

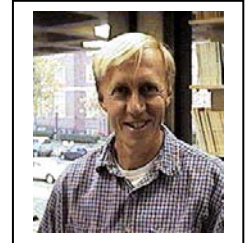


FROM THE PRESIDENT: UPDATE ON NCME ACTIVITIES!

Terry A. Ackerman, University of North Carolina - Greensboro

NCME members,

In my second letter I'd like to update you on several NCME Board activities and also share with you results of a survey that members filled out when they renewed their membership last fall.



First, however, let me begin by congratulating Jimmy de la Torre. This past summer Jimmy was notified that he had won the Presidential Early Career Awards for Scientists and Engineers. This is a very prestigious award and one that is rarely awarded to researchers in educational measurement. I encourage you to congratulate Jimmy. The link to the press release is http://www.whitehouse.gov/the_press_office/President-Honors-Outstanding-Early-Career-Scientists/. It should be noted that at last year's annual NCME meeting in San Diego Jimmy was awarded the Jason Millman Promising Scholar Award. Thus, the presidential award not only cross validates Jimmy receiving the Millman Award but also speaks to the impact Jimmy has had on the national stage.

Over the past several months I know many NCME Committees have been meeting and involved in a lot of activities concerning outreach, membership, awards, standards, and website. In preparation for our 2010 meeting, Luz Bay has been busy putting together a series of workshops and John Willse and Bob Henson are busy sorting through reviews on the 425 proposals that were received. I can't tell you how valuable your work is and how appreciative I am of all your efforts. The NCME Board has been very busy as well. We will be meeting at the end of October in Washington, D.C., at the AERA offices. Plumer, Wayne and I have gone to D.C. met twice to visit with Felice Levine (AERA's Executive Director), Jerry Sroufe (AERA's Assistant Director) and Laurie Cipriano (AERA's meeting planner). I arranged these meetings to strengthen our relationship with AERA. Our contract with AERA expires in 2010 and there are several changes we are considering that will improve the NCME annual meeting and make the meeting planning process more timely and effective. We also again will partner with AERA's Division D at our Denver meeting for the Nohost reception. We are considering having a couple of bands composed of NCME and AERA members at the reception. It should be fun.

Another initiative that I have been working on is to become more engaged with the policy makers in D.C. and have NCME serve as a resource to these people. We have invited the following people come and talk with NCME Board members in October about this process: John Tanner (CCSSO), John Easton, U.S. Department of Education and Director of the Institute for Educational Studies (IES), and Becky Smerdon, Vice President and Deputy Director, U.S. Education and Workforce Development. We are really excited about this opportunity and are looking forward to these discussions. Hopefully, in my next letter I'll be able to lay out a more definitive plan and what it means logistically for NCME to be involved in the educational assessment policy decision making process.

Finally I would like to share with you some of the results of the NCME survey that members filled out when they renewed their membership. In all 88 members responded. Kelly Godfrey from the College Board helped me analyze and summarize the survey results.

1. Collaboration with AERA Division D (82 comments)

About half of the responders to this question praised the collaboration with Division D. The other half offered suggestions or criticisms, ranging from not even noticing that there was collaboration to asking for better scheduling.

2. The annual meeting (77 comments)

These were broken down and categorized into eight general categories:

- a) Timing for presentations and discussion (7 comments).* Most of the comments were requests for more time per presentation, by having either fewer presentations or longer sessions.
- b) Session topics (14 comments).* Eight comments here were regarding the diversity of topics presented at the conference, 5 comments were requests for more practical applications, and 1 was a request for more theoretical presentations. Those who commented on diversity of topics noted that many presentations focused on the standard topics and so did not have the depth they wanted or they indicated that the topics were not applicable to their work.
- c) Quality of speakers and presentations (23 comments).* There were five sub-categories in this category: technology issues (3 comments regarding PowerPoint compatibility issues and/or microphones), poor presentation skills (9 comments), language barriers (3 comments), and quality of papers presented and/or inferences made (9 comments).
- d) Proposal review process (3 comments).* There were 3 comments regarding the proposal review process, including one person who was upset that proposal and his friends' proposals, based on dissertations, were rejected. The other two comments concerned requiring more practical applications from the proposals.
- e) Program organization (10 comments).* These 10 comments tended to be about session conflicts (difficulty attending sessions because of concurrent time slots), suggestions to hold NCME at a different time than AERA (this comment occurred in responses to more than one of the five questions), complaints about not breaking for lunch, and a request for an earlier program release so that travel plans and workshop registrations could occur sooner.
- f) Materials (5 comments).* These comments tended to focus on access to presenters' papers and handouts, either suggesting that we require them to be available at the conference, or make them available online.
- h) Facilities (26 comments).* In regards to the Hard Rock Hotel (13 comments), comments tended to be about helpfulness of the staff, space and layout of the hotel, high noise levels, poor lighting, and lack of access to coffee and refreshments (this was also a recurring theme in responses). The 14 audiovisual comments tended to praise the availability of projectors and A/V staff on hand, but some complained that having only one microphone available was insufficient and made it difficult to hear.

3. New topics (68 comments)

These comments fell into 8 general categories ranging from types of assessment (including formative assessment, classroom assessment, CAT, and alternate forms of assessment), to specific topics (such as validity, theory, state assessments, programming and software, ELL, and college readiness), to the assessment process (including test security, standards setting, scoring, reporting to the public/lay-people, quality control, item and test development, equating and scaling).

