc.manuscriptcentral.com/jedm>

An account is created by clicking on “Create Account” on a small tab near the upper right of the first screen. Please note that if you are working with both *EMIP* and *JEM* you must create separate accounts within each. You may, however, use the same login IDs and passwords on both, if you choose.

Both authors and reviewers must complete four screens including contact information, and it is imperative that all information with a small red “req” after it be filled in. This includes from 3 to 5 keywords indicating the users’ areas of interest. You can edit your account at any time, for example to update your contact information, by logging in and clicking on the “Edit Account” tab near the top right of the first screen. You must remember to click on “Finish” at the bottom of the last screen to save your changes.

### Authors

After you have successfully created an account don’t forget to click on “Finish” or your entries will not be saved. You will receive an email (assuming you have entered your e-mail address correctly which, by the way, must be entered 3 times when creating the account) indicating that the new account has been created.

You then log into the system and go to the Author Center where you will be provided with instructions for uploading your manuscript. It is very important to be familiar with the requirements of *EMIP* or *JEM* with respect to what must be entered into the system and how. For example, all manuscripts must be free of author identification, and *JEM* requires an abstract.

Please also note that the system allows you as an author to also specify that you are willing to be a Reviewer. If you have not been a reviewer for either publication in the past and you wish to be considered a reviewer in the future you should contact the editor (Sue Brookhart for *EMIP* [susanbrookhart@bresnan.net], and Jim Carlson for *JEM* [jcarlson@ets.org]) and send her or him a copy of your vita.

Authors are able to log into the system to determine the current status of their manuscript.

### Reviewers

Once you are in the system as a reviewer and recognized as such by the Editor, she/he can “Invite” you to review a manuscript that they deem to be in your area of expertise. You must respond to that e-mail and there are three choices near the end, one of which you must select. One indicates you “Accepted” the reviewing task, the second that you “Declined”, and the third that you “Are Unavailable.” There is no stigma attached with declining, the Editors recognize that you are not always able to do reviews, but we hope that you are most of the time because good reviewers are a scarce commodity. The system also allows editors to create reviewers’ accounts, so you may receive a request to review that includes an account already set up for you. You may edit the information in those accounts but don’t forget to click on “Finish” to save the changes.

Once you have accepted the invitation to review, the manuscript will be made available to you (assuming you have correctly entered all the required information in your Account Profile). You are given a deadline for your review (specific to each publication) and the system will generate reminders at specific times (again specific to each publication). You may, if you wish, print out the manuscript and create your review off-line. You do need, however, to upload your review in the system so it is available to the editor. And you do need to select one of the decisions that each review must recommend about disposition of the manuscript (again, specific to each publication).

*JEM* also publishes book reviews and has a book review editor, Mary Pommerich, Defense Manpower Data Center ([mary.pommerich@osd.pentagon.mil](mailto:mary.pommerich@osd.pentagon.mil)). In general book reviewers are invited to provide this very valuable service to NCME. If you know of professional books that have not been reviewed in *JEM* please contact Mary.



# ASSESSING COLLEGE READINESS: A CONTINUATION OF KIRST

Scott F. Marion and Brian Gong, *The National Center for the Improvement of Educational Assessment*

Having recently become involved in college readiness assessment issues, we appreciated Michael Kirst's (2007) call for more attention to this topic by NCME members in the last issue of this newsletter. We agreed with much of Professor Kirst's summary of the issues and call to action. On the other hand, we found ourselves disagreeing with several of key points in his analysis and with some of his suggested actions. This is not surprising since we come at this from somewhat different perspectives. We think this is why Professor Kirst called for more attention to college readiness testing by NCME so that we may bring the multiple perspectives of our membership to bear on this important problem. To that end, we view this response to Kirst as a continuation of his initial framing of the college readiness assessment problem rather than a critique.



## Standards

We agree with Kirst's (2007) general sentiment that "[e]ducation standards and tests are created in different K-12 and postsecondary orbits that only intersect for students in Advanced Placement courses" (p.4). While we might quibble and suggest that postsecondary's influence is a bit stronger than Kirst's characterization, especially for those students with four-year college aspirations, there is a noticeable separation between K-12 and college readiness standards. While we use the same word, standards, to talk about K-12 and postsecondary learning targets, there is certainly not the same expectation for *standardization* that we have come to expect with K-12 content standards<sup>1</sup>. Higher education systems within many states are largely independent of one another, and many universities do not even have coherent internal policies (standards) regarding the meaning of readiness across colleges and degree programs. State K-12 systems, conversely, have largely pursued a standards-based approach (common content and performance standards for all students). Most higher education institutions simply have not articulated a coherent set of content and skills students should have learned either upon entrance or college graduation.



Perhaps the most important reason for the divide between K-12 and postsecondary standards is a gap in governance structures in most states between the K-12 system and the higher education systems. Few states have a common governing body, and even fewer have coherent working relationships between the state agencies responsible for K-12 and post-secondary education. These unique governing boards respond to different political pressures, which leads to non-overlapping priorities, even though both sets of priorities might be entirely logical. It makes sense for postsecondary institutions to focus on having students enter with the skills necessary to avoid remedial coursework for both financial reasons and to improve retention rates. On the other hand, K-12 systems are under tremendous pressure to have all students meet a common relatively low bar (in most states) and to reduce the embarrassingly large dropout rates. Many schools and districts, particularly those located in high SES areas, have focused intently on college preparedness, but it has not been a state concentration until recently.

Kirst (2007) suggests that part of the reason for this disjuncture between K-12 and college readiness standards is because state K-12 education leaders have set standards "without talking with higher education institutions..."(p.4). Having spent much of our work lives at the state level on the K-12 side of this equation, we have experienced how incredibly difficult it is to attract and sustain the participation of university faculty members or administrators. Kirst is correct in noting that postsecondary contributors might become frustrated by the political constraints when working in the K-12 context, but walking away from the challenge does not make it go away. We and most of the state leaders we know would welcome the contributions and engaged participation of university experts to help bridge the gap between K-12 and postsecondary expectations or at least determine how much of the gap should be bridged<sup>2</sup>.

## Assessments

States typically have three main uses of large-scale assessments in high school: 1) to measure graduation standards, which typically are much lower than college readiness standards, 2) to assess grade-level or grade-span proficiency standards, which although higher than most high school graduation requirements, also are lower than college readiness; and 3) increasingly, through the use of college entrance exams by states, bringing increased attention to what it may take to be admitted to college. It is highly unlikely, despite several groups' strong advocacy (e.g., Achieve), that states will adjust their graduation standards

<sup>1</sup> We are not advocating for standardization of postsecondary learning targets or readiness standards, but this difference is important in terms of contextualizing this issue.

<sup>2</sup> It is beyond the scope of this short article to discuss whether targeting K-12 standards so that all students leave high school "college ready" is in fact the right thing to do unless we are willing to deal with significant structural and economic factors. Pragmatically, achieving universal (or at least close) college readiness anytime in the foreseeable future will make universal proficiency by 2014 seem like a walk in the park.

to reflect college readiness. The political fallout from having enormous numbers of students failing to graduate is untenable at this time. States are also unlikely to raise the current high school proficiency standards given the tremendous political pressure of NCLB. That leaves us with the use of commercial college entrance exams by states to foster college readiness. It is an option that we do not welcome, because unlike Kirst (2007), we do not find that “[a]dmissions tests send powerful and clear signals to all K-12 groups about what knowledge is most worth knowing and how it should be taught” (p. 4). Further, the use of assessments for college entrance follows a radically different path than high school proficiency and graduation tests. College entrance exams are essentially norm-referenced exams that allow institutions to sort students by relative performance, not by an assertion of performance on established content. We know that both ACT and the College Board, to their credit, have been trying to foster more standards-based interpretations of their respective test scores, but postsecondary institutions continue to use the scores normatively when considering the size and achievement level of the applicant pool along with the number of available spaces in the entering class. Kirst points out another challenge when he correctly notes that for many students, the entrance exam is less of an issue than the institutionally administered placement exam that determines whether or not students are eligible for credit-earning coursework. Both the lack of transparency of the college entrance test specifications and the idiosyncratic nature of placement exams exacerbates rather than ameliorates the K-12/postsecondary chasm. If those calling for having high schools improve students’ preparation for college, could help make it less like a game of “Marco Polo” if the expectations were more explicit so that high schools could better align their curricula (assuming that was desirable).

