

FROM THE PRESIDENT

Wim J. van der Linden, CTB/McGraw-Hill

The preparations for the 2014 Annual Meeting are in full swing. While I write this message, Paul De Boeck, Kathleen Scalise, and Billy Skorupski are in the process of finalizing their programs. The number of 772 proposals for presentations we have received this year has been amazing. Except for a few software glitches that had to be fixed on the fly, the reviewing process has been smooth. And thanks to additional meeting space and changes in our meeting format, Paul and Kathleen will be able to accept more than 500 of them, roughly the same number as the total of all proposals last year! In addition, there will be six invited lectures and two debates. Billy has received 22 submissions for his training program. Unlike earlier years, all sessions will be held on the two days before our main program, which means that he will be able to accept somewhere between 14-full-day and as many as 28 half-day sessions.



Our 2014 conference hotel is already known, it'll be the Loews, 1200 Market, in Philadelphia. It is not yet known if the training sessions will be at the same hotel. If we need to move to another location for them, I will let you know in my next message.

You may not be aware of this, but our agreement with AERA for its provision of annual meeting services to NCME ended immediately after our last meeting in San Francisco. AERA is willing to extend our agreement for three more years. But during a few recent meetings with its central office, it has suggested that we should think creatively about our long-term plans. In fact, the same discussion had already emerged during the Past President's Luncheon at our last annual meeting. A committee existing of David Frisbie (chair), Terry Ackerman, and Anne Fitzpatrick has accepted the charge to obtain an inventory of all the pros and cons for several scenarios beyond 2016, including the one of just asking AERA for another extension of our agreement. The committee will deliver its report to the Board by the end of January. Of course, during the process of decision making that will follow, the Board will rely heavily on input from all of our members, most likely in the form of a survey.

One of the recent actions of the Board has been to make the NCME Career Contributions Award more special. For instance, our future Award winners will be selected early in November, and the Award ceremony will include an invited presentation by the winner in a special session highlighted in the announcements of the annual meeting program. Also, the new Award will be bi-annual. A new committee under the leadership of Suzanne Lane will implement the new plans.

Finally, the Board has appointed Brett Foley, Alpine Testing Solutions, as our new Website Editor. Brett will succeed John Willse, our current Editor. We are grateful to John for all the important contributions he has made to NCME in his role as Editor over the several past years. The Board has used the transition to reorganize the website editorship. Basically, the Editor will now operate entirely analogously to our JEM and EM: IP Editors; that is, with an Editorial Board appointed by Brett and under the responsibility of the Publications Committee. The change will allow the Website Management Committee to focus on the more technical aspects of website management.

In my next message I will be able to inform you about another few items on the agenda of our recent Board meetings, including our plans to work on an update of the NCME strategic plan. Until then, please accept my best wishes for continued success in your work and your other services to the field of educational measurement.

GREETINGS FROM THE EDITOR

Susan Davis-Becker, Alpine Testing Solutions

Happy **Fall** to everyone! In this issue our President provides us with an update on some very recent events and news from NCME leadership. Melinda Montgomery presents some keys to the dissertation process from members of NCME who have overseen numerous graduate students navigating this process. Our Spotlight member is Liru Zhang from the Delaware Department of Education. Also in this issue, we take the opportunity to explore one of NCME's goals concerning public policy. We were fortunate to have some members provide their perspective on how NCME leadership and members can help advance this part of NCME's mission. I encourage everyone to think about goal and how they may be able to contribute to the advancement of the organization and the influence it has within public policy settings. April Zenisky from the Website Management Committee provides an inside look the glossary of terms from our field available on the NCME website – what a great resource! Looking forward to the 2014 Annual meeting, our program co-chairs have provided us with information about the identified keynote speakers and our Fun Run coordinators are looking to collect information to help them prepare for this great event.

GRADUATE STUDENT CORNER

THE DISSERTATION: CONCEPTION TO COMPLETION

Melinda Montgomery, University of Kansas

As I approach the end of my graduate student years there is one major task to complete - the dissertation. I have personally found the process to take longer than I originally planned. In part, that is due to some life issues that were outside of my control but I have also myself in a bit of a free-fall after completing the comprehensive and oral exams. The hold-up in my dissertation process tends to fall into two general areas; (1) an incomplete understanding of how to plan and scope the dissertation project, and (2) how to get back on track when external circumstances press in and derail the process. I thought it might be a nice idea to speak with several faculty members in our field who have worked with graduate students in the dissertation completion process to see if there were some keys to success or hints that they could pass along to us as graduate students. Dr. Davis-Becker arranged for me to contact the following faculty members; Dr. Sara Finney (James Madison University), Dr. Gregory Cizek (UNC-Chapel Hill), and Dr. Steve Sireci (University of Massachusetts Amherst).