4. Overall Suggestions (95 comments)

There were two major categories here:

- a) Hotel (these comments included, but were not limited to, wireless internet access in conference rooms, noise and lighting, location and floor, and access to refreshments)
- b) Conference (these comments were put into 3 categories)
 - a. Program (including comments about the NCME website for the conference, scheduling, projectors, and presentations/discussions)
 - b. Praise for the conference (including specific staff members and general praise)
 - d. Activities (including requests to continue the Fun Run, comments on the reception(s), and suggestions to collaborate the career center with AERA's)

I have shared these comments with this year's Program Chairs, Bob Henson and John Willse, our Workshop coordinator, Luz Bay, and our Executive Director, Plumer Lovelace, and his staff. To date we have 1,795 members so the survey responses represent only 4.9% of our membership. I strongly encourage more of you to share your thoughts and ideas with us. We are considering establishing a link on the website to allow members to share their ideas and suggestions at any time.

We will be coming out with a new survey this fall that will be extremely important. We plan to use the results of the survey to create two maps of the United States. One map would be for graduate students. They will be able to go to this map and click on any state and to see a listing of the measurement programs in that state. Links would then be established to the particular measurement program's website. The second map would be for measurement practitioners (school accountability personnel, school district administrators, state departments of education, etc.). When practitioners click on a state on this map, a list of NCME members (who volunteered their information and agreed that the information could be posted) would be listed along with their particular areas of expertise and their contact information. The goal is to get our membership more involved in serving local and state educational communities in the areas of assessment and educational research. We will be sending more information about this in the future.

Thanks,



NEW COLUMN... A NOTE FROM THE EDITOR

Thanos Patelis, *The College Board*

The active NCME Newsletter Advisory Board has offered some wonderful ideas that we hope you enjoy. First, we have the presidential column by our president, Terry Ackerman (University of North Carolina – Greensboro). Second, we're introducing a new column entitled Legal Corner by S.E. Phillips. Each issue we'll tackle a legal or policy issue and request commentary from members. In this issue, the Sotomayor ruling on the *Ricci v. DiStefano* case is discussed with commentary from Dave Arnold, Chad Buckendahl, and Wayne Camara, followed by concluding comments by S. E. Phillips. Third, we have our standing graduate student column by Carol Barry where the topic of taking time to reflect is discussed. Then, we have our spot light on one of our esteemed members, Jim Impara, provided and nicely written by Susan Davis. Finally, please find a variety of Calls and miscellaneous announcements. Please drop me an email with suggestions. Sincerely and at your service, Thanos.



LEGAL CORNER: JUSTICE SOTOMAYOR AND EMPLOYMENT TESTING

S.E. Phillips, *Consultant*

On August 6, 2009, the Senate voted 68 to 31 to confirm Second Circuit Court of Appeals Judge Sonia Sotomayor to replace Justice David Souter as the 111th justice of the U.S. Supreme Court. Her confirmation hearings explored her unique background and, among other cases, her participation in a Second Circuit employment testing decision (*Ricci v. DeStefano*) recently reversed by the Supreme Court.

Background

In addition to being the first Hispanic and third female associate justice, Justice Sotomayor brings multiple perspectives and varied experiences to the Supreme Court. Her Puerto Rican parents immigrated to New York City during World War II and raised her and her younger brother (now a doctor) in the housing projects of the South Bronx. Her father, who died when she was nine years old, was a factory worker with a third grade education who spoke no English. She attended Catholic schools and was diagnosed with juvenile diabetes. She received a scholarship to attend Princeton University, spent the summer between her freshman and sophomore years learning to write well in English and was awarded her undergraduate degree summa cum laude. She also joined other Hispanic students to urge the school to hire more Hispanic professors. She graduated from Yale Law School where she was editor of the Law Review. She also filed a complaint against a recruiter from a Washington law firm who asked a question implying that minorities admitted under Yale's affirmative action program might not have been fully qualified. A student/faculty review panel ordered the firm to write a letter of apology.

Justice Sotomayor began her legal career as an assistant prosecutor in the Manhattan District Attorney's Office. Initially, she prosecuted misdemeanor crimes which she said were difficult because she believed them to be mainly caused by poverty. Six months later she was reassigned to felonies and distinguished herself in several high profile cases. After five years in the DA's office and a divorce, she moved to a private law firm and spent eight years representing corporations in international business law cases. She also spent twelve years on the board of the Puerto Rican Legal Defense and Education Fund where she served on its litigation committee. Among other cases, the organization challenged New York City's police and fire department hiring and promotion testing procedures.

Justice Sotomayor also has 17 years of experience as a federal judge. In 1991, at age 38, President H.W. Bush nominated her for the federal district court. Five years later, President Bill Clinton nominated her for the Second Circuit Court of Appeals. When he nominated her for the Supreme Court, President Obama praised her experiences and her empathy. She had been rated highly as a federal judge and variously characterized as "tough," "no-nonsense," "talkative" and "domineering." One male judicial colleague declared that the latter descriptions were sexist and that she was no more assertive than other federal judges.

As a federal judge, Justice Sotomayor ruled on about 100 cases involving racial discrimination claims. One analyst concluded that she favored the discrimination claim in less than 15% of these cases and was reversed only in the New Haven case.

New Haven Case

In 2003, the city of New Haven, Connecticut administered civil service examinations to 118 firefighters seeking promotions to the rank of Lieutenant or Captain. The exam results determined rank order eligibility for promotions to fill an existing 8

Lieutenant and 7 Captain vacancies and for any additional vacancies during the following two years. The city charter required job-related examinations and the selection of a candidate for a vacancy from the top three scoring candidates on a rank-ordered list certified by the New Haven Civil Service Board (CSB). The contract between the city and the firefighters union also specified a composite total score for each candidate weighted 60% written test and 40% oral exam. To qualify for testing, candidates were required to have a high school diploma and specified vocational training together with 30 months of experience for Lieutenant positions and one year of experience as a Lieutenant for Captain positions.