On the other hand, we thought that two of Kirst’s policy possibilities offered some hope. Kirst notes that while it has not occurred yet, having K-12 education leaders negotiate with postsecondary institutions about a statewide aligned assessment (or assessment system) that both will use makes a good deal of sense. In fact, in Rhode Island there are plans to allow for the results from the Proficiency Based Graduation Requirements local assessment systems to be considered in admissions and placement decisions for applicants to the University of Rhode Island system. On a wider scale, Achieve and nine of its partner states in the American Diploma Project (ADP) have designed an end-of-course Algebra II exam. Many of these state collaborators hope to negotiate with their state higher education systems to have these test results considered in placement decisions. Since the ADP standards were created with significant participation from the higher education community, this model is a concrete example of this policy possibility.

A slight twist on the previous approach, according to Kirst, is to have higher education approach K-12 leaders in order to modify an existing K-12 statewide exam to make it congruent with postsecondary expectations. Kirst notes that California State University’s (CSU) Early Assessment Program (EAP) is a good example of this policy possibility. We agree, but we also suggest that this approach required the CSU system to be very open regarding its standards for admittance to credit-bearing classes at all of its institutions. We note, however, that the EAP is voluntary, so for now it is seen by many as an “extra” test. Having an assessment such as the CSU-EAP as a required part of the California high school assessment system might be an even better example of this policy possibility—noting again that it may serve as a college placement test, but almost certainly could not be the state proficiency test.

## **Accountability**

We turn to accountability considerations, a predictable turn if we expect—and we do—to see an increase in calls for high schools to improve the college readiness of its graduates. We agree with Kirst that there are few, if any, K-16 accountability systems or components of systems to encourage more collaboration on assessment goals or to help reduce postsecondary remediation. We support the use of accountability systems to help foster the types of behaviors valued by stakeholders. However, we have seen far too many accountability systems go awry because of rushed or otherwise poor design decisions as well as limited capacity to meet the goals of the system.

We have serious concerns about some of the guiding philosophies of a potential joint accountability system. We are fearful that college readiness will become yet one more target for which high schools will be held accountable, as if the problems reside only with high schools. How will the accountability responsibility be shared with the postsecondary system? We are not trying to establish a “blame game”, but would rather see a partnership where each side is expected to live up to its end of the bargain. For example, let’s not forget that all of the high school teachers being faulted for not adequately preparing students for college learned their craft initially at some postsecondary institution, and most did so in the state where they are now teaching.

We have several questions from an accountability design perspective about the quality and validity of various indicators and potential outcomes. For example, if postsecondary placement tests are to be used as either outcome or intermediary indicators, one would need to ensure that these tests have been validated for their use as both a placement determiner and accountability indicator. Similarly, it appears that many involved in the discussion about college readiness take it as a given that college course content is valid. In other words, building a system to ensure that students are ready for college coursework assumes that the coursework is itself valid. Many of us have sat through courses that might lead us to question this assumption. Finally, there must be a coherent or at least predictable relationship between the higher education requirements from open-admissions

community colleges through the most selective public and private institutions before we can design a shared K-12/postsecondary accountability system.

We would like to finish where we began by thanking Professor Kirst for initiating this important discussion in the NCME newsletter. This multi-faceted problem can benefit from the talents and experience of the membership. We look forward to continued thinking and writing on this important issue.

## References

Kirst, M. W. (2007). Testing in a disconnected K-16 system. *NCME Newsletter*, 15, 4, 4-5. Retrieved at [http://www.ncme.org/pubs/vol\\_15\\_num\\_1.pdf](http://www.ncme.org/pubs/vol_15_num_1.pdf).

---

## THE “INTRINSIC RATIONAL VALIDITY” OF AN INTEGRATED EDUCATION SYSTEM<sup>3</sup>

Jon S. Twing, *Pearson Educational Measurement*

Michael Kirst's column, appearing in the previous edition of this newsletter, caused me to think long and hard about a concept I work with everyday...namely the validity of assessment systems. Dr. Kirst argues that K-12 and post-secondary education systems are not integrated, and most of his column is devoted to illustrating this thesis. While I think Michael and I agree generally that our educational systems are disconnected, there are several aspects of this disconnection that Dr. Kirst did not address. Hence, I wanted to give my perspective (for what it is worth), as I have spent most of my career providing support for many large-scale, statewide assessment programs, some of which include end-of-course testing, college readiness components and consideration of school-to-work transitions.



It is absolutely true, as Dr. Kirst points out, educational standards (both content standards and performance standards) as well as the assessments themselves are "...created in different K-12 and postsecondary orbits...". I am not sure, however, if this is the cause or the consequence of what he calls the "...proliferation of tests in grades 9 through 11", or frankly, if they are even related. It seems to me that there were a lot of ACT®, SAT®, and AP® assessments long before NCLB and the standards-referenced movement. Add to this the PACT and the PSAT (or whatever they evolved into), the PLAN®, Explore®, the SAT Subject Tests™, COMPASS™, WorkKeys® and clearly there is a lot of testing explicitly linking what is done in one system (secondary school) with another (college).

I also agree with Dr. Kirst when he argues that "For example, some colleges in California complain that secondary tests do not emphasize trigonometry enough." I have no doubt, having built mathematics assessments for years, this is true. Many students do indeed lack the skills required to be ready for college because their high school courses and exams are different from the expectations at college. This could be due to the watering down of high school curricula, the "Jekyll and Hyde" pressure most state programs have regarding high and rigorous standards and low failure rates, or the fact that students know they will get accepted into college without the pre-requisite skills anyway.

Yet, high school educators are working on ways to improve students' preparation for college. The high school educators I speak with are often talking about what they can do to improve ACT® and SAT® scores, how they can get students college level credit via dual enrollment, CLEP® or AP®, and how they counsel students to take the rigorous courses they will need for college. There are many examples of such notions being supported in the literature. The ACT documents: "Crisis at the Core: Preparing all Students for College and Work" (2004; <http://www.act.org/path/policy/reports/crisis.html>) and "Courses Count: Preparing Students for Post-Secondary Success" (2005; <http://www.act.org/path/policy/pdf/CoursesCount.pdf>) tell us that the rigor of the courses taken in high school is insufficient to prepare students for success in college. The expansion of end-of-course testing seen in several states might well be an attempt by secondary systems to compensate. The stated goal of the American Diploma Project by Achieve (<http://www.achieve.org/>) and its sponsoring of the Algebra II End-of-Course Assessment (<http://www.achieve.org/node/842>) is an example of explicit attempts to prepare students for the rigors of post-secondary study.

Dr. Kirst's three state policy possibilities, which I read as his recommendations, to rectify this situation however, seem to ignore the very thing that I think is the cause of the problem, namely that we have not defined what it is we are trying to

---

<sup>3</sup>The author would like to thank the reviewers of this text for their very helpful suggestions.



accomplish. I am not a student of Robert Ebel nor did I know and speak with him as many of my colleagues and readers of this newsletter might have. However, I have studied his work, and I can't help but think that Bob might call what we are trying to accomplish, the notion of a linked academic system, an intrinsically rational system. An intrinsically rational academic system would be required to survive the scrutiny of a rational human asking if such a linked system was valid. For example, if Algebra II skills are required for success in post-secondary efforts, an intrinsically rational system would suggest secondary school systems would prepare students by teaching them the same Algebra II enabling skills post-secondary requires. Similarly, such an intrinsically rational system would require the primary school system teach the content standards required as pre-requisites by the secondary school system such that students coming to high school would be ready to learn the information they needed to be successful when they got to college.

Several obstacles appear to stand in the way of achieving an intrinsically rational "all encompassing" system. The lack of a linked system is not simply the lack of the secondary system understanding or preparing students for success in college. First, the post-secondary system may have difficulty agreeing on a single set of educational standards. In working with college faculty and others in the post-secondary arena, I have observed that their fundamental independence might just preclude a single set of such well articulated standards from being developed in the first place. I doubt universities are able or inclined (perhaps not even willing) to prescribe curriculum for not just their introductory courses but subsequent courses that build on the skills acquired in these beginning courses. Just how likely are universities going to be in persuading all faculties who teach, say, introductory sciences courses, to ensure they cover the same big ideas, let alone the same scope and sequence? Compared to, arguably the sometimes flawed but nonetheless vertically articulated, statewide curriculum the post-secondary corpus of content seems disjointed.

