

The NERA Researcher

JUNE 2007

VOLUME 45, ISSUE 2

Newsletter of the Northeastern Educational Research Association

Editors: *Geraldine Mongillo, William Paterson University, and Heejung An, William Paterson University*

THE PRESIDENT'S MESSAGE

STEPHEN G. SIRECI



NCLB: The Good, The Bad, and The Ugly Unanswered Research Questions

This fall, Congress will vote on reauthorization of the No Child Left Behind Act (NCLB). Like most mammoth Federal laws, there are some good things, and some bad things, about it. In this column, I present my thoughts about what I see as the positive and negative consequences of the law thus far, and I call attention to some important issues regarding the assessment policies associated with this law that should be addressed in its reauthorization.

NCLB: The Good

There is so much criticism against NCLB that it may seem unfathomable that I think there are positive aspects of this law. The most obvious positive consequence I see is the increased attention it has brought to the education of minority students and students with disabilities. Advocates for students with disabilities seem particularly blown away with the amount of attention paid to these students under NCLB. Not only are students with disabilities one of the subgroups of students for whom the reporting of test results is mandated, NCLB also addresses test accommodations for such students, as well as alternate assessments for severely disabled students. Perhaps for the first time, the vast majority of students with disabilities are being included in statewide assessments, and schools are being held accountable for get-

ting them to meet educational goals. In addition, NCLB has awarded over \$11 billion in special education grants.

Schools are also held accountable for the performance of other important subgroups of students such as ethnic subgroups, students from low socioeconomic status households, and English language learners. Where particular subgroups do not demonstrate "adequate yearly progress," it is expected that improvement plans will be developed to increase the number and proportion of students from these groups that meet the academic achievement standards set in the state (i.e., attain "proficient" in each subject area). As I explain later, I think the accountability mechanism of NCLB is unfair to schools and districts with large numbers of subgroups of students. However, it is hard to argue with mandated reporting for these student subgroups and the requirement that schools and districts develop improvement plans to close achievement gaps. Through the data provided by NCLB, educators, parents, policy makers, and the general public can see the

(Continued on page 7)

TABLE OF CONTENTS

Message from the Editors	2
NERA Officers, Board Members & Chairs	2
2007 Conference Highlights	3
Conference Pre-Session Announcement	4
Conference Hotel Information	5
Teacher-as-Researcher Award 2007	6
Important Dates	6
The Graduate Lounge	7
Donlon Memorial Award Call	7
Conference Registration Forms	insert

Message from the Editors

Hello NERA Members! It is time for many of us to wind down from our hectic schedules and ease into summer mode. We hope that you had time to submit your proposals for the conference...our chairs tell us that there was an impressive response to the call! This year's conference offers several interesting pre-sessions and distinguished keynote speakers. Check this month's issue you will find all the information you will need regarding our annual conference and be sure to make your reservations early for the new venue in Rocky Hill, Connecticut.

It was great to see so many of you at AERA! Why not share some news about your trip...Did you hear a great speaker? Discover interesting new research? Make new friends? We would love to print your insights so please send us your thoughts!

Have a restful and enjoyable summer!

- Gerri & Heejung

The Editors

Geraldine Mongillo

William Paterson University
Wayne, NJ 07470
mongillog@wpunj.edu

Heejung An

William Paterson University
Wayne, NJ 07470
anh2@wpunj.edu

NERA Officers, Board Members & Chairpersons

Stephen G. Sireci, President

[University of Massachusetts, sireci@acad.umass.edu](mailto:sireci@acad.umass.edu)

Theresa Rooney, Past President

[York College/CUNY, rooney@york.cuny.edu](mailto:rooney@york.cuny.edu)

David M. Moss, President-Elect

[University of Connecticut, david.moss@uconn.edu](mailto:david.moss@uconn.edu)

Darlene Perner, Secretary

[Bloomsburg University, dperner@bloomu.edu](mailto:dperner@bloomu.edu)

Barbara J. Helms, Treasurer

[Education Development Center, Inc., bjhelms@gmail.com](mailto:bjhelms@gmail.com)

Amy Dresher, Program Co-Chair (2006-2007)

[Educational Testing Service, adresher@ets.org](mailto:adresher@ets.org)

April Zenisky, Program Co-Chair (2006-2007)