The city contracted with a testing company specializing in police and firefighter entry-level and promotional exams. The test development process included a job analysis involving interviews of job incumbents, ride-along observations and a job analysis questionnaire administered to most incumbents and their immediate supervisors in the department. Minority firefighters were oversampled at all stages of the process. Content rated “critical” or “essential” was emphasized. A list of study materials from which test questions were drawn was compiled and distributed to all test takers three months before test administration. The written test consisted of 100 multiple-choice questions with a reading level below tenth grade and the oral exam involved responses to hypothetical job situations rated by a panel of three evaluators. Each panel consisted of one African-American, one Hispanic and one White member trained to apply checklists of desired response criteria. Test results were as follows:

Test	N _{total}	N _{minority}	AA _{pass}	H _{pass}	W _{pass}	[80% W _{pass}]
Lieutenant	77	34	32%	20%	58%	[47%]
Captain	41	16	38%	38%	64%	[51%]

Under the EEOC *Uniform Guidelines*, these data created a presumption of disparate impact because the passing rates for both minority groups were less than 80% of the majority passing rate for both tests. Prior exams administered in 1999 had demonstrated similar levels of disparate impact. Based on the “rule of three,” all ten of the candidates eligible for immediate promotion to Lieutenant were White. Of the nine candidates eligible for immediate promotion to Captain, two were Hispanic and seven were White. Rather than requesting the technical manual required by the contract, City officials met privately with the testing company’s vice president. Apparently, no differential performance statistics were reported, probably due to the small numbers of test takers. At best, a comparison of White and combined minority item performance may have been possible for the Lieutenant exam.

The CSB held a series of five hearings to consider whether to certify these results. Multiple stakeholders, the testing contractor, three experts and several test takers made presentations at these meetings which were often heated and contentious. One test taker with learning disabilities, including dyslexia, had incurred substantial financial cost to obtain tapes of the study materials and reported studying 8-12 hours daily to prepare for the test. Although he did not yet know whether he had passed, he (and others) argued that the tested material was drawn from Department procedures and accepted firefighting standards (i.e., the test was job-related and content valid). He also stated that “the people who passed should be promoted. When your life’s on the line, second best may not be good enough.” However, some other test takers believed that the test questions were outdated or irrelevant to conditions in New Haven and that the study materials were too lengthy and expensive at about \$500 for a full set.

Of the selected experts who provided opinions, only one had firefighting experience and had examined the tests in detail. He was a retired, African-American fire captain from Michigan currently working as a fire program specialist for the Department of Homeland Security. He supported the relevance and appropriateness (job-relatedness and content validity) of the test questions and stated that the test takers were given an advantage by being told in advance which book chapters from the study materials covered the tested content. The only testing expert was a competitor of the testing contractor who had not thoroughly reviewed the tests or seen the job analysis data. He indicated that the tests’ disparate impact was similar to that for other written tests and, although he recommended certifying the results, he criticized the lack of review of test questions by Department personnel (outside firefighters had been used to protect test security due to allegations of cheating on prior exams) and the higher weight given to the written portion of the test. He recommended replacing the test with assessment centers simulating real-world situations (a service his company was prepared to provide to the city). No data on relative costs, validity or reliability of the assessment centers were provided. The final expert was a professor from Boston College who declined an offer to review the exams. She concluded that the exams were unfair because the job analysis reflected a White job perspective that differed from the way other ethnic groups might perform the job. There was some evidence that City officials had exerted influence on the decision process.

Following a recommendation from the City attorney not to certify the test results, the CSB, with one member recused, split 2-2 resulting in a failure to certify the results. Seventeen White firefighters and one Hispanic firefighter who passed their tests but were denied an opportunity for promotion filed suit against the City and its officials. The City argued that the result was fair because all firefighters were treated the same by promoting no one. The firefighters argued that the decision to invalidate the exam results was reverse racial discrimination in violation of Title VII. Specifically, they alleged disparate treatment based on race. The City countered that the disparate impact statistics and threat of a resulting lawsuit justified their decision. The federal district court granted summary judgment for the City and a three-judge panel from the Second Circuit, which included Justice Sotomayor, affirmed on appeal in a short, one-page opinion stating that the City’s actions were protected because it “was

simply trying to fulfill its obligations under Title VII when confronted with test results that had a disproportionate racial impact.” A motion for rehearing of the case by the full appeals court was denied. In a 5/4 decision, the Supreme Court reversed, holding that disparate impact statistics alone were insufficient to justify disparate treatment based on racial considerations (invalidating the test results because the higher scoring candidates were White).

Recall that Title VII disparate impact challenges involve a shifting of burdens. Once the plaintiffs establish a presumption of disparate impact via statistics failing to satisfy the 80% rule, the burden shifts to the employer to demonstrate “business necessity” – the validity and job relatedness of the test. If the employer’s argument is convincing, the burden shifts back to the plaintiffs to demonstrate the existence of an equally effective but less discriminatory test that the employer refused to implement. In the New Haven case, the Supreme Court held that based on the disparate impact statistics alone, it was not a foregone conclusion that the City would be found in violation of Title VII. By failing to obtain a technical manual and disregarding the other evidence of validity and job-relatedness of the test in the record, the City had failed to adequately evaluate its likelihood of prevailing. Similarly, the Court found that one expert’s mentioning of assessment centers and the possibility of less disparate impact with different written/oral exam weights (an action that might have run afoul of the Title VII prohibition on adjusting employment test scores based on race) did not by itself constitute sufficient evidence of an equally valid, less discriminatory alternative. The Supreme Court held that to justify disparate treatment based on race, the potential threat of disparate impact liability must be based on “a strong-basis-in-evidence,” a standard that the Court found was not satisfied in the New Haven case.