Second, college admission programs may need to help the secondary systems by implementing well articulated and seldom adjudicated standards of acceptance or criteria of admission that do not require remediation. Yet, colleges seem to admit more and more students clearly in need of remediation who do not have the pre-requisite skills. Hence, I am a little disturbed by the concept or characterization that, somehow, post-secondary institutions are the victims of poor instruction or poor course requirements in high school. I find it unfortunate that colleges admit students who are under-prepared, and then charge the students tuition for remediation. This argument is then extended by some in post-secondary to imply that colleges are not culpable because they were not asked to sit at the table when K-12 content and achievement standards were devised. Perhaps if colleges did not enroll these under-prepared students it would create a demand in the secondary school market place for more trigonometry (because you need it to get into Cal State for example).

Given the plethora of different courses, different placement tests, different standards of performance and different graduation requirements and one might think linking these two systems together (secondary and post-secondary) would be impossible. But the issue would seem to be easily solved with an intrinsically rationally valid system which explicitly links learning:

- First, know what you want to accomplish
- Second, define learning programs to achieve what you want to accomplish
- Third, measure learning against what you want to accomplish
- Finally, feed back measures to improve learning to help gain what you want to accomplish

If the goal of high school is making students ready for college (i.e. what we want to accomplish), then this should be reflected in the learning programs and measures. Standards to evaluate the success of meeting this goal will include metrics of college success and hence, the systems will be explicitly linked.

---

## 2007 NCME AWARD WINNERS

### **Andrew Ho receives the Brenda H. Loyd Outstanding Dissertation Award**

*Mark Gierl, Committee Chair*

The 2006 Brenda H. Loyd Outstanding Dissertation Award was given to Dr. Andrew Ho for his study titled "Comparing Score Trends on High-Stakes and Low-Stakes Tests Using Metric-Free Statistics and Multidimensional Item Response Models". The criteria used by the committee to evaluate all applicants included the significance of the contribution to the field of educational measurement, quality of the literature review, technical quality of the research, and clarity of writing. Dr. Ho completed his research at Stanford University under the supervision of Dr. Edward Haertel and is currently employed as an Assistant Professor at the University of Iowa.

## Comments from Andrew Ho

Test score trends indicate degrees of progress or regress for students, teachers, schools, districts and states. Interpretations of trend statistics form the basis for newspaper headlines, scholarly reports, and educational policies. This dissertation describes counterintuitive features of large-scale test score trends that threaten these widespread interpretations. First, popular trend statistics are pliable under considerations that could be described as arbitrary: the choice of a passing score, the choice of a percentile for reporting, or the choice of a score scale. This pliability can warp trend magnitudes and, in dramatic situations, reverse the sign of the trend. Second, trend statistics for high-stakes tests are shown to be discrepant with and generally more positive than trend statistics for matched low-stakes tests.

This dissertation addresses these inconsistencies in turn. First, a metric-free methodology overcomes the shortcomings of popular trend statistics. Second, a conceptual and statistical framework models trend discrepancies as a consequence of numerous factors, including differences in content, examinee motivation, and sampling procedures. These contributions have implications for many educational controversies, including the validity of gains in high-stakes test scores, trends in score gaps between advantaged and disadvantaged students, and the attribution of test score trends to particular educational policies.

## **Jianbin Fu receives the Brad Hanson Award for Contributions to Educational Measurement**

*Jee-Seon Kim, Committee Chair*

The Bradley Hanson Award has been established to honor Brad Hanson's contributions to the field of educational measurement and to further advance the goals embodied in his work. The award is given in support of a project that promises to make a significant contribution to the field of educational measurement. The 2007 Bradley Hanson Award is given to Dr. Jianbin Fu at ETS for his project entitled *The development of a general purpose computer program for the estimation and equating of item response theory models*. Dr. Fu received the AERA Mary Catherine Ellwein Outstanding Dissertation Award for Quantitative Methods for his work on a polytomous Fusion model in 2006.

## Comments from Jianbin Fu

The project that I proposed involves the development of a general purpose collection of computer programs for the estimation and equating of item response theory (IRT) models using the freely available statistical software package R. By incorporating several different IRT models and equating procedures within the same software package, I am intending to create a very flexible product that can be easily adapted to accommodate the variety of different models currently used in IRT, as well as the various types of equating designs that lend themselves toward different linking procedures.

In the past four years I have written various R programs to carry out statistical computations for research projects on educational measurement that have involved item response theory (IRT), for example, Li, Bolt, and Fu (2005a, 2005b), and Li and Cohen (2003). These programs have been used for the following purposes: (a) test characteristic curve linking, followed by observed score equating and true score equating under the graded response model, the partial credit model, the generalized partial credit model, the nominal response model and the two-parameter normal ogive (2PNO) testlet model (Bradlow, Wainer, & Wang, 1999), and (b) the maximum marginal likelihood estimation using an EM algorithm (MML-EM) for the multiple-group 2PNO testlet model with parametric priors (Li, Bolt, & Fu, 2005a).

Importantly, these applications have involved IRT models of considerable complexity. Indeed, many other IRT models, and their associated linking/equating procedures, can be formulated as special cases or easily extended from the models/procedures I have already programmed. Consequently, I plan to extend this previous work in several directions: (a) expanding the program to handle other linking methods such as mean-mean and mean-sigma, and other IRT models such as the 1-3PL(PNO) models, the three-parameter testlet model (Wainer, Bradlow, & Du, 2001) and the polytomous testlet model (Wang, Bradlow, & Wainer, 2002); (b) expanding the program to accommodate the MML-EM estimation of the multiple-group versions of IRT models, including the two-parameter polytomous multidimensional IRT models (Moustaki, 2000; Muraki & Carlson, 1995) and the three-parameter dichotomous multidimensional IRT model (Beguín & Glas, 2001; Reckase, 1997) as well as sub-models, which compose the majority of popular IRT models; (c) refining and optimizing the codes to be object-oriented for easy maintenance and other extensions, and (d) creating an R package to hold all functions including the writing of help files, and uploading the package to the Comprehensive R Archive Network (CRAN, <http://www.r-project.org/>) as a contributed add-on package to the freeware R for users to download, as well as disseminating the program by other means such as publishing a Computer Software Exchange note on *Applied Psychological Measurement*, and/or uploading the program to a psychometric software exchange website.

The purposes of the current project are threefold: (a) to provide a unified implementation in R of a class of techniques to carry out the estimation and equating of a wide range of popular IRT models, such that a user with minimum knowledge of R language is able to use it; (b) to make open source code available so that users can reuse the functions and easily develop

variations and extensions to meet specific needs, and (c) to promote R as a valuable computing tool in the educational measurement research community. The ultimate goal is that all significant statistical techniques in education measurement have their implementations in R so that they become easy to find and use within a single common environment.

### **Joanna Gorin receives the Jason Millman Promising Measurement Scholar Award**

The 2007 Jason Millman award was presented to Joanna Gorin for her research on the intersection of cognitive psychology and psychometrics. The purpose of the Jason Millman award is to recognize a scholar at the early stages of his/her career whose research has the potential to make a major contribution to the applied measurement field.

#### Comments from Joanna Gorin

I am honored to receive the Jason Millman award that acknowledges early potential in a measurement scholar. Though I never had the opportunity to know Dr. Millman, the praise of his scholarship and character from those who knew him makes me all the more grateful. Thank you to NCME and the awards committee for this recognition.

The overarching theme of the majority of my research and goal for advancing our field is to examine traditional psychometric analyses through a cognitive lens. Years of exceptional psychometric research have provided models to estimate statistical properties of test questions. We can easily show how one test item is more difficult than others. Further, we can accurately identify test questions that function differently for distinct subgroups of examinees of equivalent ability. But these are only partial answers to the more important question we hope to answer with educational tests – what do students know? To answer that question we must go further. Why is the difficulty parameter for one item higher than it is for another? Why is DIF present for some items and not for other?