Choosing the Topic

All of the faculty members provided some very useful suggestions as to how to choose a dissertation topic. Dr. Davis-Becker mentioned that as a graduate student, one of the best pieces of advice she was given was not to think of the dissertation as the greatest piece of work she would ever produce because in doing so it will never get done: "Perfection is the enemy of completion." That reminded me of something that Dr. Kingston has been known to say: "The only good dissertation is a finished dissertation." If I allow myself to think of the dissertation as my greatest work, then that implies that if I were to plot my academic career over time it might look something like a right skewed distribution with my dissertation at the top of the peak. This visual image is certainly not what I want my career to look like.

Dr. Cizek said that whereas it is understandable to want the dissertation to make a contribution to the field, it is more important that the research project be well defined and demonstrate that as graduate students we can design and complete a rigorous research project. The dissertation, then, is yet another opportunity to learn and refine the skills that are needed for the career. Hence, choosing the topic is going to be an important part of the process, but the process itself is an important aspect in and of itself.

Dr. Sireci suggested four factors to consider when choosing a topic: (1) something that really interests you; (2) the significance of the topic to others; (3) "do-ability" (i.e., do you have the resources [e.g., data] and time required to do the research?); and (4) whether the topic will be of interest to one of the faculty members in your program. In choosing something that interests you, Dr. Cizek recommends reading a lot of literature and keeping a notebook of interesting

problems from that literature and from your courses. This will give you a collection of interesting problems to consider when choosing a topic. While I did not keep a notebook per se, my dissertation is born out of a test type issue that has intrigued me since my first measurement course.

Dr. Finney has three questions that she asks students to think about in order to judge the need for the study and how this study is situated within the existing literature:

What is the purpose of the study?

Why does the study need to be completed?

What theory or theories inform this work?

I used a similar set of questions to define my final topic. Once the relevance of the topic has been established, Dr. Finney recommends that students create a rough outline of the introduction (Chapter 1), as that serves to kick-start the writing process and provides a crude framework that helps to focus the dissertation. Next, she recommends that students write the literature review (Chapter 2), which, in addition to providing a summary of the domain of study, will inform the introduction and methods chapters. In some programs, a comprehensive paper written at the end of the course work might serve as the literature review. I have heard some faculty say that unless you plan to publish the literature review it only needs to be comprehensive enough to place your topic in the existing literature. I think a good start would be to have a discussion with your advisor about the scope of the literature review expectations for your individual topic and program. Completion of a high quality literature review might take several months so that needs to be factored into the timeline.

Timeline

All three of the faculty that I spoke with mentioned the importance of a timeline. Interestingly, I had never thought of planning the dissertation in this way. I am not saying that my advisor or other faculty in my program did not suggest this as an approach. Rather, I am saying that if they did, that idea did not register with me. Even if you are nearing the end of the process, as I am, I think this sort of timeline approach can be very beneficial.

The creation of the timeline is a process of working backwards from the date that the proposal is due, initially, as well as when you want to complete the dissertation process as a whole. The timeline will provide a realistic date when you can graduate. As an example, Dr. Finney said if a student wants to propose at the beginning of October that may mean the student needs to begin writing in May so that the student can engage in the iterative draft process and garner the necessary feedback along the way to create a high quality proposal.

This timeline should be created with your advisor and be specific enough to move the project forward including specific dates and deliverables. Turnaround time should be included in the timeline. Dr. Cizek and Dr. Finney, when serving as the advising chair, have students submit each chapter one at a time and help them to get it edited and locked down before the next piece is submitted. When the proposal--and the final dissertations project for that matter--is complete, Dr. Sireci recommends giving the committee at least two weeks to read and review it.

The Scope of the Project

When outlining the scope of the dissertation, Dr. Finney recommends focusing on the research question (rather than say the number of pages or number of articles cited). Your advisor can work with you to assess if the research question needs to be narrowed or expanded. With that said, she noted the work should be able to be completed within an academic year. However, as Dr. Sireci pointed out, that timeline is influenced by funding and family situations. If a student already has a full-time position after completing the proposal there may be less urgency in completing the final project; although completion of the proposal prior to exiting the graduate program is usually recommended. In the area where I live finding a full-time job will likely require moving. If moving is part of the overall plan, then the length of time taken to complete the dissertation project will be a factor in the larger plan.

In terms of the overall scope of the dissertation project, I liked the way Dr. Finney put it. She said that if the literature review chapter is removed from the dissertation, the remaining text should read much like a journal article (i.e., the dissertation should need editing, not re-writing, to create a manuscript for publication). Most of us have presented research at conferences and submitted articles for publication, thus the scope of an article should not be foreign. Although the dissertation will be larger than most of our previous work, it should only require editing to transform it into a publishable manuscript. With that in mind, Dr. Sireci said that we should think about getting publications out of it or in his words, "milking it for all it's worth."

Overcoming Setbacks

It is not uncommon for students to hit one or more obstacles along the way to dissertation completion. There are all sorts of things that can take our eyes off the goal. Dr. Finney said that in her experience these “overcoming setbacks” conversations often center on allocating and protecting “dissertation time” each week. That is, the dissertation must become a priority. She noted that you must set aside time to work on the dissertation just as you set aside time to sleep, eat, work, and exercise. Students may think it is impossible to fit this dissertation time into their schedule; however, you can, she said, carve out at least 1.5 hours each day or even every-other-day to work on the dissertation.

