INSTITUTE OF EDUCATION SCIENCES

National Board for Education Sciences Meeting

June 20, 2012
Washington, DC

Summary Report

Location

Institute of Education Sciences (IES) Board Room
80 F Street NW
Washington, DC 20001

Participants

National Board for Education Sciences (NBES) Members Present
Bridget Terry Long, Ph.D., Chair
Kris D. Gutierrez, Ph.D., Vice Chair
Deborah Loewenberg Ball, Ph.D.
Anthony S. Bryk, Ed.D.
David Chard, Ph.D.
Adam Gamoran, Ph.D.
Robert Granger, Ed.D.
Margaret R. "Peggy" McLeod, Ed.D.
Judith D. Singer, Ph.D.
Robert A. Underwood, Ed.D.
Hirokazu Yoshikawa, Ph.D.

Ex-Officio Members Present

John Q. Easton, Ph.D., Director, IES, U.S. Department of Education (ED)
Elizabeth Albro, Ph.D., Acting Commissioner, National Center for Education Research (NCER), IES
Sean P. "Jack" Buckley, Ph.D., Commissioner, National Center for Education Statistics (NCES), IES
Rebecca Maynard, Ph.D., Commissioner, National Center for Education Evaluation and Regional Assistance (NCEE), IES
Peggy McCardle, Ph.D. M.P.H., Branch Chief, Child Development & Behavior Branch, Center for Research on Mothers and Children, Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), National Institutes of Health (NIH)
Chuck Pierret, U.S. Bureau of Labor Statistics
Deborah Speece, Ph.D., Commissioner, National Center for Special Education Research (NCSER), IES

NBES Staff

Monica Herk, Ph.D., Executive Director, Designated Federal Official (DFO)

IES Staff

Sue Betka
Lisa Bridges
Wai Chow
Joy Lesnick
Ellie McCutcheon

Ruth Neild
Carol O'Donnell
Allison Orechwa
Audrey Pendleton
Chris Rakes

Invited Presenters

John Hutchins, Communications Director, MDRC
Douglas Staiger, Ph.D., John French Professor of Economics, Dartmouth College
Helen Ladd, Ph.D., Edgar T. Thompson Distinguished Professor of Public Policy and Professor of Economics, Duke University
Anne Ricciuti, Ph.D., Deputy Director for Science, IES

Members of the Public

Anne Chamberlain, Social Dynamics
Kim Hymes, Council for Exceptional Children
Becca Jackson, Social Dynamics
Carla Jacobs, Lewis-Burke Associates
Angel Lopez, Jr., Council for Exceptional Children
Myrna Mandlowitz, Southwest Educational Development Laboratory
Augustine Mays, WestEd
Michele McLaughlin, Knowledge Alliance
Siobhan Mueller, Hager Sharp
Virginia Neale, Northwestern University
Sarah Sparks, EdWeek

Call to Order, Chair's Remarks

Bridget Terry Long, Ph.D., NBES Chair
Dr. Long called the meeting to order at 8:34 a.m., and Dr. Herk, Executive Director and DFO, called the roll. NBES members unanimously approved the agenda for this meeting. Dr. Long welcomed four new Board members: Dr. Gamoran, who served on the Board previously, and Drs. Chard, Singer, and Yoshikawa. Dr. Long said the new members attended a half-day orientation session on June 19, 2012. Two additional Board candidates have been approved by a Senate committee and are awaiting approval by the full Senate; if approved, they would join the Board in October. In November, several Board members will complete their terms. At a conference call in April, the Board unanimously approved the 2012 Annual Report.

Dr. Long noted that a May 2012 memo from the Office of Management and Budget (OMB) encourages all executive departments and agencies to use evidence and incorporate evaluation in program planning and budget-setting. The memo refers to the Coalition for Evidence-Based Policy, an organization headed by former NBES Chair Jon Baron. Dr. Long said that some have suggested that the memo signals the coming of age of evidence-based policymaking; she said it is clear that Washington, DC, is paying attention to concerns about evidence.

The Board generally focuses on three areas, said Dr. Long:

- Disseminating and scaling up promising practices
- IES research funding
- Advocating for the support and use of research

The agenda for the meeting reflects all three areas. In terms of dissemination, two presenters will talk about communicating research effectively to a wide range of audiences. Regarding funding, IES staff will discuss the selection and role of reviewers in evaluating grant applications. With respect to support for research, the Board will review proposed changes made by a previous Board in 2008 to the Education Sciences Reform Act (ESRA), which is up for reauthorization.

Update: Recent Developments at IES

John Q. Easton, Ph.D., IES Director

Dr. Easton welcomed the new Board members. He applauded the Board for taking a new approach to the Annual Report, working with Dr. Herk to put together a summary that represents the Board's perspective, instead of simply approving a summary written by IES staff. Regarding the IES budget, Dr. Easton said the proposed fiscal year (FY) 2013 budget represents a small overall increase for IES but includes both increases and decreases in specific programs. The biggest change would be increased spending on special education research. While nothing is settled, the current proposal is promising, said Dr. Easton.