My research to date has focused on general implications of cognitive psychology for principled assessment design that yields meaningful and educationally useful test scores. Specifically, I have explored how principles of cognitive psychology, as described in the NRC's report *Knowing What Students Know: The Science and Design of Assessment*, can provide a strong theoretical basis for examination of test score meaning on standardized measures of reading comprehension. With the goal of more fully understanding the processes measured by these tests, I have employed multiple methods – correlational designs (Embretson & Gorin, *JEM*, 2001; Gorin & Embretson, *APM*, 2006), experimental manipulations (Gorin, *JEM*, 2005), and eye-tracking techniques (Gorin, *EMIP*, 2006). Results from these and related studies have identified both construct relevant and irrelevant sources of variance in student scores. Most significantly, issues have been raised regarding the emphasis on decision processes related to item structure in lieu of text representation processes more associated with reading. As greater understanding of reading assessment is gained, our assumptions regarding the meaning of reading comprehension test scores and our design of test items will likely be challenged. I am excited to see what the future may bring and look forward to contributing to innovations in this area.

Finally, it is without a doubt that any success I have had is due to my mentors and their expertise - Dr. Barabara Dodd, on polytomous IRT, CAT, and DIF; Dr. Anne Seraphine, on dimensionality and factor analysis; and most of all Dr. Susan Embretson, on cognitive models and issues of construct validity. I thank these exceptional scholars for inspiring my scientific curiosity and providing me with the skills to pursue my goals.

### **Christina Wikström receives the Alicia Cascallar Award for an Outstanding Paper by an Early Scholar**

*Mei Liu, Committee Chair*

The Alicia Cascallar Award for an Outstanding Paper by an Early Career Scholar has been established to honor Alicia's professional commitment and accomplishments and to continue her practice of mentoring and encouraging promising new scholars in the area of educational measurement.

The 2007 Alicia Cascallar Award was presented to Dr. Christina Wikström of Umeå University in Sweden in recognition of her paper entitled *Classroom Assessment and Grading - Validity Issues in the Process of Selection to Higher Education*. The paper by Dr. Wikström integrates findings from a number of studies on validity issues. It deals with complex validity issues and highlights for policy makers and educators the importance of validity considerations. The award committee believes that Dr. Wikström's work has the potential to have a wide impact on educational measurement.

#### Comments from Christina Wikström

I am very honored and delighted to have received this year's Alicia Cascallar Award for my paper *Classroom assessment and grading – validity issues in the process of selection to higher education* presented at the NCME conference in San Francisco 2006. Alicia Cascallar was a great scholar and a wonderful person and receiving an award in her memory is very special. I am truly grateful to those who nominated me and who supported the nomination, for their time and effort.

My paper discusses validity issues related to the upper secondary school grades when used for selection to higher education. The paper is based on empirical evidence provided by a Swedish research programme called VALUTA. VALUTA was a joint project between Gothenburg University and Umeå University, led by professor Jan-Eric Gustafsson and professor Christina Stage. VALUTA had the purpose to validate the Swedish admission system to higher education by using the validity model introduced by Samuel Messick (described for instance in the third edition of Educational Measurement) as an overall theoretical framework and the idea was to conduct a number of studies that illuminated various validity aspects and hence achieve a more thorough and faceted picture of this quite complex system than. I joined VALUTA in 2001, shortly after the project had started, and my part in the project was to study the characteristics of the criterion-referenced grades.

It was particularly interesting to study the characteristics of the grades since the Swedish school system was heavily reformed during the 1990s. The entire school system was decentralized and also opened up for competition between schools. The former norm-referenced grading system that had been heavily criticized was changed to criterion-referenced. This naturally changed the prerequisites for both education and assessment in all schools and at all levels. The grading of students in Swedish schools is solely based on classroom assessment, with no external control mechanisms. It is simply assumed that teachers are skilled enough to assess students' performance in an appropriate manner and assign fair and comparable grades. The public and academic discussion about school grades has traditionally mainly focused on how they affect student learning and classroom climate but until recent years, very little evidence has been known about what really is going on when grades are to be used for educational evaluation and selection.

The paper I presented at NCME discusses the reforms and the prerequisites teachers have for their grading and the empirical evidence primarily provided within the VALUTA project. The purpose was to follow the intention of the VALUTA project, by integrating the evidence from various studies (my own and other researcher's) and forming an integrated evaluative judgment about their validity as selection instruments. Even though a number of conclusions could be drawn from these studies, pointing out how the grades are lacking in both reliability and validity, it was found to be very hard to form an integrated evaluative validity judgment. Messick's model is however found useful for systemizing validity evidence. The paper concludes that adapting a broad validity perspective has resulted in that we now have a more faceted and hopefully also useful view on how the selection instruments work.

My present research is also focusing on the selection instruments to higher education. I am currently in the process of starting up a research project including also researchers from the department of Economics which aims to further investigate how upper secondary GPA and SweSAT scores predict performance in higher education, but also how the selection to higher education and traditional and non-traditional criteria for study success relates to future employment and wages.

### **Howard Wainer receives the NCME Award for Career Contributions to Educational Measurement**

*Bruno Zumbo, Committee Chair*

The 2007 NCME Award for Career Contributions to Educational Measurement was presented to Howard Wainer. This award is presented to an individual in the field whose publications, presentations, and professional activities over a career have had a widespread positive impact on the field of educational measurement. This award is in recognition of Howard's remarkable substantial and creative theoretical and technical developments, his innovative ideas that have significantly affected measurement practices, and his many applications of theory that have influenced the very nature of educational tests and measurements. Howard Wainer is a Distinguished Research Scientist for the National Board of Medical Examiners and Adjunct Professor of Statistics at the Wharton School of the University of Pennsylvania.

#### Comments from Howard Wainer

First, let me say how honored and delighted I am to receive this award. It really means a great deal to me. I wish its name were changed a little, so as to contain the phrase "Career Contributions SO FAR". Without that it seems like a retirement notice. Happily I noticed that Ron Hambleton, who received it in 1993, has in the intervening 14 years done enough new things to justify giving it to him again. I'm not finished yet either.

Henri Lombard was a Swiss physician who, in 1835, published a paper on life expectancy as a function of profession. He had accumulated over 4,000 death certificates, which had on them the man's profession (they were all men) and his age at death. From these he calculated the average age at death for each profession. He found that for professors it was 60.1, businessman 57.3, soldiers 44.5, and the title of the most dangerous profession, with an average age of death of 20.2 was reserved for students. I previously used this result in a graduation address in which I congratulated the new graduates for their having safely surviving this, the most dangerous of professions. Obviously one's profession changes over time. What you start out in is rarely how you end up, unless you perish before you can change. Which brings me to the point of my remarks about this award – a career is not a unitary thing, it changes in character over time.

When I first began I was like a car with a high-powered engine and no steering wheel. My intellectual pursuits mirrored Stephen Leacock's young man who ran out of his house, jumped on his horse and galloped off in all directions at once. I had no taste in problems (see my PhD thesis as an unhappy example). Fortunately, I recognized this and quickly indentured myself to Darrell Bock who pointed me in sensible directions. With his help I joined up with David Thissen to help Alex Roche of the Fels Institute make better use of the 50 years of longitudinal growth data they had gathered. This collaboration with Thissen, begun in the early 1970s, has continued for the intervening 35 years. A major part of this award is rightfully his.

After Chicago I went to Washington where I fell under the tutelage of Al Biderman, who explained the larger world of how statistics and the social sciences were in service to larger goals, and my view was expanded further. There is an expression I like because it says much in few words "A man with one watch knows what time it is, a man with two watches is never sure." My time in Washington gave me my second watch and made clearer the breadth of possibilities of a career in research. I began to develop good taste in research problems. Then I spent more than two decades at ETS working with Sam Messick, Henry Braun, Paul Holland, Charlie Lewis, and a host of others (many, but not all, of whom were named Wang). These were collaborative efforts in which we often had substantial overlaps in skills, but enough uniqueness to provide a synergy that yielded a sum greater than merely totaling the parts. This was a fine way to make professional contributions, and I recommend it to anyone who is fortunate enough to be a part of such a talented crew.

But things changed. I got older. Some horsepower might've leaked away, but was more than compensated for by the increase in vision. A better sense of what was important. In addition I noticed young scholars arriving with enormous horsepower but the some of the same cluelessness that characterized my own beginnings. When Eric Bradlow and Xiaohui Wang showed up at ETS in the mid 1990s I pounced. I helped set direction and they ran down the path dragging me along in their wake. And in that wake are dozens of papers and our newest book (hot off the Cambridge University Presses) on testlet response theory. This partnership has lasted more than a decade and has been a symbiosis in all regards.