[University of Massachusetts, azenisky@educ.umass.edu](mailto:azenisky@educ.umass.edu)

Dolores Burton, Director (2006-2009)

[New York Institute of Technology, dburton@nyit.edu](mailto:dburton@nyit.edu)

Kerry Dunn, Director (2006-2009)

[Marist College, kerry.dunn@marist.edu](mailto:kerry.dunn@marist.edu)

Kristen Huff, Director (2004-2007)

[College Board, khuff@collegeboard.org](mailto:khuff@collegeboard.org)

Katharyn Nottis, Director (2005-2008)

[Bucknell University, knottis@Bucknell.edu](mailto:knottis@Bucknell.edu)

Brian Preston, Director (2005-2008)

[Southwestern Boces, bpreston@swboces.org](mailto:bpreston@swboces.org)

Barbara Yingling Wert, Director (2006-2009)

[Bloomsburg University, bwert@bloomu.edu](mailto:bwert@bloomu.edu)

Michael Young, Director (2004-2007), Tech. Committee Chair

[University of Connecticut, myoung@uconn.edu](mailto:myoung@uconn.edu)

Dolores Burton, Chair, Teacher-as-Researcher Award Committee

[New York Institute of Technology, dburton@nyit.edu](mailto:dburton@nyit.edu)

Dianne Zager, Chair, Leo D. Doherty Memorial Award Committee

[Pace University, dzager@pace.edu](mailto:dzager@pace.edu)

Diane Liebert, Co-Chair Distinguished Paper Award Committee

[DLE Associates, diane444@comcast.net](mailto:diane444@comcast.net)

Douglas Penfield, Co-Chair Distinguished Paper Award

[Rutgers University, doupen@rci.rutgers.edu](mailto:doupen@rci.rutgers.edu)

Liora P. Schmelkin, Chair, Tom Donlan Memorial Award Committee

[Hofstra University, Liora.P.Schmelkin@hofstra.edu](mailto:Liora.P.Schmelkin@hofstra.edu)



38th Annual Conference, October 17-19, 2007

Selected Program Highlights include:

▶ Keynote Speaker: **Cora B. Marrett**
Assistant Director of Education and Human Resources, National Science Foundation

▶ Keynote Speaker: **Wayne J. Camara**
Vice President for Research and Development, The College Board

▶ Conference Pre-Sessions
Educational Program Evaluation with Robert Gable
Item Response Theory with H. Swaminathan and Jane Rogers

▶ Featured Sessions
Using Assessment Data: Practice and Research
Teacher-as-Researcher Award Winner
Graduate Student Issues Committee Session
Reflections from NERA Past Presidents

New Location for 2007!

Hartford Marriott Rocky Hill in Rocky Hill, Connecticut

Book your hotel accommodations online at <http://cwp.marriott.com/bdlrh/nera/>
or by calling 1-800-228-9290 and mentioning NERA (or the NERA block)

For more information about the 2007 Conference,
contact the Conference Co-Chairs, Amy Dresher and April Zenisky
neramembers@gmail.com

**Northeastern Educational Research Association (NERA)
38th Annual Conference, October 17-19, 2007**

**CONFERENCE PRE-SESSIONS
October 17, 2007**

**PRE-SESSION A:
12:15 p.m.-2:15 p.m.**

Program Evaluation Methodology: Are We Thinking Clearly?

Robert K. Gable
Emeritus Professor, Neag School of Education
University of Connecticut
Director, Educational Leadership Doctoral Program
Johnson & Wales University

This pre-session will focus on current Federal policies and controversies in program evaluation methodology. Topics will include: the history of program evaluation models, the role of quantitative and qualitative research strategies for effective program evaluation, and the implications of the US Department of Education and Institute of Education Sciences promotion of Evidence-Based Education (EBE).

**PRE-SESSION B:
12:15 p.m.-2:15 p.m.**

Introduction to Item Response Theory

H. Swaminathan
Professor, Neag School of Education
University of Connecticut

H. Jane Rogers
Associate Professor, Neag School of Education
University of Connecticut

This session covers the basics of item response theory and its use in educational testing applications. Topics include the models available for dichotomous and polytomous responses, computer programs available for IRT scoring, procedures for assessing goodness of fit, and an overview of applications of IRT to equating, test construction, and computerized adaptive testing. The session is aimed at meeting the needs of individuals who are or expect to be using test scores based on IRT procedures and need an understanding of the basic concepts and issues. Participants with BILOG and/or PARSCALE are encouraged to bring their laptops, but laptops are not required.