Dr. Finney suggested finding a place where you can work uninterrupted. For her that place was not her graduate assistant office or her home; it was the 2nd floor of the library. It was that quiet place that allowed her to read and write with focus for allotted time each day. Second, she had to identify the time of the day when she was most productive and then allocate that specific time of day to working on the dissertation.

Dr. Cizek said that he began during his own graduate program to block off large chunks of time for writing—a practice that he has worked to continue throughout his career. For example, he might block off every Wednesday on his calendar for writing and, when other meetings or requests come up, he would say that he already had a commitment. Another aspect of scheduling and time monitoring is keeping track of one’s commitments. It is easy to get involved in too many projects and presentations. I have heard that from my advisor as well to be choosy in the projects that I work on. Not every project is worth the time that it will take away for other, more important or interesting work.

On the other hand, as Dr. Sireci pointed out, some work can be really good for the career. It seems to me that the ability to recognize and focus on the most lucrative work is an important skill that we can learn as graduate students and carry into our future careers.

General Advice

In reflecting on his graduate experience, Dr. Cizek said that it was passively communicated to him that this was an independent project so he had to be assertive to push the work forward. He said, “Faculty are looking for students to make the transition from student to faculty member, independent scholar or researcher.” Dr. Cizek also recommended that we be proactive and “bang on doors to make things happen.” I heard something similar from a mentor at the conference last year. Advocating for ourselves is something worth learning and practicing.

Following up on this idea, although most of the responsibility lies with us as the graduate students, Dr. Cizek also said that we need a clear advocate for the research project, and that this is a key role that an effective advisor should play. Dr. Sireci echoed that by saying that as a graduate student we are dependent on our advisor so due consideration should be taken in choosing a committee chair. I would suggest talking to other graduate students about their selection process. It is much better to have selected a chair who supports your research interests than to find that the person you chose is uninterested in your topic and does not have time to work with you on it.

With respect to interest in the topic, one piece of advice Dr. Finney shared was to select a topic that interests *you* (in addition to your advisor) and that you will enjoy working on, in depth, for a full year. It is difficult to engage for a long period of time in an arduous task, like writing a dissertation, if the task isn’t of interest to you or you don’t feel as if it’s important work. Although it is likely you will tire of the topic by the end of the dissertation, the topic should motivate and excite you to some extent in order to sustain *consistent* effort. Thus, of that long list of potential dissertation topics, ask yourself which one or two ideas excite you the most.

The dissertation process is not a project in isolation. It can help us develop skills that will carry us into our careers as I have already mentioned in terms of learning how to prioritize. Dr. Sireci said that his advisor, Kurt Geisinger, helped him learn how to write through the dissertation process. His advisor also helped him prepare for his dissertation defense. Another suggestion that Dr. Sireci had was to use the dissertation to connect with other scholars outside the university that are working on the same topic. In this sense, the dissertation can be a great networking opportunity.

Some Final Thoughts

Given, as Dr. Cizek put it, “the purpose of entering the PhD program is exiting the PhD program” we do need to make this project a priority and put time into it. However, he also recommended that we take a “hygiene day,” or in other words, take time for ourselves – go out to dinner, take a break. I have personally found that when I take break – go for a walk or something not dissertation and not work related, I am more productive when I get back. As for me, this process is teaching

me about balance – finding that point where I complete the work in a timely manner and do not burn out. That is definitely a skill worth taking into my future career.

SPOTLIGHT ON THE PEOPLE WHO MAKE OUR ORGANIZATION GREAT – LIRU ZHANG, EDUCATION ASSOCIATE (PERFORMANCE ASSESSMENT AND EVALUATION) WITH THE DELAWARE DEPARTMENT OF EDUCATION

This month our member spotlight is on Liru Zhang, Education Associate (Performance Assessment and Evaluation) with the Delaware Department of Education.



How did you get into the field?

Although my interest in measurement and quantitative analyses could be traced back to my Master's program in China, the beginning for me to systematically learn measurement theories in education was when I was a graduate student at the University of Iowa. During my doctoral dissertation, I worked full-time as a Program Associate at ACT, which provided the valuable opportunity to continue my learning in the real world. I became involved in the process of test development for MCAT and other large-scale tests. I would say the working experience at ACT helped create my firm foundation in test development and surrounding technical issues. After graduation, I started working at the Delaware Department of Education for the statewide student assessment program, a position I still hold today. My primary responsibilities are psychometrics in the design, development, and implementation of assessments. To me, the great advantage of serving at the state educational agency is the tremendous opportunities to extend my knowledge in measurement and beyond. Through this experience, I believe that I have developed a comprehensive understanding of content standards, assessment, accountability, and educational policy and their impacts on educational measurement and testing. More importantly, I have improved my knowledge and skills in measurement and psychometrics.