And thus the final part of my avuncular peroration; aimed at scholars of my generation as well as those a little junior – let your generosity of spirit increase with your experience. I had this epiphany in 1991 when an ETS predoctoral summer student came to work with me. I had an idea for measuring testlet DIF and knew just how to go about it. I directed the student in this project and after he did the analyses I wrote it up and prepared to send it out for publication. But before I could, my wife Linda read it and said, "why did you put yourself as first author?" I replied, "because it was my idea, I told him how to do it, and I wrote it." But she was firm. "Why do you need a 201<sup>st</sup> paper as first author? How much is enough? Put Steve first, it will mean much more for him than for you." This was an academic version of "from each according to their abilities, to each according to their need." At the time I didn't think Marxism had any place in modern scholarship, but Linda can be very stubborn when she believes herself to be right and so when the article appeared it was Sireci, Thissen & Wainer. Linda's wise counsel has manifested itself ever since, as almost all papers that I've co-authored with students since then have been "and Wainer". It was a good idea in 1991 and a better idea now. I urge all of you who are of a certain age to consider following it. This policy is not merely generosity but also a reflection of your changing career roles. Academic jobs evolve from being a discoverer to being a guide and facilitator of younger scholars. Before you can do your job well you must recognize what your job is – and it isn't what it once was. Being generous will yield wonderful and unanticipated rewards. I urge you all to try it. Finally, thank you all for this fine recognition and especially thank you to all those who helped along the way: Linda, Dave, Paul, Eric, Xiaohui, Henry and all the rest. I couldn't have gotten here without you.

### **Gregory Cizek receives the NCME Award for Outstanding Dissemination of Educational Measurement Concepts to the Public. Annual Award**

The NCME Award for Outstanding Dissemination of Educational Measurement Concepts to the Public was awarded to Greg Cizek of the University of North Carolina, Chapel Hill. This purpose of this award is to honor significant contributions to the field of educational measurement and the winner is chosen based on quality, innovation, and importance of the contribution.

#### Comments from Gregory Cizek

"It is a great honor that NCME has selected my work for this award. In today's educational environment, it is important that the teachers, parents, and policymakers all have a clear grasp of how testing is essential for helping every student to achieve." *Addressing Test Anxiety in a High-Stakes Environment*, which was co-authored with Samantha Burg, provides research-based recommendations for dealing with test anxiety and covers suggestions on how to control test anxiety, information about which students are affected, and situations in which test anxiety can be helpful. "The book represents the continuation of an enduring professional goal of mine: promoting public understanding of testing. I hope that the book fosters improved testing practices and that the award stimulates others to join in achieving that goal."

## Certificates of Appreciation for Service to NCME

NCME presented certificates of appreciation to eight members for their service to the organization this year. This included:

Linda Cook	In recognition of dedicated service and outstanding leadership as a member of the Board of Directors of the Council, 2004 to 2007
Terry Ackerman	In recognition of dedicated service and outstanding leadership as a member of the Board of Directors of the Council, 2004 to 2007
Jim Impara	In recognition of dedicated service and outstanding leadership as a member of the Board of Directors of the Council, and for your service to the Council in the offices of Vice President, President, and Past President, 2004 to 2007
Mary Pitoniak	In recognition of dedicated service to NCME as Co-chair of the 2007 Program Committee
Michael Jodoin	In recognition of dedicated service to NCME as Co-chair of the 2007 Program Committee
Cheryl Cardell	In recognition of dedicated service to NCME as 2007 Training Chair
Steve Ferrara	In recognition of dedicated service to NCME as EM:IP Editor, 2004 to 2006
Robert Brennan	In recognition of dedicated service to NCME as Editor, Educational Measurement, 4 <sup>th</sup> Edition

## Awards for Classroom Assessment Practice

NCME presented awards for excellence in classroom assessment practice to the following educators:

Sean Daleiden	Elementary Teacher, Reading Among Disabled Learners, Aurora, Illinois
Suzy Dees	8th Grade Writing, Bloomington, Illinois
Debbie Davis	8 <sup>th</sup> Grade Reading, Bloomington, Illinois
JoLynn Plato	8 <sup>th</sup> Grade Math, Bloomington, Illinois
Julie Dawson	3 <sup>rd</sup> Grade Reading, Bloomington, Illinois
Melanie Stanley	2 <sup>nd</sup> Grade Math, Bloomington, Illinois
Diane Steminski	Administrator, Bloomington, Illinois
Paul Kimpton	High School Music, Hinsdale, Illinois
Jim Gardner	6th Grade Math, Naperville, Illinois
Katie Sheehan	6th Grade Math, Naperville, Illinois
Lisa Forsythe	6th Grade Math, Naperville, Illinois
Joan Benson	2 <sup>nd</sup> Grade Literacy and Math, Naperville, Illinois
Jean Carson	2 <sup>nd</sup> Grade Literacy and Math, Naperville, Illinois
Paula Smith	6 <sup>th</sup> Grade Math and Social Sciences, Naperville, Illinois
Mathew Gerwig	6 <sup>th</sup> Grade Math and Social Sciences, Naperville, Illinois
Jeff Van Harlingen	7 <sup>th</sup> Grade Writing, Naperville, Illinois
Katherine Barr	7 <sup>th</sup> Grade Writing, Naperville, Illinois
Anthony Romanelli	8 <sup>th</sup> Grade Language Arts, Naperville, Illinois



NCME presented awards for supporting sound classroom assessment practice:

Jay Linksman	Executive Director, Professional Development Alliance
Bob Nielsen	Superintendent, Bloomington Public Schools
Rebecca McCabe	State Assessment Director, Illinois Department of Education

---

## **NCME AWARDS CALL**

### **Call for the NCME Award for Career Contributions to Educational Measurement**

#### Description of the Award

The award honors living persons whose publications, presentations, and professional activities over a career have had a widespread positive impact on the field of educational measurement. These contributions may include theoretical or technical developments, service to professional organizations, conceptualizations of educational measurement that have enhanced public understanding of measurement problems, applications of theory that have influenced the nature of educational tests and measurement, or innovative ideas that have significantly affected measurement practices. Award recipients receive a check for \$1,000 and a commemorative plaque from NCME. In addition, recipients are invited to provide an address at the next year's NCME Annual Meeting.

#### Description of a Nomination

A nomination must include two items: (a) a 1-or-2 page summary of the nature, significance, and impact of the nominee's contribution to the field of educational measurement, and (b) an up-to-date copy of the nominee's vita.

#### Application Procedure

Submit eight (8) copies of all materials by November 30, 2007 to:

Krista Breithaupt, Ph.D.,  
Director, Psychometrics and Research  
American Institute of CPAs  
Parkway Corporate Center  
1230 Parkway Avenue, Suite 311  
Ewing NJ, 08628-3018

For additional information go to the NCME Website or e-mail Krista Breithaupt [kbreithaupt@aicpa.org](mailto:kbreithaupt@aicpa.org)

### **Call for the NCME Award for an Outstanding Example of Application of Educational Measurement Technology to a Specific Problem**

NCME is accepting nominations for an award for an outstanding example of educational measurement technology applied to a specific problem. The award will be based on a project completed in 2005, 2006 or 2007. Examples of applications to specific problem areas include, but are not limited to, selection or classification of students, measuring a hard-to-measure trait, evaluating an educational program or product, integrating testing and learning, or applying technology in a new way to a current problem. Selection criteria are quality and inventiveness of the application on the practice of educational measurement. Selection criteria for the award will include, quality, innovation, and importance of the contribution.

Self nominations are encouraged as are nominations for others. Individuals or groups are eligible for this award. Nominees need not be NCME members. A nomination consists of 6 copies of a 3-5 page statement summarizing the application of educational measurement technology to a specific issue, as well as an electronic version of the statement. Applicants should clearly describe and demonstrate the importance of the contribution to the field of educational measurement technology. Additional supporting documentation is welcome. Applications should include the names and addresses of two persons familiar

with the specific application and its results. The committee may request further materials and may contact others who are likely to be able to evaluate the contribution.

Nominations should be sent by December 7, 2007, to:

Dehui Xing, PO BOX 9143, STN PROV GOVT, Victoria, BC V8W 9H1, Canada  
Tel: (250) 356-7203, Fax: (250) 387-3682, Email: [dehui.xing@gov.bc.ca](mailto:dehui.xing@gov.bc.ca)

The award will be presented at the NCME Annual Meeting in New York, April 2008.

For additional information go to the NCME Website or e-mail Dehui Xing.