**Northeastern Educational Research Association (NERA)
38th Annual Conference, October 17-19, 2007**

**Hotel Information and Registration Deadline
Hartford Marriott Rocky Hill**

<http://cwp.marriott.com/bdlrh/nera/>

100 Capital Boulevard, Rocky Hill, CT 06067
Phone: 1-860-257-6000; Fax: 1-860-257-6060

NERA's Annual Meeting has a new venue for 2007! Located in Rocky Hill, Connecticut, the Hartford Marriott Rocky Hill will host this year's conference. Among the amenities of this location we'd like to share with you are to-be-newly renovated meeting rooms, a variety of transportation options for getting you to the conference in comfort, and the choice of booking your accommodations either online or by phone.

**Important Notice regarding 2007 Conference Rates:
Make Your Hotel Reservations Early!**

The following hotel rates are in effect for the 2007 NERA Conference, **but you must book by September 26, 2007 to obtain these rates:**

- Single occupancy room: \$175.00 per night
(includes dinner on night of arrival and breakfast and lunch the next day)
- Double occupancy room: \$127.25 per person, per night (total: \$254.50 per night)
(includes dinner on night of arrival and breakfast and lunch the next day for each person)

Hotel rooms booked after **September 26, 2007** are charged at the following rates (and include meals as well):

- Single occupancy room: \$254.00 per night
- Double occupancy room: \$167.00 per person per night (total: \$334.00 per night)

You can book your hotel accommodations online at <http://cwp.marriott.com/bdlrh/nera/> or by calling 1-800-228-9290 and mentioning NERA (or the NERA block)

Traveling to the Hartford Marriott Rocky Hill

The Hartford Marriott Rocky Hill is located just off Route 91 in Rocky Hill, CT. Rocky Hill is ten minutes from Hartford, including the train and bus stations. In addition, Bradley International Airport in Windsor Locks, CT is just 25 miles away (estimated one-way taxi fare is \$40). The Hartford Marriott Rocky Hill provides complimentary on-site parking for conference attendees.

We look forward to seeing you in Connecticut at the 2007 Conference!
April Zenisky and Amy Dresher, 2007 NERA Conference Co-Chairs

Teacher-as-Researcher Award 2007

The Teacher-as-Researcher award, established by NERA in 1993 to promote educational research and to encourage the development of research among junior researchers, focuses on recognizing teachers for outstanding efforts to conduct Action Research in their classrooms and to use the outcomes to improve teaching and learning. Action Research has been defined as, "a systematic, sustained, and publicly shared way of learning and improving one's self and one's practice" (Arhar, Holly, & Kasten, 2001). Each year we honor a teacher researcher that has identified a question, issue, or problem, defined a solution or intervention, applied the intervention, collected data regarding the intervention, analyzed the findings and used the information to take action.

Classroom teachers are invited to apply directly for this award or be nominated by NERA members, school administrators, faculty mentors, or others familiar with the teacher's work. All applications/nominations should follow the format of the Teacher-as-Researcher Award application form and be postmarked or emailed as an attachment in Word no later than June 1, 2007. Information regarding the award is available on the NERA website at http://www.nera-education.org/teacher_as_researcher_award.html.

The 2007 awardee will be invited to speak about the research project at a special session at the annual NERA meeting in October and be presented with the award at that time. The award includes a plaque, NERA membership, \$150 toward travel, and two full-days of meals and lodging at the Hartford Marriott Rocky Hill Hotel, Hartford CT, NERA's conference site.

As you think about Teacher Researchers you may nominate for the Teacher-as-Researcher award you may wish to share some information on teacher research available on these internet sites.

- http://www.teach-nology.com/currenttrends/teach_as_rese/: A web portal sponsored by TeAch-nology.com that contains free resources on teaching with technology and links to articles on the history and practice of action research.
- <http://www.scu.edu.au/schools/gcm/ar/arp/arphome.html>: Contains links to papers and resources on action research.