If you weren't in this field, what would you do?

I could be an interior designer. I love art, especially modern art. Whenever I travel to large cities, I always find time to visit museums, galleries, and shops. Through the Internet I have access to online art galleries. I have also learned key concepts and ideas from experts through professional magazines in architect and interior design. I have spent years remodeling and re-decorating my home and enjoyed the process very much.

What advice would you have for graduate students who want to get into this field?

Keep learning is my recommendation for graduate students. Learning should be a passion and a lifetime commitment for success. Conducting research, perhaps, is one of the best ways to continue learning. Through literature review, research design, data collection and analyses, and generating results and findings, you will learn a great deal. During this process, you may apply the knowledge you have learned, update your knowledge and skills, and improve your communication skills. If you get the opportunity to use real data from an assessment, I suggest you pay attention to the quality of test items. Sometimes, poorly constructed items are the sources of problems.

What do you like to do for fun outside of work?

For enjoyment, I like reading books in English and in Chinese, such as literature, history, mysteries, and true stories. I love ancient Chinese poems; they are amazing. I love movies, particularly dramas and scientific stories. And I like travel to spend quality time with my family and friends, explore different, even exotic culture and traditions.

What would you say has been one of the biggest innovations in psychometrics in the last decade or two?

In psychometrics, I think that one of the biggest changes (innovations) in the last decade or two is the prevalence of standards-based assessment for achievement testing. Some may argue that standards-based assessment is not an innovation in psychometrics, but a policy-driven product. Perhaps, this is true. The fact is that the educational reform movement over the past decades has challenged the traditional concepts and practice in educational measurement and testing, which promote the improvement of measurement theory and innovations in psychometrics. For example, the performance-based tasks used in assessments at the state, local, and national levels mirror the classroom instruction; and the promotion of accessibility for students with disabilities in large-scale assessment has considerably improved the validity of achievement tests.

When you go to conferences, how do you pick what sessions to attend?

In my review of the conference program, I usually select sessions based on (1) the topics that I am more interested in; (2) the issues that are related to my work -- student assessments, and (3) new approaches/methods that could have impact on educational measurement in the near future.

Who has been a significant influence in your professional life?

First of all I would say that I have been very fortunate to learn from many measurement experts in my professional life, including professors at the University of Iowa, the Technical Advisory Committee members for Delaware student assessments, and my colleagues. Among them, Dr. Timothy N. Ansley, Chair of Department of Psychological and Quantitative Foundations at University of Iowa, has been the most significant influence in my professional life. Dr. Ansley is an outstanding professor. He is not only an expert in educational measurement and statistics; he is kind and thoughtful as well. His advice and support encouraged me to create my professional life, for which I am always grateful.

NCME AND PUBLIC POLICY

One of the stated goals of NCME is to “*Influence public policy and practice concerning educational measurement*”. For many of us, this may seem like a daunting task and as individual members, we may wonder how this should be interpreted. Therefore, we have asked several NCME members who work in public policy contexts to provide us with their thoughts on this goal by responding to a series of related questions. We hope this collection of thoughts and ideas helps NCME members think about the influence that NCME can have on shaping and improving public policy as it relates to educational measurement.

Barbara Plake

Professor Emeritus, University of Nebraska-Lincoln



1. What types of “public policy” should NCME have influence over?

Obviously, the public policy issues should be related to the interpretation and use of test scores. This would include policies about actionable decisions related to students and educators based on test scores.

2. What areas of public policy do you see as having the greatest need for influence from expertise in test development and test use?

This is especially crucial for high stakes tests and high stakes uses of test scores. High stakes tests include those that have major implications for test takers, such as graduation and certification tests. High stakes uses of test scores include the use of student test scores for evaluation educator performance.

3. What level of influence should NCME have within these domains?

This level of influence should be educational and advisory. NCME could serve an important role in educating policy makers (including legislators) on the psychometric integrity of their proposed decisions and rules.

4. How can NCME members help advance this particular goal of NCME?

One way NCME members can help advance these policy-related goals is to promote the use of the Standards for Educational and Psychological Testing by policy makers in their states and jurisdictions.

5. Other comments?

NCME has in the past tried to serve in an active role to inform policy and policy makers, but has not been very effective. This is in part due to lack of continuity in outreach efforts. NCME doesn’t have the resources to be active in lobbying with legislators or federal government.

Ellen Forte

President, edCount, LLC



1. What types of “public policy” should NCME have influence over?

First, I need to disclose that I am such a firm believer in bringing measurement expertise together with policy involvement that I have literally made it my business to do just that. That accounts for why, should you read on, you may think that I am waving this flag really hard.

And, at the risk of appearing pedantic, I propose that there are really three questions here, one concerning the nature of the public policy that warrants NCME involvement, one the nature of NCME influence, and a third the distinction between NCME as an entity and NCME members as individuals. Over and above these questions, we need to take great care to maintain our non-partisanship in any policy position we take as an entity so we don’t lose our credibility with the next election. Policy NCME should engage with; politics NCME should not.