## **Call for the Jason Millman Promising Measurement Scholar Award Description and Application Procedures**

In 1995, the Department of Education at Cornell University initiated the Jason Millman Promising Scholar Program to honor the lifetime work of Dr. Jason Millman, to recognize his contributions to the field of applied measurement, and to continue Dr. Millman's support of scholars in their formative years who are just beginning their research careers.

Since 2003, the National Council on Measurement in Education with the support of the Millman endowment has continued the tradition of this award. As in the past, it is designed to honor Dr. Millman's work by recognizing a scholar at the early stages of his/her career whose research has the potential to make a major contribution to the applied measurement field. In addition to recognition by NCME, the successful candidate will receive \$1000. Only one candidate will be chosen to receive the award each year.

### Criteria for Eligibility

To be eligible for the award in a given year, the candidate must have:

- received the doctorate within the last five years;
- two (2) or more unique papers either accepted for presentation at an NCME annual meeting or published in NCME publications within the last five years; and
- the support of his/her professional colleagues indicating that his/her work represents a significant contribution to the field of applied measurement.

### Application Procedures

For full consideration, applications and nominations must include the following items.

1. A letter of nomination from a professional colleague who is an NCME member in good standing.
2. At least two (2) additional letters of recommendation from persons other than the nominator that address (1) the candidate's contributions to the field of measurement as a teacher, and/or as an applied measurement practitioner, and/or as a measurement researcher; and (2) the reasons that the candidate's work represents a significant contribution to the field of applied measurement.
3. Two (2) or more unique papers presented at any of the last five NCME annual meetings or published in the last five years in an NCME publication. Please note the following:
  - A. NCME annual meeting papers may be in a revised format.
  - B. Non-published or published dissertations do not qualify as papers, although papers developed on the same or similar topics may qualify.
  - C. Candidate must be the first author on all multiple-author papers. In this case, the candidate must provide a statement that defines his/her contributions to the paper.
4. Candidate's current curriculum vita.
5. A letter from the candidate outlining his/her career goals and how his/her work contributes significantly to the field of measurement.

### Deadline

Deadline for submission is November 5, 2007. All materials must be submitted together as a set for receipt no later than this deadline date. Only complete sets of materials will be considered. The Committee will acknowledge receipt of candidates' sets of materials.

### Method of Submission

Submit materials to the Committee Chair at the address given below that is appropriate to your method of delivery. Candidates may submit materials using one of the following methods of delivery:

- A. **Electronic.** One (1) copy is required for materials submitted electronically. These should be sent by email to [lgamache@the-areta-group.org](mailto:lgamache@the-areta-group.org)
- B. **Standard mail or overnight courier delivery.** Seven (7) copies are required for materials submitted as hard copy.
- C. **Both methods.** If multiple methods will be used for delivery, please notify the Committee Chair of the methods and expected arrival dates for the materials. Arrival dates must precede the November 5, 2007, deadline.

Electronic submissions should be sent to:

[lgamache@the-areta-group.org](mailto:lgamache@the-areta-group.org)

Hard-copy submissions should be sent to:

LeAnn M. Gamache, PhD  
Assessment, Research, and Evaluation  
The ARETA Group  
Box 2553  
Centennial, CO 80161

The award will be presented at the NCME Annual Meeting in 2008.

## **Call for the Brenda H. Loyd Outstanding Dissertation Award**

### Description of the Award

The Brenda H. Loyd Award honors an outstanding dissertation in the field of educational measurement. The winner of the award will receive \$1,000 and a commemorative plaque from NCME. In addition, the advisor or committee chair for the award-winning dissertation will receive a letter of congratulations.

### Description of a Nomination

Nominations will be accepted for dissertations completed between July 1, 2005, and June 30, 2007. The author of the dissertation need not be a member of NCME. However, the author's advisor must be a member of NCME. Nominations must include: (a) a letter of nomination from the author's advisor; (b) a summary of the dissertation research (up to 10 double-spaced pages), including the rationale for the study, research questions, methodology, results, and conclusions; (c) a table of contents (including a list of table and figures); and (d) a statement from the graduate school confirming the date of completion and acceptance of the dissertation.

### Application Procedure

Submit seven copies of all materials by Friday, November 16, 2007 to:

Mark D. Shermis  
Professor & Chair, Department of Educational Psychology  
1403 Norman Hall  
P.O. Box 117047  
University of Florida  
Gainesville, FL 32611-7407

**For Additional Information:** Go to the NCME Website or e-mail Mark Shermis at [mshermis@ufl.edu](mailto:mshermis@ufl.edu)

## **Call for the Bradley Hanson Award for Contributions to Educational Measurement**

### Description of the Award

The Bradley Hanson Award has been established to honor Brad Hanson's contributions to the field of educational measurement and to further advance the goals embodied in his work. Nominees must propose a project that promises to make a significant contribution to the field of educational measurement and/or to the development/instruction of new professionals in

the field. A typical time frame for the completion of the project is one to two years. The recipient will be awarded \$1,000 and a commemorative plaque from NCME.

#### Description of a Nomination

Nominees may be non-NCME members. However, the recipient must be an NCME member by the time the award is presented at the Annual Meeting in 2008. The award is open to individuals at any stage in their careers, including graduate students, and may be awarded to either a single individual or a group of individuals.

Nominations for the award must include: (1) a letter of nomination. The nomination letter must describe both the candidate and the specific project to be supported, including how the \$1,000 will be spent, and should be 1500 words or less. Self-nominations are welcome; (2) the candidate's curriculum vitae; and (3) at least one additional letter of recommendation addressing the qualifications of the candidate and the importance of the project.

#### Application Procedure

All materials should be received by November 1, 2007. The nomination letter and curriculum vitae may be submitted electronically to [j.delatorre@rutgers.edu](mailto:j.delatorre@rutgers.edu). The recommendation letter(s) should be sent as hard copies by regular mail to the following address:

Jimmy de la Torre  
Department of Educational Psychology  
Rutgers, The State University of New Jersey  
10 Seminary Place  
New Brunswick, NJ 08901  
Attention: NCME Bradley Hanson Award

#### For Additional Information

Go to the NCME website or refer to Brad Hanson's homepage at <http://www.b-a-h.com/> or e-mail Jimmy de la Torre at [j.delatorre@rutgers.edu](mailto:j.delatorre@rutgers.edu).

The award will be presented at the NCME Annual Meeting in New York City, March 2008.

Donations to the Bradley A. Hanson Memorial Fund may be made payable to the Arizona Community Foundation and sent to:

Arizona Community Foundation  
Bradley A. Hanson Memorial Fund  
2122 E. Highland Ave., Suite 400  
Phoenix, AZ 85016

### **Call for the Alicia Cascallar Award for An Outstanding Paper by an Early Career Scholar**

#### Description of the Award

Alicia Cascallar, who published most of her work as Alicia P. Schmitt, is best remembered for her prolific work in the area of differential item functioning (DIF). She brought passion to her work, and was determined to make sense of DIF findings in an empirically sound manner. Her legacy of methodological and empirical contributions to this area of research provides an important base for future explorations in the area of test fairness. Alicia was an active and visible member of NCME who mentored and encouraged the participation of individuals in the early stages of their careers in educational measurement.

The Alicia Cascallar Award for an Outstanding Paper by an Early Career Scholar has been established to honor Alicia's professional commitment and accomplishments and to continue her practice of mentoring and encouraging promising new scholars in the area of educational measurement. The award will be given to an early career scholar who presented an outstanding paper at the Annual Meeting. A cash award of \$500, a citation, and a waiver of NCME conference fees for the following year will be provided as partial support for an early career member of NCME to travel to the annual meeting.

#### Criteria for Eligibility

To be eligible for this award the individual must have presented a paper at the most recent NCME meeting in a paper session or as part of a symposium or panel discussion. The author(s) must be an early career member of NCME (received their doctoral

degree within 5 years of the annual meeting). In addition, professional colleagues who believe that this work represents a significant contribution to the field of applied measurement must endorse the paper.

### Selection Criteria

Nominations must include the following items.

1. A letter of nomination from a professional colleague who is a member in good standing of NCME.
2. At least two additional letters of recommendation (in addition to the nominator) are required.
3. All endorsements of the papers should list the reasons why this candidate's paper represents an important contribution to the field of applied measurement.
4. The candidate's current curriculum vita.

Papers will be evaluated for their scientific merit, clarity and completeness, the extent to which the material is redundant with previous publications and presentations, and the relevance of the work to practitioners in the field.