Dr. Easton announced that Dr. Maynard will leave NCEE at the end of the summer to return to a position at the University of Pennsylvania. Dr. Maynard has made an enormous positive contribution to the use of evidence, not just to NCEE but to the entire federal government, said Dr. Easton. She has changed the culture of NCEE and encouraged cross-center collaboration in a way that IES had never done before, he continued. Dr. Maynard revitalized the Regional Education Laboratories (RELs) and made a wonderful contribution to the OMB memo mentioned by Dr. Long. She is widely considered across government as an expert on evidence use, said Dr. Easton. He congratulated Dr. Maynard and thanked her for her work. Dr. Maynard responded that she feels very good about the work that NCEE has achieved and she hoped it would continue to improve.

Commissioner Updates

National Center for Education Evaluation and Regional Assistance (NCEE)
Rebecca Maynard, Ph.D., NCEE Commissioner

NCEE is launching several new impact studies, including one on higher education and one on Pell Grant expansions under the Experimental Sites Initiative. In the latter, states are required to cooperate in an evaluation with IES in order to receive funding—a mechanism that Mr. Baron and the previous Board supported. NCEE is also evaluating the effectiveness of promising strategies for federal college access programs; extended learning time, which will be linked with waiver authorities under the 21st-Century Comprehensive Community Learning Centers Initiative; and math professional development for elementary grades. NCEE is beginning a congressionally mandated study of the District of Columbia's school choice program, as well as some feasibility studies of parent information for effective school choice and of positive behavioral interventions and supports.

NCEE has released three evaluation reports since the last NBES meeting: one on moving high-performing teachers to low-performing schools, an impact study of Title I supplemental educational services, and a descriptive study of inclusion of students with disabilities in school accountability systems. Dr. Maynard noted that these reports reflect the diversity of NCEE's evaluations; not all are impact studies.

Nine of the 10 RELs are operational (one award is being contested). NCEE held two kick-off events for the new RELs in January and plans another for REL directors in July. The nine RELs have linked with 68 research alliances, of which 30 focus on special populations, such as rural areas, English language learners, or special education. The alliances are focused on the five priority areas for ED: early childhood education, identifying and retaining effective teachers and principals, adopting and implementing rigorous academic standards, increasing college readiness and access to vocational training, and improving performance among lower-achieving students. Each alliance involves 8 to 10 schools, so the RELs are extending their reach. Dr. Maynard said the program is shaping up as intended, and she credited Dr. Neild, associate commissioner of NCEE, and the rest of the NCEE staff for the success. The RELs will produce a range of products and services, including technical assistance, tools for data collection and analysis that simultaneously build capacity, descriptive analyses of local data, exploratory studies, impact evaluations, and more. The range is wider and the RELs are more integrated with their regional partners than in previous iterations of the program.

The directors of the National Library of Education and the Education Resources Information Center (ERIC) retired in December, along with other permanent staff. NCEE is using the opportunity to develop a new 5- to 10-year vision for the programs. Efforts are underway to fill the leadership positions. Dr. Maynard said NCEE is working to make both programs even more service-oriented.

National Center for Education Research (NCER)
Elizabeth Albro, Ph.D., NCER Acting Commissioner
Dr. Albro said that in March, NCER announced 26 grant awards for research on a range of topics, and more will be announced in July. The request for applications (RFA) for 2013 was released and includes two new research programs under NCER. First, the new Research-Practitioner Partnerships in Education Research support partnerships between research institutions and state or local education agencies around student achievement issues that matter to states and districts. Together the partners will identify areas for research, develop a plan, and subsequently apply for an IES grant. The effort has received lots of interest already. Second, the new Researcher and Policymaker Training Program in the Education Sciences will support preparation for research for the kind of work IES funds and provide opportunities for researchers to share their findings and cutting-edge information with policymakers and practitioners. Both new grant programs reflect interest in partnerships that emphasize collaboration and aim to support the IES goal of funding relevant research and building capacity.

Dr. Albro noted that 90 million Americans are at or below the threshold for basic reading skills. Adult education research poses unique methodological challenges—for example, adults drop in and out of educational settings, resulting in high attrition rates. Implementation of interventions and teacher quality in adult settings are also concerns. NCER gathered a small group of experts in adult education, research methodology, and state education policy to discuss ways to improve adult education. Dr. Albro said her office is still digesting the results of the meeting and determining the next steps. She stated that the gathering represented a good example of the benefits of bringing researchers and practitioners together. Dr. Herk added that the meeting participants demonstrated much energy and excitement around the topic.

Discussion

Dr. Gamoran asked whether IES plans to post a new competition for predoctoral research training; Dr. Albro said that a decision had not yet been made. Asked by Dr. Granger whether the Centers are trying to create mechanisms that help researchers and practitioners learn from each other as they go, Dr. Albro said that NCER's RFAs on partnership do not include formal directives for grantees to learn from each other, but the Centers hold meetings regularly with grantees and ensure that the grantees are talking with each other about their work. Dr. Granger said the William T. Grant Foundation and the Spencer Foundation have funded learning communities to discuss scaling up promising practices and research-practitioner partnerships. Dr. Granger suggested the Centers think about how to capture learning as programs move forward.