Based on its holding for the firefighters on the Title VII claim, the Supreme Court did not reach the constitutional question of whether the disparate impact provisions of Title VII are consistent with the equal protection clause, which prohibits only intentional discrimination. In a concurring opinion, Justice Scalia stated that this issue will be confronted by the Court in the future, perhaps during Justice Sotomayor’s tenure.

Implications

In two other notable cases involving claims of racial discrimination, Justice Sotomayor sided once with the employee and once with the corporation being challenged. She dissented in a case in which the court upheld the firing of a New York City police officer for sending hate mail on his own time, reasoning that under the U.S. Constitution; even a White bigot had the right to speak his mind. When American Airlines was sued for bumping a black couple from an international flight, she joined the court majority in holding for the airline whose actions were governed by clear language in an international treaty. According to the court, this policy could only be reversed by an act of Congress or a federal agency. The views of Justice Sotomayor in the three racial discrimination cases discussed above suggest a general tendency to uphold the letter of existing law and precedent.

Justice Souter, whom Justice Sotomayor is replacing, joined the dissent in the New Haven case. Given her support of the reversed appeals court decision in that case, the result would likely have been the same had she been on the Court (assuming she had not already participated in the lower court ruling). Thus, if an employment testing case with similar facts alleging reverse discrimination should come before the Court in the future, her participation on the Court in place of Justice Souter is not likely to significantly change the balance of viewpoints on the Court. However, in the absence of a definitive prior record on other testing issues (e.g., accommodations), her potential future effect on assessment law generally is as yet unclear.

LEGAL CORNER: COMMENTARY

David W. Arnold, Wonderlic, Inc.

S.E. Phillips’ article regarding Justice Sotomayer and the *Ricci* decision is one of the most objective and accurate representations of fact that I have read. Moreover, from my perspective, I view the extensive attention given to and concern expressed regarding Sotomayer’s appointment, and the Supreme Court’s opinion in *Ricci* as being pretty much a tempest in a teapot. Regardless of whom President Obama would have identified for appointment to the Supreme Court that individual’s rulings would have tended to consistently fall in line with the philosophy of the current administration. This is to be expected regardless of whether the then-current administration is liberal or conservative. As pointed out by Phillips, the replacement of Justice Souter by the present administration does not impact the Court’s balance of power.



As for the *Ricci* opinion, the media has provided us with a plethora of conjecture regarding its impact. Speculation has ranged from the decision being extremely unfavorable to employers to it being very positive.

In reality, the *Ricci* decision has little or no impact on the vast majority of employers (especially those in the private sector) and their employment practices. It doesn’t raise the standard for justifying the use of hiring/promotional practices (or testing) even for those practices that tend to exhibit disparate impact (e.g., background checks, credit checks, physical requirements, drug

tests, certain written tests, educational requirements). With respect to testing per se, the Court's comment exclusive thereto was: "Employment tests can be an important part of a neutral selection system that safeguards against the very racial animosities Title VII was intended to prevent."

Bottom line, the *Ricci* opinion simply mandates that employers cannot process a pool of candidates and then discard the results (intentionally discriminate against a protected subgroup) simply due to the existence of disparate impact. In order to justify such an approach, the employer must have a strong basis in evidence that the employment practice was flawed because it was not job related OR that a less discriminatory practice with equal validity was available.

LEGAL CORNER: COMMENTARY

Chad W. Buckendahl, Alpine Testing Solutions

I appreciate the opportunity to respond to Phillips's discussion of Justice Sotomayor's background and more important, key issues involving the recent Supreme Court decision in *Ricci v. DeStefano*. Although the potential impact of this decision on the testing community has been downplayed by some commentators, there are some lessons and reminders for the measurement community in light of the Court's holding.



Defining the practice domain

For employment and credentialing exam programs, a job analysis serves as a primary source of validity evidence because it specifies the job-related content, cognitive demand, performance demand, and environment. Subsequent steps in test development and validation are built on this foundational information. In the *Ricci* case, some concerns were raised by that the population of incumbents may have been too small or unrepresentative of the intended population of candidates to confidently define the domain. However, New Haven apparently did not systematically evaluate whether the validity evidence collected by the test developer on this or other components addressed this challenge was valid.

When similar concerns arise in the future, test developers should consider alternative study designs that collect and evaluate additional validity evidence to better support the argument that scores can be used and interpreted as intended. For example, because there are presumably other fire departments in Connecticut, the Northeast, or other municipalities nationally of similar size and demographic composition, additional sources of domain representation may have been available to cross-validate. For testing programs that may not have opportunities for collecting representative or comparative data, there is a greater burden to justify the resultant content specifications.

Standard setting

A second measurement consideration relates to the decision points that are used with these testing programs. Within the employment and credentialing examination program industries, defining the performance standard and setting cut scores is often driven primarily by legal and policy considerations. The 80% rule to evaluate potential disparate impact that Phillips discussed is often used within these types of testing programs as a starting point to establish, rather than more appropriately evaluating cut score recommendations.

Although these requirements are based on the EEOC's Uniform Guidelines, the rule perpetuates a common, but intuitive misconception that differences in subgroup performance without conditioning for ability suggest prima facie evidence of bias or unfair treatment. As a result, adherence to this and other legal expectations (e.g., legislatively defined cut scores) unintentionally encourages poor measurement practice in developing and evaluating cut scores because program sponsors may not want to gather the evidence necessary to defend a disparate impact challenge. The *Ricci* case provides such an illustration. Although fixed in a generation of practice, the expectations of the Uniform Guidelines could also be revisited through future legislation, regulation, or caselaw. Revisions to current requirements may become necessary as the demographics of national populations continue to change. These shifts in composition will contribute to increasing challenges in classifying individuals into discrete subgroups. Extending the concept, these changes would have related implications for other areas in measurement such as sampling, differential item functioning, and disaggregated score reporting.