With regard to the nature of public policy that NCME should consider involving itself with, even a quick look at the broad range of research, evaluation, and practice areas within our organization suggests that NCME could engage in some form of influence in any policy that promotes the use of indicators for making education-related decisions about individuals, groups, programs, or entities. That encompasses, for example, policies related to large-scale academic assessments, individually administered diagnostic assessments, classroom grading, teacher evaluation, school accountability, graduation, grade promotion/retention, and teacher licensure and certification. That’s a lot of ground. Further, policies in any of these areas are multi-level; that is, policies are established at the federal, state, and local levels and always meant to guide or effect change in practice (and policy) at the next level down.

Once we see that the realm of possible involvement is very large, the nature of involvement and the distinction between organizational and individual capacities and obligations become very important. The type or intensity of influence NCME might wish to exert as an organization depends on the urgency and potential impact that involvement or non-involvement may have. Much like a triage process, NCME could identify a range of pro-active and responsive forms of influence and call upon tools from this repertoire that best fit the target. For example, NCME may wish to publish and actively disseminate position statements/papers or to liaise actively with other entities that are policy-focused (as opposed to measurement focused; examples would be the National Governors’ Association, the Council of Chief State School Officers; state and national associations of Title I Directors; local, state, and national parent-school organizations; state and national associations for special educators; local and state school boards). The former type of influence would likely be limited to highly urgent, big impact situations while the latter would be a nice long-term approach with increasing, albeit perhaps less tangible, impact over time.

Any type of formal organizational engagement will rely on the interests and capacities of at least the particular members who raise their hands or are otherwise tapped to represent NCME on policy issues. While some of us might wish NCME to engage loudly and proudly, who among us is ready, able, and willing to take on that work and what role should our employers have in supporting that in theory and in practice? We have a terrific Assessment Policy and Practice Committee, but the members of that committee already have jobs that far exceed typical working demands. Further, if amplified activity is an expectation for our leadership then we need to make that clear before we elect or nominate or volunteer.

NCME members as individuals should consider their own capacities for engaging in policy issues, as asked to by NCME leadership or on their own. Technical advisory committee membership is one excellent way to promote good measurement practices as would be participation in local school boards. Writing articles (say, in newspapers or non-professional journals) or books for more popular consumption could also spread quality measurement concepts more broadly and perhaps, thus, more effectively. Let’s face it though: just as we do not have consensus in our field about whether consequences belong under the validity evaluation umbrella (they do), we will not all share a belief in our individual obligation to get involved. Those who chose not to are not (necessarily) ninnyes.

2. What areas of public policy do you see as having the greatest need for influence from expertise in test development and test use?

Right now, the most critical areas of influence from the NCME organization and community are (1) the use of test scores and other indicators in teacher evaluations, (2) the process and criteria that the US Department of Education

will use to evaluate states' academic assessments (typically referred to as "peer review"), and (3) the evolving claims about college and career readiness based upon test scores.

Some teachers are losing their jobs when they should not while others are staying in their jobs when they should not because of extremely faulty measurement and teacher evaluation practices. NCME should take a stand on this timely, national, great import issue. Heed the joint standards.

NCME has been notably absent from the federal standards, assessment, and accountability peer review processes over the past 15 years, although many NCME members have individually been involved in individual reviews. The joint standards are not enough here; they must be translated into a more concrete framework for evidence collection and evaluation. Otherwise we are just handing people a box of really excellent watch parts and asking them what time it is.

As the state testing consortia move into their operational phases, they will lose their cohesion and individual states or groups of states will create their own versions of consortium-designed tests. What will the scores from those tests mean within and across states? How will that compare with how they will actually be interpreted and used? What implications will this have for our existing indicators of college and career readiness as well as for assessment programs that are meant to support comparisons across states and nations? Great research questions of the type that NCME ought to be raising to increase awareness and perhaps research funding.

3. What level of influence should NCME have within these domains?

In at least the three areas I indicated above, NCME as an entity ought to have a vocal, national level of engagement via at least the publication and liaison processes. NCME probably already has some degree of influence in these areas through the efforts of some of its individual members, but will never have organizational influence unless it engages selectively and purposefully and soon.

4. How can NCME members help advance this particular goal of NCME?

Uh-oh. Dealing with logistics when the budget is zero. NCME leadership, including the Assessment Policy and Practice Committee, would have to identify a small number of actionable issues and liaison entities and members would have to be willing to take on the job of writing or liaising when asked. Maybe get travel costs covered, but not be paid. Over time, it would be lovely to see more measurement programs including policy in their programs of study so the pool of eligible and qualified suspects expands beyond the handful it includes now.

5. Other comments?

I think I've said enough.