### Application Procedures

After the Annual Meeting, the nominator will submit his/her nomination letter and the curriculum vitae of the candidate to the address of the Committee Chair. The letters of the other endorsees of the paper should be submitted to the Committee Chair under separate cover. The candidate will submit their paper electronically to the e-mail address listed below.

For multi-authored papers, the candidate must be the first author and must provide a statement that defines his/her contributions to the paper. All other co-authors must provide a statement indicating that the first author was responsible for at least 75% of the effort.

### Deadline

Deadline for submission is **November 1, 2007**. All materials must be submitted by this deadline date. Only complete sets of materials will be considered. The Committee will acknowledge receipt and notify the nominator and candidate if any materials are missing.

### Method of Submission

The candidate should submit the paper electronically to this e-mail address: [mary.pommerich@osd.pentagon.mil](mailto:mary.pommerich@osd.pentagon.mil). Fax versions of papers will not be accepted. Nominators and other endorsers of the paper may submit their letters either by e-mail to [mary.pommerich@osd.pentagon.mil](mailto:mary.pommerich@osd.pentagon.mil), or by mail to the Committee Chair at the address listed below.

Mary Pommerich, Chair  
Alicia Cascallar Award Committee  
Defense Manpower Data Center  
DoD Center Monterey Bay  
400 Gigling Rd  
Seaside, CA 93955-6771

For additional information go to the NCME Website or e-mail Mary Pommerich at [mary.pommerich@osd.pentagon.mil](mailto:mary.pommerich@osd.pentagon.mil)

Contributions to the Fund can be made to:

*The Alicia Cascallar Memorial Fund/PACF  
15 Princess Road  
Lawrenceville, NJ 08648*

The award will be presented at the NCME Annual Meeting in 2008.

# CHINESE COLLEGE ENTRANCE EXAMINATION: PROBLEMS AND REFORMS

*Ruoling Zheng, Institute of Education, Xiamen University, China*

## History of Chinese College Entrance Examination (CCEE)

CCEE is probably one of the most unique systems in the world, it was established in 1952, just 3 years later than the establishment of P.R. of China. During these three years, conformed to the needs of political and economic transition of our new country, the college entrance policy had also experienced a period of transition, from separate mode to united one.

In 1949, New China just followed college entrance policy from Jiang Jieshi's government, which was a separate mode because of the need to maintain the continuity of Higher Education (actually, the Jiang Jieshi's government ever tried the national examinations in 1938-1940, but the experiment was cut off by war). Under the separate mode, the admission size, recruiting criterion and other policies were all decided by colleges themselves, similar to US policy.

In 1950, colleges were required by the Ministry of Education to adopt united examination within the regions they located, but separate admission was also allowed, just in case some colleges might have difficulties in practicing the united examination. Since people were accustomed to the separate admission, only 36% of the colleges adopted the united examination that year.

In 1951, colleges were required to follow the admission policy of last year, besides, the Ministry of Education issued a new command, required that united admission must be finished before separate one, which means your competition on freshman might be disadvantaged if you didn't participate in the united admission. Under this circumstance, the amount of united-recruiting colleges increased sharply to 70% this year. Based on the successful experiences of these two years, the Ministry of Education made a decision that all universities and colleges must take part in national examination in 1952 (except some special colleges, like music or fine arts institutes).

From 1952 to 1965, the national examination system had been running very well (except 1958, the college entrance mode was changed from national one to separate one, because of some political reasons). In 1966, with the terrible political movement "Cultural Revolution" erupting, the national examination system was unfortunately chosen as a breakthrough in the area of cultural and educational "revolution", and was broken off from 1966 to 1976. Actually, almost all normal educational activities were stopped at that time.

Although some colleges resumed recruiting students in 1972, they used recommendation system instead of national examinations, which including four steps " " (register freely—recommended by common people—approved by leadership—reviewed by colleges). Under this policy, one's political performance, so called " " (class consciousness) and practical experience were most important criteria for admission. On the contrary, the academic requirement was placed at a very low level (the recommended candidates only needed to be graduated from junior high school or up).

In 1977, with Deng Xiaoping's effort and most people's desires, the national examination system which had been broken off for 11 years long was readopted, this action was a very important part of eliminating the bad influence of Cultural Revolution. The year of 1977 is so important in the history of Chinese higher education or even in the history of China society that it is regarded as a milestone. We had 5,700,000 examinees cherished their hopes for many years took the national examination in 1977, and 5,900,000 similar examinees in 1978. Some of them were brothers and sisters, students and teachers, husbands and wives, or even sons and fathers! From then on, everyone is equal to compete for Higher Education opportunities, based on their literacy, not on their parentage or so-called political performance anymore.

## Why Established National Examination System?

First, it was the needs of political and economic construction of our new country. In the beginning of establishing new China, we needed to rebuild almost every aspect of the society, and thus needed all kinds of professionals. The past experiences of higher education showed that, national examination system could solve the unbalanced recruiting problems between different regions, colleges and disciplines, and could increase the use efficiency of limited higher education resources, especially under the very poor educational condition at that time.

Second, it was the need of educational adjustment. In 1952, we had a very large-scale adjustment of academy and discipline which involved many universities and colleges. So, we needed a more scheming entrance policy to ensure all kinds of institutions get eligible and sufficient freshman, in order to solidify the fruit of this nationwide educational adjustment.

Third, it was the need of the development of exam system itself. China has such a big population, the scale of college entrance exam is very large, and seeking equity and efficiency is one of the most important considerations for large-scale exam.



Obviously, the national exam organized by the government is more efficient than the separate one held by colleges themselves, as concerning economic efficiency, scientificity and authority as well. Besides, it can also get rid of jamming stuff effectively.

Fourth, it was the desire of common people. In hundred years before the establishment of new China, our country had been being in situation of falling to pieces and infirmness. Everyone desired for the country's unification and mightiness under this situation. This kind of desire had very deep effect on the implementation of centralism in many aspects in China, including the college entrance policy.

Finally, it was influenced by culture. One factor was the idea of union which caused by thousands of years of centralism, as I mentioned above. Another factor was the "gene" of united exam originated from the Imperial Examination System, which had been existing for 1300 years long, from Sui to Qing dynasty. These factors have rooted deeply on Chinese people, mentally and culturally, which made them easily accept and recognize the national exam. And this might be a deep-seated reason for implementing the national exam smoothly. And also, this might make sense to explain why China Mainland and Taiwan chose the national examination system almost at the same time, though the two sides were totally isolated from each other at that time. Taiwan established its united exam system in 1954, just 2 years later than Mainland did.

## **Importance of National Examination in China**

The national examination has a high reputation in China. Many people especially those who had experienced the Cultural Revolution, cherish a very deep emotion with it. People even think readopting the national examination in 1977 saved our nation from collapse.

The national examination has a very close relationship with politics. Every time it was abolished or readopted, one can find a political reason. In some cases, the national examination even caused the highest level of leadership's attention. It also caused a continuous argument from academe or common people, and the viewpoints of both sides debate furiously. Whether its close relationship with politics, or the furious arguments it caused, both of them reflect its huge influence on China society.

As to its effect on education, the national exam is actually a "baton" of education, and has caused the phenomenon of " " (extremely pursuing the admission ratio) since 1960s, which hurts the education badly, and brings very heavy physical and psychological pressure on students.

The main reason why the national examination is one of the most focused and important system in China is because, the performance of exam is the most important or even the only criterion of college admission, in other words, it is related directly with one's possibility of entering into colleges, and entering what kind of or which level of colleges, thus it was titled as like " " (National Great Examination) by media, or things like that, the period of exam is also compared to " "(Examination Festival). During those 3 days, you can see all kinds of news or programs about it, from CCTV to local TV, from newspapers, radios to internet. Almost all important media focus on it. Sometimes, the bus routes or even air lines were changed temporarily, just in case of noise. And, inquiring about how well their children did in the exam is also a very popular topic when people meet their friends, relatives or colleagues. Anyway, college entrance examination is one of the hottest and most enduring topics in China.

Since the national exam is so highly competitive, the influence is so huge, the relationship between the exam and individual future is so close, and the examination scale is so large (in 2006, we had nearly 10 million examinees), any tiny reform will influence millions of examinees and their families, even only one province's reform, might also influence hundreds of thousands of examinees, to say nothing of the huge population of students from other levels. Thus, reforms of college entrance examination system are very complicated and hard, we have to be very cautious and considerate. If we only pay attention to one side and ignore the other, or just abolish it simply, it would be very easy to cause huge emotional vibration from some social classes, and hurt social stabilization.