Concluding thoughts

After following the *Ricci* case, I enjoyed reading the Court's holding and some of the supplemental briefs and documents to try to understand the evidence the Court prioritized in their deliberations. As discussed above, there were measurement issues that this case highlighted that inform practice. However, one lingering question for me was whether the plaintiffs had sufficient legal grounds for the action. Specifically, whether the City's ultimate non-decision regarding promotion (i.e., deciding that no

one would be promoted) represented sufficient harm to the plaintiffs for the benefit of the defendants or the firefighters who were not recommended for promotion.

In reverse discrimination cases, an argument is often made that a member of a protected class has unfairly received some benefit (e.g., admission into a professional training program, employment) that was subsequently denied to a more qualified member of an unprotected class. In this instance, the implication is also that potentially withholding a position or promotion for qualified applicants without validity evidence to support an intentional inaction violates individual rights.

In summary, this case serves as a reminder that the measurement community should continue to promote and implement its professional standards, but to also understand when legal issues need to be considered in the test development and validation process.

LEGAL CORNER: COMMENTARY

Wayne J. Camara, *The College Board*

The facts of the case (*Ricci* decision) are quite unique in employment selection and would be difficult to envision in educational or licensing testing contexts. The “rule of three” established by the city requiring hiring one of the three highest ranking candidates. This is a modified top down hiring procedure that only considers the three highest scoring candidates at one time for a position. Once a candidate is selected then this sliding scale allows the next highest ranked candidate to qualify within the top three candidates. There were 8 vacancies for Lieutenant and 7 vacancies for Captain so there were relatively few selection decisions. Rather than address the legal issues I would like to discuss one fact of the case – the use of test scores as the sole basis for promotion – since this issue does generalize to other settings.



Several screening factors were used to qualify candidates for testing (e.g., HS diploma, experience), but test scores were the sole basis for hiring decisions among those who took the promotional examination. This is not unusual in civil service and other employment (and licensing/certification examinations). Did New Haven rely solely on test scores for promotion and is this consistent with the Testing Standards? Because other factors were required to qualify for testing I do not believe the test was the sole determinant of promotion, yet it was the only determinant for incumbents who met the entry hurdles. In education there is an increased uneasiness about basing individual decisions solely on assessment results (e.g., college admissions, scholarships) and holistic review is increasingly popular in making such decisions. In employment or credentialing contexts there is much less acceptance for subjective factors or human reviews. The difference in attitudes about objective vs subjective factors in these settings may largely rest with expectations and trust. Large and persistent gaps on cognitive ability tests have led most educators to believe that subjective factors will likely benefit underrepresented students while organizations may be less benevolent in considering an incumbent’s background and circumstances. The public may be less trusting that organizations would weight individual factors appropriately and there has been historical comfort in objective data to evaluate adverse impact or discrimination. In fact, research has suggested that there is a relative strong inclination for individuals to hire and promote individuals with similar backgrounds as themselves (e.g., education, ethnicity, values and experiences). Promotion and selection in police, fire, and other similar government sectors has been a controversial and litigious area for decades. Basing hiring decisions on an objective test is defensible, but only when the test is job related and measures the appropriate knowledge, skills and abilities required. In *Ricci*, the city decided it did not like the outcome of the selection system it had approved and did not take the time to evaluate the evidence or process. If assessments are appropriately developed and validated then invalidating results at the end of the day is bad practice and leads to the cynicism we see in the media about measurement and assessment practices. For reasons other than the legal issues Susan Phillips and other commentators have noted, the Supreme Court got this one right.

LEGAL CORNER: JUSTICE SOTOMAYOR AND EMPLOYMENT TESTING

S.E. Phillips, *Consultant*

Concluding Comments

Thanks to the ABC commentators (Arnold, Buckendahl, and Camara) for their time and effort in responding to the *Legal Corner* column. They have provided multiple perspectives and thoughtful observations on the appointment of Justice Sotomayor and the *Ricci* case.

Camara indirectly raised the issue of conjunctive versus compensatory criteria, a favorite issue for plaintiffs seeking an alternative more favorable to minority candidates. While the use of compensatory criteria may qualify more minority candidates, conjunctive criteria may be more cost effective for test administrators because tests are given to only those

candidates who have met all other qualifications. Camara also identified an intriguing policy issue regarding the potential for public cynicism when an arguably fair and previously approved testing process is abandoned by policymakers due to disfavored results. As indicated, the perceived subjectivity and political implications of such actions by policymakers may in part account for the public's apparent preference for objective tests rather than more holistic candidate assessments.

Buckendahl reminded the measurement community that systematic and thorough test development practices consistent with the *Test Standards* are the best proactive defense to potential legal challenges to testing programs. Buckendahl also raised the issue of the sufficiency of harm to the plaintiffs (or unjustified benefit to minority candidates) to sustain a legal challenge when test results are abandoned after the fact. Although the Supreme Court found reverse discrimination in the *Ricci* case, its reasoning suggested that the result might be different if the evidence of job relatedness or test score validity were weak or inconsistent with professional standards. In such cases, cancelling the test results might be viewed as akin to dropping an ambiguous or flawed test item post administration.

Arnold points out that the *Ricci* decision did not change prior standards for employment testing challenges. Moreover, because discarding test results after the fact was disfavored as a remedy for disparate impact, the *Ricci* decision may encourage employers to (1) more carefully evaluate their test development procedures to ensure that adequate evidence of job relatedness and test score validity is collected and documented, and (2) more thoroughly review possible assessment alternatives before final implementation of a testing program.