Phoebe Winter and Tia Sukin

Pacific Metrics

1. What types of "public policy" should NCME have influence over?

With the increased politicization of educational measurement, policymakers need to understand the ethical and technically sound use of educational tests. NCME should provide solid information that educates decision makers about appropriate test development and use. We recognize that NCME is a diverse organization, with members holding a variety of viewpoints on policy topics. However, there are measurement principles that NCME endorses, including those in the *Code of Professional Responsibilities in Educational Measurement*¹, the *Standards for Educational and Psychological Testing*, and the *Code of Fair Testing Practices in Education*². If we see policy taking shape that we



¹ <http://ncme.org/resource-center/code-of-professional-responsibilities-in-educational-measurement/>

² <http://ncme.org/linkservid/991A492D-1320-5CAE-6E971C543DE4CA6D/showMeta/0/>

feel should be informed by these principles, NCME has an ethical obligation to supply information to assist in its development and implementation. A recent example of such an approach is the document *Testing and Data Integrity in the Administration of Statewide Student Assessment Programs*, issued by NCME in October 2012 to assist local, state, and federal educational agencies to develop policies and procedures that safeguard the integrity of test data.

NCME influence is especially important for policy related to externally mandated tests that are used to make decisions directly affecting students, teachers, principals, and schools. Some states and districts use the results from external tests to determine whether a student should be promoted to the next grade or graduate, for instance. The United States Department of Education (USED) requires that states use assessments to determine whether schools are meeting goals and apply sanctions based on how students perform on these tests. States are also required by the federal government to use student achievement growth, as indicated by test scores, as a “significant factor” in their teacher and principal evaluation systems in exchange for flexibility in how they implement some of the mandates of the Elementary and Secondary Education Act (also known as No Child Left Behind).

2. What areas of public policy do you see as having the greatest need for influence from expertise in test development and test use?

While there may not be agreement within NCME on what process or framework to use to collect validation evidence, the organization’s foundational principles are based on validity. We agree that certain types of evidence are necessary to support certain interpretations of test scores. Central to validity is the principle that test design is driven by test purposes—that developers have determined what inferences will be made about test takers and how test results can be used. If different uses and interpretations are proposed after test development, validity evidence for those uses and interpretations must be evaluated. In sum, policymakers do not need to understand the details of test development; rather, they need to realize that (1) test development is based on a thorough consideration of inferences and purposes with respect to test scores, and (2) collecting evidence to support those inferences and purposes constitutes a critical, non-negotiable part of appropriate test use.

Perhaps the broadest inference policymakers and test users are making based on externally mandated tests is student standing on state academic content standards. (Increasingly, these include the Common Core State Standards.) Based on that inference and other data, test users make inferences about how well schools and districts are doing in implementing the standards, and how well teachers are teaching and principals are leading. These tests are typically developed, and validity evidence collected, with the first inference—student standing—in mind. The uses of test scores as part of other consequential evaluation systems are rarely, if ever, considered when the tests are designed, so validity evidence with regard to these additional uses is not collected.

Right now, policymakers and state agencies need recommendations for how to collect validity evidence for high-stakes teacher and principal evaluation systems that incorporate measures of student achievement. While the use of test scores in this way has high common-sense appeal, there is not much evidence that systems incorporating student test scores provide a basis for valid inferences about teacher and principal effectiveness. On the contrary, a number of studies show that using test scores alone does not provide a basis for appropriately defining levels of teacher quality (for example, see Lockwood et al., 2007; McCaffrey et al., 2009; Newton et al.; Rothstein, 2010; Schochet & Chiang, 2010). Given this evidence, NCME should make recommendations for studying how (or whether) incorporating test scores into an evaluation system that includes other data, as required by the USED, supports the targeted inferences and decisions about quality.

3. What level of influence should NCME have within these domains?

Although we cannot force policymakers to abide by the *Standards for Educational and Psychological Testing* or the *Code of Fair Testing Practices in Education*, we can make policymakers aware of how these principles apply to policy decisions. NCME needs to do all it can to make sure that important, life-changing policy decisions are made with an understanding of the relevant professional standards and how these standards affect the scope of appropriate decisions. We may not agree with the final policy decisions, but we will have fulfilled one of our responsibilities as a professional organization by making sure policymakers had the information they needed to make appropriate decisions.

NCME recommendations to policymakers must be based on and refer to the *Standards* and the *Code of Fair Testing Practices* and be limited to issues for which the principles in these documents have clear applicability. The organization’s approach should not be one of supporting or disavowing a policy for test use (unless that policy is

clearly unethical or immoral); rather, NCME should explain what is needed in test development procedures and validity investigations to evaluate whether the proposed uses are supported.

How can NCME do this? NCME can develop and distribute accessible issues briefs to the USED, Congress, state departments of education, state legislatures, and non-governmental organizations (NGOs) concerned with test use in education. We can work with AERA and APA to coordinate educational lobbying efforts on issues clearly affected by our organization's principles. We can publicize our activities on our website and encourage NCME members to have their organizations link to our site. Most importantly, we can educate our members. NCME's *Code of Professional Responsibilities in Educational Measurement* has clear implications for how members can and should respond to situations in which policy has implications for test use.