Dr. Gamoran said that in the past the IES Research Conference was an opportunity for discussion among researchers and practitioners. Dr. Granger said such conferences can prompt discussion but mostly provide a platform for researchers to promote their successes. Dr. Granger hoped for a forum in which people can air their struggles and get advice from their peers, and he believed that a federal agency could play a key role. Dr. Albro noted that NCER and NCSER touch base with grantees four times a year; it may be feasible to arrange one quarterly meeting with all of the grantees who are funded through the new Researcher-Practitioner Partnership Program and encourage them to address challenges and barriers.

National Center for Special Education Research (NCSER)
Deborah Speece, Ph.D., NCSER Commissioner

Dr. Speece said NCSER has funded 28 grants so far in FY 2012. NCSER has two new initiatives in 2013. The Early Career Development and Mentoring program will assist new investigators in making the transition from postdoctoral to faculty positions. The grantees will be required to identify a new mentor (with whom he or she has not yet worked) and to reach out for expertise from other communities. Grantees will develop a training program and a research plan. Dr. Speece said the goal is to encourage new researchers in special education to increase their skills and build their research portfolios. Ten awards will be given. Dr. Speece noted that the program represents a modest investment but it is modeled on successful efforts by other agencies to bolster early career development.

A second new program, Accelerating the Academic Achievement of Students with Learning Disabilities (A3), seeks to promote rigorous development and testing of interventions for children in grades 3–8 who have or are at risk for learning disabilities. NCSER will provide awards for up to three A3 centers to develop and test both reading and math interventions. The goal is to foster intense, interdisciplinary research around interventions for children with the most intractable problems, said Dr. Speece.

NCSER is working closely with the National Institute on Disability and Rehabilitation Research, part of ED's Office of Special Education and Rehabilitative Services, to develop an initiative on improving outcomes for adolescents with disabilities. Evidence-based practices are lacking for this group, said Dr. Speece; high school graduation rates for those with disabilities and rates of entrance to postsecondary education are poor. The IES role is to develop tests and tools to help improve outcomes. The effort begins with a technical work group meeting this fall with experts and practitioners in the area of education of adolescents with disabilities to begin developing research models. Dr. Speece emphasized that single interventions will not work; a range of approaches is needed.

In late June, NCSER will host the second Summer Research Training Institute on Single-Case Design Research in Wisconsin. From a pool of 80 applicants, 40 early and mid-career researchers were selected to participate. A group of well-known experts in the field will serve as instructors for the week-long program.

Discussion

Dr. Gutierrez asked whether applicants for the NCSER A3 competition were encouraged specifically to include experts in language and discourse. Dr. Speece responded that NCSER promoted interdisciplinary research and left it to applicants to identify the makeup of the team needed to get the work done.

Dr. McCardle noted that the NIH has an upcoming RFA that complements NCSER's work; it will award grants for looking at the neurobiology and genetics of, for example, children with comorbid learning disabilities. She suggested that NICHD and NCSER stay in touch as the programs get underway. Dr. Speece agreed, noting that there has been some discussion of joint meetings of the principal investigators (PIs) for the two programs. She also suggested bringing ED's Office of Special Education Programs into the conversation, as they provide technical assistance and have a center on intensive intervention. Dr. McCardle agreed that with the complementary initiatives, the three organizations working together could produce exciting results.

National Center for Education Statistics (NCES)
Jack Buckley, Ph.D., NCES Commissioner
In May, Dr. Buckley said, NCES released the 2011 results on science performance among eighth-graders; this was a special collection of data limited to eighth-graders in support of the project to link Trends in International Mathematics and Science Study (TIMSS) data with National Assessment of Educational Progress (NAEP) data. These data also represent the first collection using NAEP's new science framework. The results showed increases since 2007 among all groups except the highest-performing eighth-graders. More recently, NCES released the results of the 2009 Interactive Computer Tasks and Hands-On Tasks assessment of students in grades 4, 8, and 12. The assessment involves more hands-on tasks and in-depth learning, reflecting the direction that other tests and standards are going (e.g., Race to the Top, Common Core State Standards, and Next Generation Science Standards).

Dr. Buckley said the Program for the International Assessment of Adult Competencies (PIAAC), which assesses cognitive and workplace skills, allows comparisons with 26 other participating jurisdictions. NCES has completed its data collection for the program and hopes results can be linked to the National Assessment of Adult Literacy. The Early Childhood Longitudinal Study for the kindergarten class of 2010–11 has been completed, as has the National Household Education Survey. The household survey transitioned from a telephone to a mail survey, which improved response significantly.