WHAT? SO WHAT? AND NOW WHAT? THE VALUE OF SELF-REFLECTION

Carol L. Barry, James Madison University

This issue finds us at the end of another summer and at the beginning of another school year. Some of us are just starting our programs and courses of study, some of us are entering our second or third year, and others of us are beginning new phases of our graduate lives, perhaps beginning careers while still focusing on the ever-daunting dissertation. Regardless of where we are in our graduate studies, times like this (i.e., in between or at the beginning of semesters) are perfect for taking a step back and reflecting on where we are, where we are going, and what it means.



Self-reflection requires students to slow down and take the time to look inward, and, unfortunately, time is not something that graduate students have an abundance of. However, the value of reflection in learning has been well documented (e.g., Klenowski & Lunt, 2008; Langer, 2002). In fact, reflection seems to be an integral part of many doctoral programs, especially within the educational field (e.g., Klenowski & Lunt, 2008). Thus, some programs require their students to write reflections several times a year. However, even if this is not a requirement of your program, you might still find it beneficial to engage in these activities on your own time.

Advantages of Reflection

There are several advantages of reflection that are worth noting. First, it can help us identify personal areas of weakness. It's important to identify areas upon which we need or would like to improve. We can then use this information to help set appropriate goals (i.e., what do you want to improve and how will you do so?). Second, reflection also gives us the chance to identify areas of strength, which is especially important for graduate students who may tend to overlook the good and focus on the bad. Reflection gives us the chance to celebrate our accomplishments. Third, reflection provides an opportunity for us to think metacognitively about ourselves, to keep track of where we are, and to ensure that the path we are on is the path we want to be on.

Techniques and Tips

Although the benefits of reflection may be apparent, it can still be a difficult task to do. Given this, perhaps some "user-friendly" tips would be helpful. A colleague of mine recently told me that he has used the "What, So What, and Now What" method. The first step is to focus on the "What." For example, what did you do over the last semester with regard to coursework, research, and professional development? What did you not do? Next, the "So What" question involves thinking about what this means in regard to your development as a graduate student. Did you reach goals that you set previously? Were there things that you failed to accomplish? How did you grow as a result of these experiences? Finally, the "Now What" question involves using this information to set new goals for yourself. That is, having had these experiences, what do you hope to accomplish now? These three questions are good starting points for engaging in reflection.

It may also be helpful to focus on all areas of your life as a graduate student. That is, focus your reflective thoughts not only on coursework, research, and your assistantship/job, but also on your professional and personal development. Although it may require a little more effort, focusing on yourself holistically may better allow you to step back and see the big picture. Finally,

you may wonder how often you should engage in these self-reflective processes. It may be helpful to do this in between semesters or perhaps after big milestones (e.g., internships, thesis/dissertation proposals, etc.). In job settings, performance reviews (i.e., evaluations) tend to occur twice a year, so this might be another suggestion.

Ultimately, there are no hard and fast rules about reflection. You get to decide for yourself what works for you, what you want to focus on, and how often you want to do it. Regardless, I encourage you to set aside some time and take a good look inward.

References:

- Klenowski, V. Lunt, I. (2008). Enhancing learning at doctoral level through the use of reflection? *Assessment and Evaluation in Higher Education*, 33, 203-217.
- Langer, A. M. (2002). Reflecting on practice: Using learning journals in higher and continuing education. *Teaching in Higher Education*, 7, 337-351.

SPOTLIGHT ON THE PEOPLE WHO MAKE OUR ORGANIZATION GREAT – JIM IMPARA

Susan L. Davis, Alpine Testing Solutions

In our efforts to use this forum as a means for NCME members to get to know one another, the June 2009 NCME Newsletter featured an interview with Charlie Lewis.

This interview is with Jim Impara, currently with Plake & Impara Psychometric Inquiries after a diverse career in many areas of testing and a recent President of NCME. In an effort for us to get to know Jim, we asked him a number of questions and the excerpts of his responses are provided below.



How did you get into this field?

Jim: Like many people in this field, by accident. As an undergraduate at Florida State, I was living next door to a guy who was a graduate student in the Educational Research and Testing program. After learning more about what he was studying I found it to be very interesting work. As I was getting ready to graduate, I interviewed with the Department chair in this program and was offered a spot in the program and an assistantship.

How did your career path lead you to where you are now?

Jim: In 1966, after getting my masters degree from Florida State, I worked as a statistical clerk and evaluation specialist in the Florida Department of Education. The state of Florida, in 1971, began to implement its first state-wide assessment program and the design of this was assigned to the Bureau of Research, where I worked. Once the design was completed and approved by the Florida Legislature, I was assigned to direct this project. It was interesting building an assessment program from the ground up as I was entering this field. I later moved to Oregon to do the same type of job – leading a state-wide assessment program. From there I went to Virginia Tech to teach in the Research Methods program. I then moved on to the Buros Institute for Mental Measurement (BIMM) at the University of Nebraska-Lincoln as the co-editor of the Mental Measurements Yearbook and also did some teaching in the graduate program. While at Buros, I started the Buros Institute for Assessment Consultation and Outreach (BIACO). In 2003 I realized I was too busy and travelling too much so I decided to slow down to half-time and stepped down as director of BIACO. Although I was planning on not working too much I was invited to be a part of Caveon, a private company that focuses on test security, as a founder, Senior Director, and investor. For three years I was splitting my time between Caveon and Buros. I retired in 2006 from Buros and in 2008 from Caveon. Now I work as a consultant to testing programs in education and credentialing. Through all this experience I am still learning from every new project and testing program.

If you weren't doing this what would you do?

Jim: If I had not found my way to psychometrics I think I might have been teaching math as this is something I did before and have an undergraduate degree in Math Education. Counseling psychology would be another choice, as I enjoy working with people.