4. How can NCME members help advance this particular goal of NCME?

NCME members work in the USED, state departments of education, testing companies that develop externally mandated tests, NGOs, and perhaps even on congressional and legislative staffs. We serve on state and national technical advisory committees, and we are called on for advice from state and federal policymakers. In these roles, we can take advantage of materials that NCME has produced and, especially, give advice that explicitly references the *Standards*. In this way we raise the profile of the *Standards* and provide notice that educational measurement is guided by important principles, with practices that govern test development and use. These principles and practices can be applied and implemented in a variety of ways, but they should underlie test development and test use.

5. Other comments?

The *Code of Professional Responsibilities in Educational Measurement* provides very clear guidelines for responsible practice in educational measurement. Several of these guidelines apply directly to responses to policy for test development and use. Examples of guidelines related to this responsibility include the following:

General Responsibilities

. . . NCME members involved in educational assessment will . . . adhere to the highest standards of conduct and promote professionally responsible conduct within educational institutions and agencies that provide educational services;

Section 1: Responsibilities of Those Who Develop Assessment Products and Services

1.8 Caution users, in clear and prominent language, against the most likely misinterpretations and misuses of data that arise out of the assessment development process.

Section 2: Responsibilities of Those Who Market and Sell Assessment Products and Services

2.9 Inform potential users of known inappropriate uses of assessment products and services and provide recommendations about how to avoid such misuses.

Section 3: Responsibilities of Those Who Select Assessment Products and Services

3.4 Inform decision makers and prospective users of the appropriateness of the assessment for the intended uses, likely consequences of use, protection of examinee rights, relative costs, materials and services needed to conduct or use the assessment, and known limitations of the assessment, including potential misuses and misinterpretations of assessment information.

Section 6: Responsibilities of Those Who Interpret, Use, and Communicate Assessment Results

6.2 Provide to those who receive assessment results information about the assessment, its purposes, its limitations, and its uses necessary for the proper interpretation of the results.

As we educate others on appropriate practices in test development and the validation of interpretations and uses of test data, we should be sure that we are not perceived simply as naysayers. We should help policymakers and those who implement policy determine what information and procedures are necessary in order for the policy and its implementation to provide sound information.

References

Lockwood, J. R., McCaffrey, D. F., Hamilton, L. S., Stecher, B., Le, V., & Martinez, J. F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement*, 44(1), 47–67.

- McCaffrey, D. F., Sass, T. R., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy*, 4(4), 572–606.
- Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives*, 18(23). Retrieved from <http://epaa.asu.edu/ojs/article/view/810>
- Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, 125(1), 175–214.
- Schochet, P. Z., & Hanley, C. S. (2010). *Error rates in measuring teacher and school performance based on student test score gains* (NCEE 2010–4004). Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/pubs/20104004/pdf/20104004.pdf>

NCME WEBSITE TOOLS: GLOSSARY OF IMPORTANT ASSESSMENT AND MEASUREMENT TERMS

April L. Zenisky, NCME Website Board Member

From *ability* to *variance*, and plenty in between, the “Glossary of Important Assessment and Measurement Terms” located on the NCME Website is a resource that is intended to provide an accessible dictionary of assessment-related terms for use by a non-technical, lay audience. It offers definitions for a number of technical or specialized terms in a way that is purposefully straightforward, to maximize access for persons such as K-12 educators, parents of students, media writers, adult test takers, researchers who use assessment instruments or existing assessment results in their work, policy makers at the state and national levels, and higher education administrators in areas such as admissions or institutional research.



The glossary can be found on the NCME Website directly through this link (<http://ncme.org/resource-center/glossary/>), or by clicking on “Glossary” under the Resource Center link on the NCME homepage. It is searchable by word or can be browsed by letter. As users type a search word or term, a list of matches in the glossary appears.

There are approximately 150 entries in the glossary, and the terms included were selected to reflect a broad range of concepts and labels across the realms of achievement, aptitude, and psychological assessment. The glossary was initially developed by Glossary Editor David Frisbie in collaboration with an advisory board, who together sought to cut across testing contexts and perspectives to identify terms meeting the broad intended scope of the glossary tool. All definitions for glossary terms were developed by the Glossary Editor and were fully reviewed by the advisory board.

To NCME Members: if you think there are terms that could be added or definitions to be modified, such suggestions can be made to the Website Content Editor through a web form (<http://ncme.org/resource-center/glossary/suggestions/>).

UPDATE ON THE 2014 ANNUAL MEETING

Paul De Boeck, Ohio State University & Kathleen Scalise, University of Oregon

The ten expert panels evaluating the submissions for next year’s meeting have finished their work. Although the number of submissions has increased by 50%, 50% to 70% of the paper submissions (including coordinated sessions proposals) can be accepted for paper presentations. The exact percentage will depend on the available space.