- <http://www.ar.blend-xl.eu/>: A website with definitions of action research and links to resources about Kurt Lewin (1890-1947), Lawrence Stenhouse and Donald Schön (1930-1997)
- <http://www.ericdigests.org/1993/researcher.htm>: Contains the Teacher-As-Researcher. ERIC Digest (Johnson, 1993) from the ERIC Clearinghouse on Teacher Education Washington DC.
- <http://gse.gmu.edu/research/tr/>: A site sponsored by George Mason University which provides resources and access to teacher researcher networks.

We thank you for your support of this important recognition of teacher practice. Please feel free to send questions or comments regarding the Teacher-as-Researcher award to dburton@nyit.edu.

Dolores Burton, Chairperson
Teacher-as-Researcher Award Committee

Northeastern Educational Research Association 38th Annual Conference

Save the Dates!

- April 1, 2007: Hotel Registration Opens (<http://cwp.marriott.com/bdlrh/nera/>)
- June 1, 2007: Proposals Due
- June 15, 2007: Proposals to Reviewers
- July 15, 2007: Completed Reviews Due to Conference Co-Chairs
- August 15, 2007: Notification of Proposal Acceptance
- September 1, 2007: Conference Schedule Notification
- September 26, 2007: Hotel Registration Ends
- October 17-19, 2007: NERA's 38th Annual Conference



The Graduate Lounge



Graduate Student Issues Council of the Northeastern Educational Research Association (NERA) cordially invites all conference attendees (especially graduate students!) to attend the inaugural

Graduate Student Issues Session

at the 38th annual NERA Conference
October 17–19, 2007
Hartford Marriott
Rocky Hill, Connecticut

First of its kind at NERA, the **Graduate Student Issues Session** was organized by the Graduate Student Issues Council for all NERA conference attendees, with special focus on common interests of graduate students. The Session will have well-known expert guest speakers to talk about their experiences and recommendations in publishing in professional journals, exploring career options, and other issues related to graduate student life. There will be Question and Answer section at the session that will allow session participants to address their individual needs and concerns.

Thomas F. Donlon Memorial Award for Distinguished Mentoring Call for Nominations

Make a nomination! Tell us 'who made it happen' for you. Send your nomination with a description of why that person is instrumental in your academic/professional life to: Liora Schmelkin, Chair, Donlon Award Committee.

All nominations should be e-mailed to her at Liora.P.Schmelkin@hofstra.edu.

(Continued from the President's Message—page 1)

great differences that exist in student performance across subgroups and discuss what can be done to improve the education of these students. Even in my little, predominantly Caucasian, district of Northampton, Massachusetts, teachers and administrators are discussing strategies for improving the academic achievement of students from specific ethnic groups. In general, a positive outcome of NCLB is that schools are using student achievement data to make instructional changes targeted to improved student learning and achievement.

Most readers will probably agree that the increased attention NCLB has given to traditionally underperforming or excluded student groups is a positive consequence of this law. However, the claim that NCLB has improved teaching and learning is less defensible and so it requires explanation and qualification. The benefits I see to teaching stem from the consensus processes states used to develop curriculum frameworks in several subject areas. The curriculum framework development process includes educators throughout the state to specify the educational objectives deemed most important in specific subject areas and grade levels. State-wide tests are aligned with these frameworks and are designed to measure the objectives within them. How has this curriculum-assessment alignment improved

(Continued on page 8)

(Continued from the President's Message—page 7)

teaching? Teachers who did not have comprehensive lesson plans or who were teaching skills that were far too basic for their class have been forced to alter their teaching to address the objectives on which their students are tested. For such teachers, and there are more of them than we would like to admit, the mandated curricula coupled with mandated assessments has forced them to improve their lessons.

The flipside of this issue is that there are numerous skilled teachers who were already teaching important skills listed in the curriculum frameworks, and much more. The mandated curricula may hamstring such teachers by forcing them to cover certain topics in certain ways, at the expense of their more creative and insightful lesson plans. Thus, NCLB has probably improved the teaching practices of substandard teachers, but frustrated or reduced the effectiveness of our best teachers. I see this as an important qualification, or at least a plausible qualification, of any claim that NCLB has improved instruction.