NCES is completing follow-up data collection from the High School Longitudinal Study of 2009 of ninth-graders (now eleventh-graders). The study emphasizes science, technology, engineering, and math course-taking and postsecondary education planning. It will collect data from participants through mail surveys to identify where they went to college and will collect high school transcripts for review in 2013. The Education Longitudinal Study of 2002 begins its third round of follow-up data collection in July, 8 years after the subjects graduated from high school.

The National Postsecondary Student Aid Study focuses on how students pay for college. It will serve as the base year for the Beginning Postsecondary Students Longitudinal Study. The 2008–12 Baccalaureate and Beyond Longitudinal Study examines employment and education after completing a bachelor's degree and includes a special emphasis on the experiences of new teachers. The Integrated Postsecondary Education Data System (IPEDS) will begin collecting data on enrollment in distance education this year. The IPEDS Committee on Measures of Student Success concluded its efforts this year; it offered recommendations regarding ways to give more credit to 2-year institutions for outcomes beyond graduation rates (such as transfer rates). Technical panels are discussing how to implement those recommendations and incorporate them into IPEDS.

A number of studies are under design, said Dr. Buckley:

- a congressionally mandated Fast Response Survey about the condition of public school facilities;
- the Teaching and Learning International Survey, which includes 33 countries and in which the United States will be participating for the first time;
- the Longitudinal Study of Early Adolescence, which represents the first-ever data collection from students in grades 6–8; when combined with other data collection efforts, it will provide information about the first cohort tracked all the way from kindergarten through postsecondary education;
- the National Adult Training and Education Study, a pilot test to evaluate response rates for a household study of U.S. adults; and
- the National Survey of Teachers and Principals, which is being considered as a replacement to the Schools and Staffing Survey; NCES is reengineering the survey to collect more data faster, ideally through administrative systems, and possibly using a biennial survey of teachers and principals.

Dr. Buckley noted that changes expanding eligibility for the free and reduced price school lunch program have made it more difficult to use program eligibility as a proxy for socioeconomic status. NCES is piloting some methods to work around the problem, such as collecting the catchment zones for public schools and working with the Census Bureau to compute special tabulations using the American Community Survey at the individual school level.

NCES just announced 24 new state-level grants to support the Statewide Longitudinal Data Systems (SLDS), ranging in size from $2.6 million to $5 million and totaling $97 million over the life of the grants. The grants encourage state education agencies to build or bolster their K–12 data systems and link them to information sources for early childhood, postsecondary education, and labor markets. NCES's work on the Common Education Data Standards (CEDS) continues with an upcoming meeting to discuss improving and expanding CEDS. A new effort related to CEDS involves assessing and supporting interoperability for the Race to the Top assessments.

Discussion

Dr. Buckley clarified that while SLDS grants do not require states to share their data with researchers or support research, IES urges states not to build "data mausoleums." The systems should demonstrate utility to researchers and also to the field (e.g., practitioners and school districts). The RELs are involved in disseminating data, and NCES provides technical assistance. NCES is developing new guidance to states on making data accessible to researchers, said Dr. Buckley.

Dr. Singer raised concerns that IES does not seem to be tracking the growing influence of online learning in higher education and also in K–12, including education by non-credit-granting organizations and participation of students around the world. She suggested simple surveys to provide a baseline assessment of how many students are taking some kind of online class. Dr. Buckley pointed to some data already gathered by IES showing that most students are taking at least one credit hour online.

In response to Dr. Yoshikawa, Dr. Buckley said NCES is disseminating methodological findings related to its National Household Education Survey, particularly the good news about the improvement in response rates. He acknowledged that many successful methods for improving response are very expensive. NCES uses cash incentives as one mechanism. Dr. Buckley said Congress is considering removing the ability to use cash incentives, which is a cause for concern. Dr. Bryk noted that the growth in programs and findings over the past few years at NCES is testimony to how much IES has improved in the past few years.

Communicating Research Effectively to Diverse Audiences
Updates to the What Works Clearinghouse

Rebecca Maynard, Ph.D., Commissioner, NCEE

Dr. Maynard explained that NCEE seeks to make the *What Works Clearinghouse* a central, trusted source of information about what works in education. NCEE first identifies relevant evidence, then sifts through it to find the most credible evidence and disseminates its findings through Single-Study Reviews, Intervention Reports, Practice Guides, and other tools and resources that help users review evidence critically.

Of nearly 6,500 studies reviewed, just over 5,000 were deemed ineligible for further formal review because they did not evaluate effectiveness. Dr. Maynard stressed that those studies may provide useful information that contributes to the *What Works Clearinghouse* products or complements effectiveness research, so those studies are not dismissed completely. Of the remaining 1,500 studies, nearly 1,000 did not meet the *What Works Clearinghouse's* standards of evidence for effectiveness. Thus, about 500 studies have been fully coded as meeting *What Works* standards of evidence with or without reservations. Those 500 have been translated or summarized into 143 Intervention Reports, 73 "Quick" or Single-Study Reviews, and 15 Practice Guides.