What advice would you offer a graduate student who is thinking about psychometrics?

Jim: Do it! This is a great field to be in. However, do not set yourself up with expectations that you are going to learn everything you need to know in school. Get as broad of a base of experience and education as possible so you can take advantage of new opportunities. Such experiences may come in the form of internships, research, or even through volunteer activities.

When not consulting, what do you do or like doing?

Jim: Out here in Colorado, I enjoy hiking, fishing, tying flies, and playing Farmtown on Facebook. I also do volunteer work for Trout Unlimited.

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

Jim: I would say the innovations in online testing. This is already working successfully for many testing programs including international programs. This is likely to become more commonplace for all types of testing programs in the future. The next step in online testing will be to resolve some of the outstanding security issues and concerns. Luckily there are some individuals in our field who are working on new innovations that will address these security issues. I see this eventually working well for most credentialing and educational testing programs.

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

Jim: Mostly I network and only attend a few sessions. There are several criteria I use to select a session. First, I like to go to sessions when I want to support someone who is presenting (friend, close colleague). Second, I would attend a session if one or more papers appeared to be relevant to a project that I am working on or contemplating working on. Third, I attend sessions that are about topics of interest to me.

Who has been a significant influence in your professional life?

Jim: There are three people I think of. First would be Howard Stoker, who was my doctoral chair and a major influence on my early professional career. Second would be Dick Jaeger who I met while working on a project for Florida Department of Education and who continued to be my friend until his untimely death. At the time we met we were both working on our doctorate and he mentored me in many ways. Third would be Barbara Plake who has been (and continues to be) a mentor to me. I have learned a lot from her both personally and professionally.

Where do you think the field is going next?

Jim: I don't know, the field is so diverse. More specifically, the underlying aspects of measurement are very common but when you look at the various testing fields such as educational, credentialing, and psychological, they appear to be very different. As professionals, it is very easy to get locked in to one area of testing. I would like to see greater unification across these areas of testing so professionals in one area can learn from professionals in another area. However, I don't know if most of us are getting the diverse range of experiences we need right now to realize the learning opportunities that cut across the other areas in testing. We need to step back and see how we can take the best of all possible strategies from the different areas of testing and apply these across the field.

If you were starting in this field now, what would you want to specialize in?

Jim: I probably would not specialize; I would do the same thing all over again! I have been so incredibly lucky to have such a diverse set of experiences. Knowing how many different types of work there are within the field I would miss the challenges that come with adapting to different areas of work in each phase of my career.

You are currently balancing part-time retirement and multiple consulting projects – how do you keep this balanced?

Jim: For the most part I try to keep a small number of clients so I don't have to worry about work every day. Any time I do not have specific work responsibilities I try to spend hiking or fishing!

CALL FOR NOMINATIONS

AERA, Division D**Call for Nominations for the 2010 Award for Significant Contribution to Educational Measurement and Research Methodology – Due by November 30, 2009**

Division D of AERA welcomes nominations for the 2010 Award for Significant Contribution to Educational Measurement and Research Methodology. This annual award recognizes published research judged to represent a significant advancement in

theory and practice of educational measurement and/or educational research methodology. This award is not considered a lifetime achievement award; the significance of the contribution will be the primary consideration for this award. The research may be the work of an individual or a team of researchers. The winner will be announced and honored at the 2010 AERA annual meeting with a plaque and a \$1,000 award.

GUIDELINES

In selecting a winner, the following guidelines will apply:

- Quality and potential impact of the research on educational measurement and research methodology are the primary criteria for this award.
- The recognized publication may be, but is not limited to, a refereed research article in either a print or online journal, a paper published in a refereed conference proceeding, monograph, book chapter, and/or book. The work must have been published between August 1, 2007 and July 31, 2009.
- The nominee(s) must be the first or sole author(s) of the work and must be a member of Division D of AERA.
- The nomination should be submitted electronically unless the material is not available in that format.

APPLICATION PROCEDURE

A complete nomination consists of:

- The nomination letter (self nominations are welcome);
- A copy of the nominated research publication including its bibliographic citation. If the publication is a book or monograph, the nominator should indicate which portion of the book or monograph is nominated for this award;
- At least one additional letter of recommendation (from person(s) other than the nominator) addressing the quality and potential impact of the research; and
- The nominee's vita.

The nominator agrees that this work will not be nominated for a similar award that will be given during the 2010 Annual Meeting of either AERA or NCME.

Submit the complete nomination for the above awards by **November 30, 2009** to

Phoebe C. Winter, Chair
Pacific Metrics Corporation
585 Cannery Row, Suite 201
Monterey, CA 93940
pwinter@pacificmetrics.com

NCME – Please go to the following link for details: <http://www.ncme.org/about/awards.cfm>

Call Bradley Hanson Award for Contributions to Educational Measurement – Due by November 13, 2009

The Bradley Hanson Award has been established to honor Bradley Hanson's contributions to the field of educational measurement and to further advance the goals embodied in his work. Applicants must propose a research project that promises to make a substantive contribution to the field of educational measurement and / or the development, instruction, or mentoring of new professionals in the field. A typical time frame for the expected completion of the proposed project is one to two years. The recipient will be awarded \$1,250 and a commemorative plaque from NCME, which will be presented to the recipient at the 2010 NCME Annual Meeting in Denver from April 29 – May 3. For additional information on the award, and past recipients and projects, please go to Bradley Hanson's homepage (<http://www.b-a-h.com/>).

Eligibility: To be eligible for the award a candidate must (1) be a member of NCME or be nominated by an NCME member and become a member by the time the award is presented at the 2009 NCME Annual Meeting; and (2) be working on a project that promises to make a significant contribution to the field of educational measurement and or a significant contribution to the development of new professionals in the field. Individuals or groups of individuals at any stage in their careers are eligible.