Returning to positive consequences, another positive aspect of NCLB is the information it provides to parents. I previously mentioned the value of student assessment data to teachers, schools, and school districts with respect to identifying subgroups of students who need attention. However, the information provided to parents is particularly invaluable. Consider the score reports parents receive from statewide testing programs relative to the score reports our parents received when we were in school. 20th-century score reports focused on norm-referenced information, particularly percentile ranks in specific subject areas and sub-domains within an area. Such information is useful, but I find the criterion-referenced information I receive for my sons more useful. Let me give you an example of what I mean by "criterion-referenced" information. Score reports for the statewide tests in Massachusetts assign a scale score and an achievement level score for Math and English Language Arts, number correct sub-scores within these subject areas (e.g., composition, number sense, etc.), students' responses to all multiple-choice test questions, and their scores on the open-response items. If I want, I can go onto the Department of Education's web site and see the actual items my sons got right or wrong. Furthermore, I can talk to the teachers or principal about things that were covered on the test (and hence, in the state curriculum frameworks), but were not covered by the teacher. Clearly, this test score information elevates the conversations parents can have with their teachers about what is happening in the classroom, and more importantly, gives parents a good idea of how their chil-

dren are doing with respect to clearly defined targets of academic achievement defined by curriculum experts within the state.

The last positive consequence of NCLB I will point out is the professional development opportunities it provides for teachers and administrators. This benefit, and others related to statewide testing, was pointed out by Cizek (2001). Professional development for teachers has included training in item and test development, developing scoring rubrics and standardized methods for scoring open-response items, reviewing test items for content quality and alignment to the curriculum frameworks, and how to use assessment results to improve instruction.

NCLB: The Bad

It is hard to talk about the positives of NCLB without thinking of the negatives because there are definitely two sides to the NCLB coin. I already mentioned that the effectiveness of many teachers might be reduced due to constraints on their prior teaching practices. Much worse, however, is the instructional time that has been lost due to the volume of tests students are required to take. Some may say that I am an advocate for tests and testing, but the truth is I am an advocate only for *sensible* testing. In my opinion, we are currently testing our students too much. Again using Massachusetts as an example, in some grades up to 11 school days are taken up by assessments. Can you believe that? I do not need 11 days to report how well elementary students are doing in math and reading. The accountability demands of NCLB, and state accountability demands, currently involve testing students not only in math and reading, but in writing, science, social studies, and other areas. You can add English proficiency testing to the list for English language learners, too. Sensible educational assessment policy requires balancing accountability needs and the cost to instruction due to the time it takes to administer tests. I do not think we need to test all subjects and all students in all grades for accountability purposes. Sampling strategies that would sample students within grades, or sample subjects across grades, should be used to reduce the testing burden at the student and classroom levels. We can and should be useful to policy makers on that issue.

Another negative consequence of NCLB is the incredible stress it has placed on everyone throughout the educational community. Principals and other administrators certainly feel it, and that is no accident. They are probably the appropriate targets for accountability decisions. The teachers certainly feel stress on many levels. They

(Continued on page 9)

(Continued from the President's Message—page 8)

do not want to be portrayed as doing a bad job and they do not want to see the students with whom they are working miss achievement targets. I think some level of pressure on teachers to raise student performance is appropriate, but teachers should not be held accountable for the performance of their students in a strict sense, because they are dealt a certain hand and are limited in what they can do. That is, students are not randomly assigned to teachers and schools are not equivalent with respect to the resources available to them for improving instruction. The idea of using student achievement test data to evaluate teachers is a bad one, unless it is used as a relatively small component in a larger evaluation effort conducted by the principal or someone else familiar with the particular characteristics of the teachers' students, the school environment, the community, and available resources.

Students are also stressed out about all the tests they take. This baffles me because, at most grade levels, there are absolutely no consequences for students. Nevertheless, I see the stress and it is widespread. When I asked one principal where this stress comes from, she replied "everywhere." Clearly, the stress teachers feel is being passed to their students. I am not aware of any training programs for teachers to help them deal with new accountability pressures and how to prepare students to take tests. Such training should be provided and it should include ways of relaxing students and appropriate ways to engage them in positive discussions about tests and testing. The stress NCLB-related tests have caused throughout the school system has led to a perceived demoralization of teachers (Sireci, Martone, & Lewis, 2006) and may result in increased test anxiety for students.