Focusing on the Intervention Reports, Dr. Maynard summarized the breakdown of evidence ratings on a scale that describes whether a given intervention is positive, negative, potentially positive or negative, or indeterminate. She pointed out that 73 percent of the evidence used to support the 143 intervention reports comes from small bodies of evidence (e.g., one study or a few small studies); the rest come from medium to large bodies of evidence (e.g., multiple studies or studies with large sample sizes).

Dr. Maynard walked Board members through the various mechanisms for searching the *What Works Clearinghouse* website. Users can look at everything the *What Works Clearinghouse* has reviewed or filter results by key words or evidence ratings, for example. The results include full citations with information about the *What Works Clearinghouse's* evidence ratings. Tools allow users to refine or revise their search strategy as they go and to export the results into commonly used software applications.

Dr. Maynard said the *What Works Clearinghouse* receives 200,000 page views per month. The Practice Guides are by far the most popular product; they can be downloaded into a printable format if desired. The most popular topics relate to reading and math, but interest in *What Works Clearinghouse* reports varies a great deal by audience. NCEE has forged communication partnerships with education and professional organizations to reach more people, and most of the organizations include a link to the *What Works Clearinghouse* on their websites.

Analysis of customer feedback reveals that users like the Find What Works search tool, the effectiveness ratings, and the details about interventions. Users suggested that the layout be improved and definitions clarified. A lot of feedback comes from inquiries to the *What Works Clearinghouse* help desk; about one quarter of those seeking help are educators—a major target of the *What Works Clearinghouse*, said Dr. Maynard.

NCEE hopes to enhance the *What Works Clearinghouse* website and make it more user-friendly. Discussions are underway about incorporating single-case design research, and NCEE is working closely with NCSER and consultants in this area. NCEE has also forged some collaborative agreements to gather cost-effectiveness data that would feed into the *What Works Clearinghouse*. NCEE is developing new evidence synthesis products. It also hopes to offer more online training and data sharing.

Communicating Research to Diverse Audiences

John Hutchins, Communications Director, MDRC

Mr. Hutchins explained that MDRC seeks to identify what works in social and education policy through evaluations and demonstration projects and to communicate its findings to policymakers, practitioners, and other influential people. MDRC is trying to address the challenges of communicating to different audiences with different objectives; one mechanism is a new website that will launch later this year.

To improve policy and practice, it is important to answer the questions that policymakers and practitioners are asking, which requires an understanding of the context, said Mr. Hutchins. Consulting frequently with the target audience(s) increases that understanding and also builds the audience for the findings, but it takes years. The findings must be made clear to smart lay people who understand policy but may not understand complex statistics. Furthermore, Mr. Hutchins said, MDRC often aims to reach multiple audiences with one document:

- Researchers who want to see all the data and who lean toward cautious interpretations
- Policymakers who focus on the bottom line, i.e., what works and how much it costs
- Practitioners who want to know why an intervention works and how to apply it
- Advocates who seek supporting evidence to build a case
- Media sources looking for stories and narratives that appeal to consumers

Mr. Hutchins described some ways MDRC addresses these challenges. In written communication, for example, products are simple without being simplistic. They should avoid jargon, and follow a clear story line. To reach multiple audiences, one product can have many layers, including a very short description on the website, a one-page overview, a five-page executive summary, a more detailed report, and appendices for the most complex supporting data. In addition, reviewers who are not researchers should read the products before publication to ensure they are clear to lay audiences.

MDRC also suggests helping researchers learn how to talk about their work with different audiences, including policymakers, practitioners, and reporters. When communicating, researchers should understand the policy and political contexts that surround their work, and developing that understanding takes time, said Mr. Hutchins. Researchers tend to limit their discussion to the confines of a given report; with support, they can learn to speak more broadly about research implications—which inform policymakers' decisions. Researchers may not be comfortable with how research findings are reported or used; by improving their own communication skills and understanding, researchers can do a better job explaining their own research.

Findings can be disseminated "wholesale"—that is, through broad communication such as websites and large e-mail lists, social media, and print mailings. They can be targeted to influencers, such as the trade press and bloggers. On the "retail side," said Mr. Hutchins, findings can be communicated to specific policymakers or to advocacy groups and membership organizations at conferences and meetings.

Mr. Hutchins noted that some other challenges further complicate communication. For example, a finding may speak to a question that was raised 5 years ago. The policy and research calendars rarely coincide, he said.

Although MDRC's current website is very text-heavy, its new site will be more visually appealing and use less text. It will be easier to navigate and provide the user multiple entry points to get to MDRC's publications. The site uses tagging and meta-data to help users find materials and enables users to filter searches by areas of interest, populations, etc. Mr. Hutchins noted that one section highlights the broader questions that researchers hope to answer, which showcases MDRC's work in a different light than typical approaches. Users will have the opportunity to comment on MDRC's work in an online public forum.