Six invited speakers are scheduled for next year's meeting. This list includes (bio information taken from speaker's websites):



Ryan Shaun Joazeiro de Baker
Teachers College, Columbia University

About Ryan: Research interests include Educational Data Mining, Learning Analytics, Learner-Computer Interaction, Gaming the System, Student Modeling (Disengagement, Affect, Robust Learning, Knowledge, Strategic Behavior, Meta-Cognition), Intelligent Tutoring Systems, Serious Games/Educational Games, Science Microworlds, Off-Task Behavior, Cultural Factors. (Read more: <http://www.tc.columbia.edu/academics/index.htm?facid=rsb2162>)



Angela Duckworth
Psychology Department, University of Pennsylvania

About Angela: “She studies competencies other than general intelligence that predict academic and professional achievement. Her research centers on self-control (the ability to regulate emotions, thoughts, and feelings in the service of valued goals) and grit (perseverance and sustained interest in long-term goals). She is particularly interested in the subjective experience of exerting self-control and grit – and conscious strategies which facilitate adaptive behavior in the face of temptation, frustration, and distraction.” (Read more: <http://psychology.sas.upenn.edu/people/duckwort>)



Eric Hanushek
Hoover Institution, Stanford University

About Eric: “He has been a leader in the development of economic analysis of educational issues with his work frequently entering into the design of national and international educational policy. His research spans such diverse areas as the impact of teacher quality, high stakes accountability, equity and efficiency in school finance, and class size reduction along with the role of cognitive skills in international growth and development.” (Read more: <http://hanushek.stanford.edu/>)



Donald Hedeker
School of Public Health, University of Illinois at Chicago

About Donald: “His main expertise is in the development and use of advanced statistical methods for clustered and longitudinal data, with particular emphasis on mixed-effects models.” (Read more: <http://tigger.uic.edu/~hedeker/>)



Jim Popham
Emeritus Professor, UCLA

About Jim: “He has spent the bulk of his educational career as a teacher. His first teaching assignment, for example, was in a small eastern Oregon high school where he taught English and social studies. In 1978, he was elected to the presidency of the American Educational Research Association (AERA). In 1968, he established IOX Assessment Associates, an R&D group that formerly created statewide student achievement tests for a dozen states. In 2002 the National Council on Measurement in Education presented him with its Award for Career Contributions to Educational Measurement.” (Read more: <http://www.ioxassessment.com/who.php>)



Valerie Shute
College of Education, Florida State University

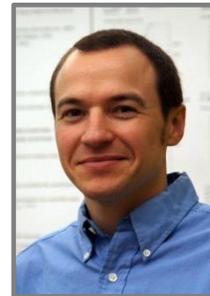
About Valerie: “Her general research interests hover around the design, development, and evaluation of advanced systems to support competencies. She and her colleagues are developing a suite of model-based tools that are used to assess understanding and provide the basis for informative and reflective feedback during instruction.” (Read more: <http://myweb.fsu.edu/vshute/index.html>)

NCME FITNESS WALK/RUN CORNER

Brian French, Washington State University & Jill van den Heuvel, Alpine Testing Solutions

We hope your summer has been outstanding! We know all of you are looking forward to the beautiful fall walking and running weather. We are in the planning stages for the event in Philadelphia. In doing so, we would appreciate your response to a brief survey which will assist us in planning for the coming year. It will take less than 7.5 minutes of your time. We will select 2 responders randomly to receive free registration (with shirt) this year. Keep moving!

Survey link: <https://www.surveymonkey.com/s/NCMEWalkRun2013>



ANNOUNCEMENTS: NCME

MEMBERS ON THE MOVE

Andrew Wiley -> Alpine Testing Solutions

Thanos Patelis -> The National Center for the Improvement of Educational Assessment

OTHER CONFERENCES OF INTEREST

- Midwestern Educational Research Association: Annual Meeting, November 6-9, 2013, Evanston IL
- Mid-South Educational Research Association: Annual Meeting, November 6-8, 2013, Pensacola Beach FL
- New England Educational Research Organization: Annual Meeting, April 30 – May 2, 2014, Dover VT

**To get the NCME Newsletter four times a year
(March, June, September, and December) go to
<http://ncme.org/publications/newsletter/>**

Newsletter Advisory Board

LANIE BRADSHAW, University of Georgia
JOHN DENBLEYKER, Minnesota Department of Education
ELLEN FORTE, edCount
JOANNA GORIN, Arizona State University
SARA HENNINGS, MetaMetrics, Inc.
JOAN HERMAN, CRESST/UCLA
ANDREW HO, Harvard Graduate School of Education

THEL KOCHER, Walden University
GERRY MELICAN, The College Board
MELINDA MONTGOMERY, University of Kansas (Grad Student Rep)
THANOS PATELIS, The Center for Assessment
S.E. PHILLIPS, Consultant
CHRISTINA SCHNEIDER, CTB
XIAN (BO) WANG, The College Board

SUSAN DAVIS-BECKER, Editor, Alpine Testing Solutions
Send articles or information for this newsletter to:

Susan Davis-Becker
Alpine Testing Solutions

Phone: 402-483-5898
e-mail: susan.davisbecker@alpinetesting.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.