Call for the 2010 Jason Millman Promising Measurement Scholar Award – Due by November 13, 2009

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 and will be presented the award at the annual meeting.

Eligibility: To be eligible for the 2010 Jason Millman Award an applicant must have received the Ph.D. after April 1, 2005. If confirmation is required, the review committee will use the date printed on the nominee's diploma. The nominee must also have two or more unique papers either accepted for presentation at an NCME annual meeting or published in NCME publications within the last five years; and have the support of his/her professional colleagues indicating that his/her work represents a significant contribution to the field of applied measurement.

Call for the 2010 Brenda H. Loyd Outstanding Dissertation Award – Due by November 13, 2009

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.

Eligibility: Nominations will be accepted for dissertations completed between July 1, 2007, and June 30, 2009. The author of the dissertation need not be a member of NCME. However, the author's advisor must be a member of NCME.

Call for the 2010 Alicia Cascallar Award for Outstanding Paper by an Early Career Scholar – Due by November 1, 2009

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(s) who presented an outstanding paper at the most recent Annual Meeting. A cash award of \$1,000, a citation, and a waiver of NCME conference fees for the following year will be provided as partial support for an early career member(s) of NCME to travel to the annual meeting. The award will be presented at the NCME Annual Meeting in 2010.

Eligibility: To have presented a paper at the most recent (2009) 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).

Call for the 2010 NCME Award for Career Contributions to Educational Measurement – Due by November 30, 2009

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 present an address at the next year's NCME Annual Meeting.

Eligibility: Nominations should be made by an NCME member.

Call for the NCME Award for Outstanding Dissemination of Educational Measurement Concepts to the Public – Due by December 4, 2009

In the year 2010, NCME will honor outstanding dissemination of educational measurement concepts to the public in 2007, 2008, or 2009. This award is to recognize persons who have made an exemplary contribution to the public's understanding of information concerning tests and measurement issues and practices. Examples of dissemination of measurement concepts include, but are not limited to, books, workshops, articles in newspapers and popular magazines, radio or television presentations, and legislative testimony. Selection criteria include judgments about the significance of the contribution to the public's understanding of measurement issues and practices, the breadth of the dissemination across the public, and the magnitude of the impact of the contribution in increasing public understanding of educational measurement concepts.

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 a 3-5 page statement summarizing the nature of the dissemination of educational measurement concepts to the public. Nomination statements should clearly describe contribution and its significance, as well as the means of dissemination, and how it has increased public understanding of educational measurement concepts. Additional supporting documentation of the contribution (e.g., the relevant book, article, video of presentations and workshops, etc.) is recommended. Nominations should also include the names and addresses of two persons familiar with the contribution. The review committee may request further materials and may contact others who are likely to be able to evaluate the dissemination. The award will be presented at the NCME Annual Meeting in Denver, April 2010.

OTHER CONFERENCES

Eastern Educational Research Association

President: Abbot L. Packard, University of West Georgia

Annual Conference: February 10-13, 2010

Location: Hyatt Regency, Savanna, GA

Web page: www.eeralonline.org

Mid-Western Educational Research Association

President: Dimiter Dimitrov, George Mason University

Annual Conference: October 14-17, 2009

Location: Sheraton Westport Chalet, St. Louis, MO

Web page: www.mwera.org

Mid-South Educational Research Association

President: Dana Thames, University of Southern

Mississippi

Annual Conference: November 4-6, 2009

Location: Hilton, Baton Rouge, LA

Web page: www.ms-era.org

New England Educational Research Organization

President: Patricia Paugh, University of Massachusetts,
Boston

Annual Conference: April 7-10, 2010

Location: Harborside Hotel, Portsmouth, NH

Web page:

<http://faculty.education.uconn.edu/edlr/cdc03005/neero/about.html>

Northeastern Educational Research Association

President: Kristen Huff, The College Board

Annual Conference: October 21-23, 2009

Location: Marriott Rocky Hill, Rocky Hill, CT

Web page: www.nera-education.org

Northern Rocky Mountain Educational Research Association

President: Suzanne Young, University of Wyoming

Annual Conference: October 8-10, 2009

Location: Snow King Resort, Jackson Hole, WY

Web page: www.nrmera.org

Southwest Educational Research Association

President: Mary Margaret Capraro, Texas A&M Univ.

Annual Conference: February 17-20, 2010

Location: The Hotel Monteleone, New Orleans, LA

Web page: www.sera-education.org

NEWSLETTER ADVISORY BOARD

CAROL L. BARRY, James Madison University (Grad Student Representative)
SCOTT BISHOP, Data Recognition Corporation
MARY LYN BOURQUE, Mid-Atlantic Psychometric Services
SUSAN M. BROOKHART, Consultant
SUSAN L. DAVIS, Alpine Testing Solutions
ELLEN FORTE, edCount LLC
EDWARD H. HAERTEL, Stanford University
SARA S. HENNING, Consultant

JOAN HERMAN, CRESST/UCLA
JOANNA GORIN, Arizona State University
THEL KOCHER, Edina Public Schools, Minnesota
GERALD MELICAN, The College Board
S.E. PHILLIPS, Consultant
CHRISTINA SCHNEIDER, CTB/McGraw-Hill
DONNA L. SUNDE, James Madison University
XIANG (BO) WANG, The College Board

THANOS PATELIS, Editor, The College Board

Send articles or information for this newsletter to:

Thanos Patelis
The College Board
45 Columbus Avenue
New York, NY 10023

Phone: 212.649.8435
Fax: 212.649.8427
e-mail: tpatelis@collegeboard.org

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.