Discussion

Dr. Lesnick, who administers the *What Works Clearinghouse*, said that the *Clearinghouse* has more than 700 followers on Facebook, and interest has been rising steadily. Dr. Singer pointed out that both the *What Works Clearinghouse* and the MDRC websites are extremely text-heavy; both would benefit from including more graphical and audiovisual communication, such as podcasts and videos. The *New England Journal of Medicine's* website, for example, posts videotaped interviews with investigators when their papers are published, and those interviews become part of video news releases used by news outlets. Dr. Singer suggested that IES consider media training for researchers and consider how to distinguish good research from the clutter of sources that are available to the public. She noted that one can be so focused on getting the research right that it doesn't reach anyone (because there is so little attention to effective communication of the findings). Dr. Long suggested MDRC set up an experimental design to test the effects of allowing public comments on its new website (e.g., posting every other public comment).

Dr. Bryk questioned the overall productivity of the *What Works Clearinghouse* effort given that so few studies meet the requirements for full coding and review. Even some research produced by IES-funded centers does not meet the *Clearinghouse's* standards, which seems like a disconnect that stems from the design of the *Clearinghouse*, he said. Dr. Maynard pointed out that the *What Works Clearinghouse* assesses research on the basis of whether it demonstrates reasonable causal validity for questions of impact, but the standards are not that exacting. Many good studies fall under the category of exploratory research or are useful for generating hypotheses or evaluating predictions. The goal of the *What Works Clearinghouse* is to assess the effectiveness of interventions in practice, Dr. Maynard emphasized. She noted that the *What Works Clearinghouse* has not had a lot of pushback from researchers suggesting that their work has been mischaracterized. She acknowledged that the assessment of effectiveness is imperfect, but when similar findings from multiple studies or multiple sites are available, those results are bundled into the intervention reports.

Dr. Ball pointed out that policymakers and practitioners draw conclusions on the basis of a range of experiences, not just from research findings, and in that sense, the premise of the *What Works Clearinghouse* is the beginning of the disconnect that Dr. Bryk described. She wondered how IES is helping people make better informed decisions that take the available evidence into account. Dr. Ball expressed concern that the education research community is not improving the quality of public decisionmaking by encouraging people to use what is known. She said the links are missing between good, worthwhile research and the application of research to decisionmaking. The field could be more effective if it considered more closely "the chain of public reasoning," said Dr. Ball.

Dr. Buckley said that unique approaches to communication can be effective, even those that appear terrible on the surface. For example, instead of writing a report and organizing a press conference for its annual *Condition of Education* report to Congress and the White House, NCES made a 7-minute animated video that was enthusiastically received. That video is available online and garnered better media coverage than the in-depth press briefings NCES had used in the past. The online video can be accessed at http://nces.ed.gov/programs/coe/slides.asp. Dr. Buckley said he was even more skeptical about using Twitter to disseminate news about NAEP releases, but the real-time dialog that Twitter offers allows NCES to answer questions and clarify information even as it is being disseminated during a press conference.

Dr. Yoshikawa advocated new ways of thinking about the kind of assistance needed in communicating findings, especially by grantees who conduct high-impact research. Media trainings and workshops for researchers typically do not involve communication staff from their institutions. Assistance should involve the communications staff at the grantee's home institution.

Dr. Granger pointed out that what is known about what works is limited—as evidenced by the fact that so many of the *What Works Clearinghouse's* intervention reports rely on a small body of evidence. He expressed some ambivalence about promoting research findings as the basis for changing policy when the findings are limited. Along the same lines, providing media training to scientific researchers to encourage them to talk about the findings of a single study flies in the face of the standard research paradigm, said Dr. Granger. He cautioned against trying to steer policy and practice on the basis of limited information. The Practice Guides are quite transparent, added Dr. Granger, in that they clearly spell out what is known and what is not known about a topic from research.

Dr. Gamoran agreed that the narrowness of the findings in the *What Works Clearinghouse* must be kept in mind. However, the research has progressed to the point that the *What Works Clearinghouse* can now identify some interventions that do work. It has also had a substantial impact on understanding what research can tell us about certain circumstances and conditions. The *What Works Clearinghouse* reflects the movement toward a larger body of research that can lead to causal conclusions, said Dr. Gamoran, and he hoped that thread would not be lost in a focus on disseminating findings.

Dr. Gamoran said the Practice Guides are valuable to both researchers and practitioners; he asked why the *What Works Clearinghouse* did not produce more of them. Dr. Maynard said many new Practice Guides are in development, but they focus only on areas where there is enough evidence to produce a meaningful guide. The Practice Guides are expensive, she added. To improve the overall effectiveness of the *What Works Clearinghouse*, Dr. Maynard said, the Quick Reviews are now produced much more quickly. Also, the coding process used in the overall review of effectiveness is being used to inform Single-Study Reviews. The results of the coding exercise are now part of the database and can be used by the public to better understand the process. Dr. Maynard noted that if researchers apply the *What Works Clearinghouse* coding protocol to their own work, NCEE is willing to include the results in the database. Thus, better reporting procedures improve the likelihood that a study will be included and decreases NCEE staff effort.