Another negative consequence of NCLB is the money it has sucked out of the school system to develop the too-many tests that are being administered. The annual budget for statewide testing in Massachusetts is in the tens of millions of dollars. Now consider the annual testing budgets for all states and territories. Although I have not figured out the total, I bet at least a half-billion dollars is spent on test development alone across the United States each year. Think of what we could do with the remaining funds if we cut that cost in half or in three-quarters. Given that many of the states are testing very similar content, the idea of a national test does not seem so bad anymore. In fact recent collaborative developments across states, such as the New England Common Assessment Program (New Hampshire, Rhode Island, and Vermont), and the consortium of 9 states

who pooled their resources to develop a common Algebra-2 test (see <http://www.achieve.org/node/836>) illustrate that some states are finally becoming sensible with respect to how to efficiently spend funds for test development.

Another particularly severe negative consequence of NCLB is the decreased attention paid to students who are above grade-level proficiency, sometimes referred to as "gifted and talented students." The accountability demands of NCLB seem to be all about getting underperforming students up to "proficient." Schools and districts can do quite well under NCLB by ignoring higher-achieving students and focusing on students below the proficient level. This irks me. In my little rural district, there are no programs for gifted and talented students until high school. In my opinion, these are the children who are left behind under NCLB.

The negative side of the NCLB coin also qualifies the claim that NCLB has helped educators use data to improve instruction. While I think this claim is generally true, the data provided by statewide assessments is useful, but extremely limited, for the purpose of instructional improvement. Statewide tests assess general proficiency in a subject area and tend not to be very useful for diagnostic purposes. The data provided to schools at the learning objective level may not be very reliable because very few items may measure an objective and the objectives measure may vary from year to year. Thus, it is possible that teachers and administrators are drawing unsubstantiated inferences from aggregated or individual student test results. Again, more training is needed regarding how teachers and administrators can use test results to improve instruction. Such training should include discussions of the limitations of these data and potential misinterpretation.

There is a fundamental problem in the NCLB accountability process that may result in a particularly troublesome negative consequence that is counter to the entire accountability enterprise—flagging the wrong schools for under-performance. This problem stems from using an accountability yardstick that focuses on achievement levels, rather than on a continuous score scale. Specifically, NCLB requires schools to demonstrate "adequate yearly progress" and defines such progress as increases in the percentages of students who are designated "proficient" in reading and mathematics in grades 3 through 8 and one grade in high school. Anyone who has taken the most elementary of statistics courses knows that when you slice up a continuous score scale into discrete groups, for example when test scores are

(Continued on page 10)

(Continued from the President's Message—page 9)

Table 1. Summary of NCLB Positive and Negative Consequences

Effect	Positive Consequences	Negative Consequences
Improving achievement of subgroups of students	Much more inclusion of SWD and ELLs in assessments; focus on narrowing achievement gaps	Schools with more subgroups have more accountability hurdles to overcome
Improved teaching and learning	Poorly organized teachers have better lesson plans	Innovative teachers may be stifled
School accountability	Pressure on administrators to improve student achievement	Wrong schools flagged or not flagged based on poorly designed accountability mechanism
Provides data for school improvement	Identifies general curricula areas in particular need of improvement	Schools institute curricula changes based on data that may be unreliable or is too general
Differentiated Instruction	More attention paid to lower-achieving students	Less attention paid to higher-achieving students
Cost	Benefits (see above) associated with test results	Less money for schools from state.
Provides data for parents	Parents can see how their children measure up to state standards. Elevates parent/teacher discussions.	
Professional development	Teachers/administrators trained in curriculum development, aligning assessment and instruction, scoring rubric development	
Increased testing time		Reduced time for instruction
Stress		Demoralization of teachers, increased test anxiety

divided into “proficient” and “below proficient,” information is lost. Thus, the high-stakes process used to evaluate schools uses a very crude metric. Students who learn a lot in a given year and who have large test score increases, but who remain in the same “proficient” or “below proficient” category, will not contribute to a positive evaluation of the school’s yearly progress. Students with very small educational gains who happen to be near the proficient cut-point in a given year will have a much bigger impact on a school’s adequate yearly progress. If that sounds counter-intuitive to you, then you understand the problem.