## **Dilemma**

We have been facing a dilemma when choosing an appropriate college entrance system all these years. On the one hand, the national exam is probably suitable for us, as I mentioned just now, but on the other hand, it does exist with shortcomings.

The biggest shortcoming of national exam is lacking flexibility. For example, different type and level of higher education institutions have different requirements on their students, but maybe they could not pick out the right students they want just according to their performance of exam, especially one-time exam, in other words, it is hardly possible for the national exam to do flexible and personal assessment. It is also difficult to examine one's morality or other special abilities. In America, perhaps you can assess one's morality or other abilities through recommendation letters or other documents, but this seldom works well in China, because of dishonesty, corruption or relationship things, which is the bad side of our traditional culture.

Actually, we have a very old saying about this situation, came from the argument about Imperial Examination, about 1000 years ago. It was pointed out by Su Shi, a very famous litterateur of Song dynasty and a cabinet minister as well. it says:“ ” ” (if the criterion of selecting Civil Servant is filial piety, the brave person will slice off their thighs as prescription, even the less brave people will also be in mourning for a few years). If the criterion is incorruptness, they will take worn-out carriages, wear rags and just eat simple food. Anyway, they will do whatever can satisfy the emperor. Under this traditional culture, we have to use rigid examination to prevent from the intervention of non-rigid things, and this is also one of the main reasons why Imperial examination system had existed so long a period in Chinese history.

Theoretically speaking, it is best if we can not only recruit students with satisfying special requirements of institutions and desires of examinees, but also ensure the efficiency and equity. But the problem is, if each college holds the entrance exam separately, the cost of exam will overload much more than its payment capability, and will also make it exhausted in dealing with all kinds of relationship stuff, and thus exists with very big risk of hurting equity.

## Reforms in Recent Years

We have been trying to reform national exam since mid 1980s because of its stated shortcomings. The intention of the reforms is to strengthen colleges' self-determination power of recruitment, and make the system more flexible. Some thought about abolishing the national exam in favor of individual admission, like colleges in US, but this view is very challenging.

As mentioned earlier, we ceased the national exam twice (in 1958 and in 1966) and replaced it with recommendation, both of which caused a sharp decline of higher education quality, so the exam was readopted. In 1986, we established the recommendation system, which was parallel to the national exam, but the proportion of this recruitment is quite small (I call it a real recommendation system, which is different from the one we used in the past, the latter was totally influenced by political factor, while the academic preparation was almost ignored).

Under this real recommendation system, those excellent or talented students can directly enter into colleges with the recommendation from their middle schools. Though they don't need to take part in the national exam, they have to pass some special evaluation or exam held by the colleges.

The recommendation system worked very well in the beginning, but it did not last. Like bogus information, recommending ineligible students, wire-pulling, corruptions, etc. To fix these problems, the Ministry of Education had to set up another special united exam which was called“ ”(Testing for integrative ability) for the recommended students ( but the winners of National Olympic Subject Games can be exempt), which means we have returned to united exam in some sense. At the same time, the Ministry of Examination had to shrink the scale, heighten the criterion and strict the administration of recommendation system, and make the process brighter and more transparent, with the intervention of media and the monitor of public opinions.

Another big step of diversification is to reform the exam subject mode which was called “3+ X” from 1999. In the past, the exam subjects were divided into two groups: one was for students who wanted to study the arts, another one was for science and engineering. Subjects in each group were fixed. In the new pattern, though the “3”, which are Chinese, Mathematics and foreign language, is still required on everyone, other subjects can be flexible to choose. Each province decides its own subject mode (most provinces choose “3+ / ”).

In 2001, we also started an experiment of “self-recruitment”, under this reform, the colleges recruit freshman with evaluating the applications and take their national examination scores into account as well. The biggest argument of this reform is still about the equity problems. The self-recruitment reform was pushed onto a higher step by Fudan University and Shanghai Jiaotong University this year, which totally got rid of the national exam scores. Without national exam, then what matters in the process of admission? High school grades and the score of testing conducted by the colleges themselves did.

In 2004, we also projected a reform on making exam questions by province, which used to be made by Ministry of Education, and this has also broken the national pattern further.

Besides, the unbalanced admission marks among provinces is also a big issue of the national exam reform, it has brought huge debates since 1990s, actually, it has been being a proposition of “ ” (The National People's Congress and Chinese People's Political Consultative Conference) in recent years. This year, China University of Political Science and Law decided to distribute the admission quota to each province, mainly based on population, which was regarded as an “ice-breaking action.”

Other reforms in the last 20 years include internet matriculation, standardized examination, score scale reform, high school graduation examination reform( ), and so forth.

## EVENTS OF REGIONAL INTEREST

## **Northern Rocky Mountain Educational Research Association (NRMERA) – 2007 Conference: Making Connections**

The 2007 NRMERA conference, *Making Connections*, will be held October 4-6 2007, in beautiful Jackson Hole, WY. The conference provides an excellent forum for graduate students and emerging researchers to present their work. The Call for Proposals will be available April 1<sup>st</sup> with proposals accepted until June 1<sup>st</sup>. Check out the NRMERA website ([www.nrmera.org](http://www.nrmera.org)) for upcoming conference announcements and details! Contact Chad Buckendahl ([cbuckendahl2@unl.edu](mailto:cbuckendahl2@unl.edu)) with any questions about the conference.

## **23rd Annual Washington State Assessment Conference**

The 23<sup>rd</sup> annual Washington State Assessment Conference will be held December 6-7, 2007 at the Seattle Airport Hilton Hotel Conference Center. A pre-conference training day will be held on December 6. This year's theme is "Testing Assumptions." The keynote speakers will be:

Laura Lipton, international consultant focusing on effective and innovative instructional practices

Tony Alvarado, former Chancellor of the New York City Board of Education and District 2 Superintendent

The conference, in addition to the keynoters, will feature over 50 breakout sessions presented by local educators, as well as members of the staff of the Washington State Superintendent of Public Instruction. Registration information will be available on the WERA Web site in September at [www.wera-web.org](http://www.wera-web.org).

## **Eighth Annual Maryland Assessment Conference: Alternate Assessment**

The Eighth Annual Maryland Assessment Conference: Alternate Assessment will be held in the Grand Ballroom, Stamp Student Union, University of Maryland, College Park, MD on October 11 and 12, 2007. Registration and breakfast is from 7:00 to 8:30 a.m. Presentations will start at 8:30 a.m. and will finish at 5:30 p.m. on both days.

The 2007 Maryland conference will bring together prominent national experts to explore (1) the nature of the construct(s) that alternate assessments are designed to measure, (2) the unique assessment challenges that alternate assessments pose, (3) a range of approaches to these challenges that have been or are likely to be successful, and (4) documentation of the success of alternate assessments.

Inquiries regarding registration and Attendance should be directed to Mr. Ricardo Morales at 301-405-3629 or [RMorales@umd.edu](mailto:RMorales@umd.edu). The registration form is on the WEB at: <http://www.MARCES.org> or at <http://www.education.umd.edu/edms/events>.

### **NEWSLETTER ADVISORY BOARD**

MARY LYN BOURQUE, Mid-Atlantic Psychometric Services  
SUSAN BROOKHART, Brookhart Enterprises LLC  
SUSAN L. DAVIS, Buros Center for Testing  
ELLEN FORTE, edCount LLC  
SARA S. HENNINGS, Pearson Educational Measurement  
JOAN HERMAN, CRESST/UCLA  
THEL KOCHER, Edina Public Schools, Minnesota

WENDY MCCOLSKEY, SERVE  
GERALD MELICAN, College Board  
CAROL S. PARKE, Duquesne University  
S.E. PHILLIPS, Consultant  
DOUGLAS RINDONE, Consultant  
GARY SCHAEFFER, CTB/McGraw-Hill  
DONNA SUNDRE, James Madison University

SCOTT BISHOP, Editor, Data Recognition Corporation  
*Send articles or information for this newsletter to:*

Scott Bishop  
Data Recognition Corporation  
13490 Bass Lake Road  
Maple Grove, MN 55311

Phone: 763.268.2029  
Fax: 763.268.2529  
e-mail: [SBishop@DataRecognitionCorp.com](mailto:SBishop@DataRecognitionCorp.com)

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.