Dr. Maynard stated that the *What Works Clearinghouse* does produce "null" reviews—that is, summaries of areas in which none of the available evidence meets the *What Works Clearinghouse* standards. Null reviews signal that a topic has been addressed.

Dr. Chard agreed that the *What Works Clearinghouse* has contributed to the focus on evidence-based thinking, and PIs now consider whether their research meets the *What Works Clearinghouse* criteria. However, competitors such as Edutopia are more nimble and update their websites more frequently. Dr. Chard proposed that IES act as the communication arm for grantees or provide technical guidance to research institutions about communicating their own work and making sure that they attribute their findings to IES funding. In other words, the grantees and their respective institutions become a distributed network for communicating about IES.

Dr. Long concluded the discussion by saying that the *What Works Clearinghouse* is improving, and Dr. Maynard has addressed many of the challenges it faces. She appreciated that both the *What Works Clearinghouse* and MDRC are working to provide users with more ways to find information and both are thinking about reaching different audiences. Dr. Long suggested that adding more graphics and color to the *What Works Clearinghouse* could make it more user-friendly. She asked for more details at a later time about who uses the *What Works Clearinghouse* website and how. Dr. Long reiterated the need to reach beyond the website by using social media and other mechanisms to promote research findings. Regarding the competition for the public's attention, Dr. Long emphasized the importance of IES distinguishing its contributions by clarifying what constitutes good research and delineating what is known and what is not known.

Recent Research on Instructional Quality

Interpreting Volatility in Annual Performance Measures of Teacher Effectiveness

Douglas Staiger, Ph.D., John French Professor of Economics, Dartmouth College

Dr. Staiger noted that volatility—that is, dramatic year-to-year changes in results—is an issue for every measure of teacher effectiveness, which raises questions about the reliability of such measures for decisionmaking, particularly high-stakes personnel decisions. To discuss his perspective, Dr. Staiger focused specifically on teacher value-added but stressed that the same questions about reliability apply to every measure of teacher effectiveness (e.g., classroom observation scores and student perceptions). He noted that all performance measures are notoriously volatile.

Even unreliable measures can identify substantial differences, Dr. Staiger noted. For example, estimates of value-added from one year or one class are predictive of value-added across teachers in different years and different classes. Dr. Staiger and a colleague published a controversial paper pointing out that if a decisionmaker relied solely on year-to-year data on teacher value-added to maximize student test scores, the decisionmaker would fire 80 percent of teachers every year. The paper was not advocating the actual use of such a decision rule but simply providing a thought experiment demonstrating that even an unreliable measure can provide useful information.

In one study based on Los Angeles data, Dr. Staiger ranked teachers into quartiles based on their value-added scores in their first 2 years of teaching. He then looked at the teachers' value-added in their third year of teaching: teachers ranked in the top quartile during their first 2 years of teaching were 10 percentage points higher in their value-added in their third year of teaching compared to teachers ranked in the bottom quartile.

The Gates Foundation's Measure of Effective Teaching (MET) study ranked teachers on a composite score (value added, observation, and student perceptions) in different classrooms. The MET composite score accurately predicted large differences in future value-added, especially when compared to cruder measures of teacher effectiveness such as teacher credentials or experience, and those are typical findings. Researchers can achieve even greater predictive accuracy if they weight the value-added measure more heavily, he said.

Dr. Staiger explained how he and his colleagues examined the link between teachers' annual value-added scores and their career average scores. Their model found a fairly strong (about 0.65) correlation between a teacher's 1-year value-added score and their career average value-added using data from at least 6 years of teaching. A correlation of this magnitude has predictive power and reduces the risk of misclassifying teachers. Dr. Staiger conceded that some high performers and low performers change places over time, but some volatility and unpredictability are unavoidable. Year-to-year comparisons of value-added scores are volatile, he said, but a given year is a strong predictor of career average scores—as good as any other measure in use now.

Assessment of teacher effectiveness could be improved by combining multiple measures across years and in different classrooms, Dr. Staiger emphasized. The best predictor of underlying performance is past performance, so accumulated data are important. Dr. Staiger noted that a lot of performance data are available, but little effort has been made to understand how to apply the data to decisionmaking. He encouraged the use of more cumulative data in teacher performance assessment, perhaps with more weight given to recent measures.

Value-Added Modeling of Teacher Credentials

Helen Ladd, Ph.D., Edgar T. Thompson Distinguished Professor of Public Policy and Professor of Economics, Duke University

Dr. Ladd focused on the relationship between teacher credentials and student achievement. She presented data gathered from North Carolina on teachers in grades 4–5 and in high school. North Carolina has a large body of data, including from high schools, which require end-of-course tests that are linked to the state's standard course of study, count for 25 percent of a student's course grade, and serve as the foundation for the school-based accountability system. Thus, teachers have a strong incentive to teach the material covered by the tests, and students have a strong incentive to learn it.

Dr. Ladd noted that all of the results she presented are expressed as standard deviations, and the findings are statistically significant (unless otherwise noted) and come from huge data sets with careful attention to causal modeling. Examining the effects of teacher experience on student achievement, Dr. Ladd noted a big jump in teacher effectiveness after the first or second year of teaching, with steady increases as the teacher's career continues. Among elementary school students, the effect is stronger in math than in reading.