A related problem associated with this negative consequence is that the accountability evaluation focuses on all students within a school and specific subgroups; but the subgroups of students are not consistent across schools. Thus, there is a negative aspect to the explicit focus on subgroups of students. Although focusing on diverse groups of students is good for the students, it may not be good, or fair, for the schools. As Linn (2005) pointed out, a school with several subgroups (e.g., high poverty, African American, Asian American, Hispanic/Latino, students with disabilities, English language learners, etc.) must surpass more accountability-hurdles (i.e., proficiency targets for each group) relative to schools with fewer subgroups. Measurement and

sampling error are involved in each subgroup that is evaluated for adequate yearly progress, and this error aggregates as the number of subgroups increases. Thus, for a relatively homogenous school or district, the accountability demands are actually less than those imposed upon a more diverse school or district. Some policy researchers have advocated for longitudinal growth models that would focus on scale score changes for common groups of students over time. Such approaches have merit and limitations, but in general, any strategy that focuses on scale scores rather than achievement level classifications will be an improvement.

Summary of the good and bad

My preceding discussion is clearly not an exhaustive list of the positive and negative aspects of NCLB thus far. Rather, the consequences I discussed are the most immediate impressions I have based on reflecting on NCLB for this column. Although not exhaustive, these thoughts are based on several years experience consulting for state departments of education, the U.S. Department of Education, and national research centers such as the National Center for Educational Outcomes. A summary of the positive and negative consequences discussed is presented in Table 1. As can be seen from this Table most of the positive consequences have negative

(Continued on page 11)

(Continued from the President's Message—page 10)

counterparts. In any case, this summary provides a helpful starting point for considering the research to be done to help inform assessment policy within the context of NCLB. I frame my discussion of possible research in terms of ugly unanswered questions.

The ugly unanswered research questions

Given that we are educational researchers, what can we do to inform sensible educational assessment policy in the context of NCLB? What types of recommendations can we make to Congress? I believe there are some recommendations we can make at this point, such as minimize testing time and consider alternative strategies for computing adequate yearly progress. However, given space constraints, I will highlight three ugly, but important, questions that require much more research before we can provide sensible advice to policy makers. Space will not permit me to discuss any of these in detail and so I pose each one with only some brief commentary.

1. What types of test accommodations increase the validity of test score interpretations and for which types of students?

All states provide test accommodations for students with disabilities, and most provide accommodations for English language learners. Examples of such accommodations include extra time, reading test material aloud, use of a scribe, and translating a test from English into a students' native language. There is currently little research on the degree to which such accommodations are helpful to the students and the degree to which the increase test score validity.

2. What are the best methods for evaluating adequate yearly progress of a school or district?

I believe the current "status" model where different cohorts of students are compared across years (e.g., grade 3 students in 2006 and grade 3 students in 2007) with respect to achievement level percentages is not the best way to identify schools in need of improvement or sanctions. Longitudinal growth models, index-based models need to be investigated for this purpose, and other ideas should be explored.

3. How can states defend (validate) the achievement level standards they have set?

Research has shown that the states differ widely with

respect to the proficiency achievement level targets they have set (Linn, 2005; McLaughlin et al., 2005). Many states use the achievement level standards set by the National Assessment of Educational Progress (NAEP) to evaluate their targets, but there are few studies that evaluate the alignment of state tests to NAEP. Aside from NAEP-state alignment studies, studies that use external data to evaluate the standards set by a state are also needed (see Kane, 1994).

Closing remarks

In this column, I argued that NCLB has resulted in both good and bad consequences, but there is much more work to be done by educational researchers. I encourage all of us to engage in research that will help inform educational assessment policy and I hope to see some papers and sessions on such research at NERA 2007!

References

- Cizek, G. J. (2001). More unintended consequences of high-stakes testing. *Educational Measurement: Issues and Practice*, 20 (4), 19-27.
- Kane, M. T. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 64, 425-461.
- Linn, R. L. (2005, June). Conflicting demands of No Child Left Behind and state systems: Mixed messages about school performance. *Education Policy Analysis Archives*, 13(33). Available at <http://epaa.asu.edu/epaa/v13n33/>.
- McLaughlin, D., de Mello, V. B., Blankenship, C., Chaney, K., Hikawa, H., Rojas, D., William, P., & Wolman, M. (2005, February). *Comparison between NAEP and state mathematics assessment results: 2003. Final report Volume II: Appendix D State Profiles*. Washington, DC: American Institutes for Research.
- Sireci, S. G., Lewis, C., & Martone, A. (2006, April). *Why can't we all just get along? How psychometricians can work with school districts to improve student learning*. Presentation delivered at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

NERA RESEARCHER
Dr. Barbara J. Helms, Treasurer
P. O. Box 66114
Auburndale, MA 02466