The same pattern applies to high school teachers, with big jumps in effectiveness in the first 1 to 2 years of teaching, with increases leveling off to very small effects after 3 to 5 years. Some might interpret the findings from the high school data as evidence that teachers do not continue to learn on the job. Dr. Ladd believes the change may indicate that teachers move away from teaching low-level math and reading classes after a few years (since Dr. Ladd's data set is limited to teachers who teach the entry-level math and reading classes). When teacher fixed effects are added, said Dr. Ladd, there is some evidence that teachers continue to learn throughout their careers.

Dr. Ladd's data indicate that licensure in a form other than the initial/regular teaching license is associated with lower student achievement. National Board certification, however, correlates with increases in student achievement. The latter findings led Dr. Ladd to ask whether the National Board of Professional Teaching Standards identifies more effective teachers (i.e., a signaling effect) or whether the certification process improves the teacher's abilities. She concluded that at the elementary school level, the signaling effect seems to be primary, as the greatest impact on student achievement is evident 2 years before teachers completed the National Board certification process. At the high school level, however, the National Board certification process itself seems to lead to improved teacher performance.

Most evidence suggests that obtaining a master's degree does not translate into student achievement. Dr. Ladd said her data showed that students with elementary school teachers who obtained their master's degrees 5 years or more into their teaching careers had lower test scores than students whose teachers did not. She believed the findings may indicate that weaker teachers seek a graduate degree in order to raise their salaries. She did find a small but significant positive correlation at the high school level between whether teachers obtain a master's degree more than 5 years into their teaching career and their students' academic performance.

Looking at teacher licensure test scores, Dr. Ladd reported positive correlations with achievement at the elementary level and smaller positive correlations with high school test scores. At the high school level, she noted, data are available on licensure scores by topic. (In North Carolina, some teachers have topic-specific certifications.) Teachers' scores on algebra and geometry correlated with the largest positive correlations with student achievement. Dr. Ladd could not explain why teachers' English scores had an unexpectedly negative correlation with student achievement; she thought perhaps the findings represent teachers who teach courses of less interest to them, such as basic English. Teacher certification in math (algebra and geometry) or English has the largest positive correlation with student achievement.

Dr. Ladd defended the small magnitude of the correlations in the data by saying that credentials represent a bundle of characteristics. When she and her colleagues looked at the combination of credentials that elementary school teachers had and characterized them according to whether the teacher had a strong or weak set of credentials, they found modestly large relationships between teacher credentials and student achievement and that the relationship was larger for math than reading. For high school teachers, Dr. Ladd and her colleagues identified the 90th and the 10th percentile of teachers in terms of their overall set of credentials and again found a relatively strong relationship between teacher credentials and student achievement.

Dr. Ladd summarized the policy implications of her findings. She noted that if one accepts the premise that teacher credentials on average matter, it is possible to look more closely at the distribution of teachers across schools than one can with value-added measures. Dr. Ladd and colleagues determined that teachers are currently maldistributed across schools and classrooms in terms of credentials, which contributes to student achievement gaps. Schools serving the most disadvantaged students have the weakest teachers in terms of credentials, said Dr. Ladd. In North Carolina, the distribution differences increase over time, she added. Thus, Dr. Ladd recommended evening out the distribution in terms of credentials of teachers across schools. More research is needed on how to improve the distribution of teachers beyond existing information concerning the effect of salary differences and bonuses. Dr. Ladd also suggested focusing on teachers' perception of leadership and working conditions in the schools where they work.

Further, school-level accountability is affected by the maldistribution of teachers. Dr. Ladd said she worries about the pressure to perform put on schools that have many teachers with weak credentials, high teacher turnover, and limited capacity to respond to these challenges. Some evidence shows that the pressure for school-level accountability increases turnover in low-performing schools. Moreover, because No Child Left Behind holds elementary schools accountable only for student performance in grades 3–5 (since these are the tested grades), principals have incentives to move better teachers into these grades. However, Dr. Ladd said, what happens in the early grade classrooms is important.

Finally, Dr. Ladd suggested rethinking the salary schedule for teachers, particularly raising salaries at the bottom level to attract candidates with higher abilities. Increasing salaries more quickly at the bottom levels should also be considered, because teachers no longer stay in the field for life. Salary increases should encourage good teachers to stay in the field, and salaries should be more consistent with patterns recognized in other professions.

Dr. Ladd concluded that it is important to identify good teachers, but she believes most principals and colleagues already know who the good teachers are. She suggested identifying both strong and weak teachers. She also recommended paying teachers for taking on additional work, such as mentoring—a mechanism that teachers would see as more fair than paying for performance.

Discussion

Dr. Ladd noted that she and her colleagues have compared their North Carolina data with that of Florida, where the correlation of teacher credentials with student performance has not been as strong on some dimensions, such as National Board Certification. However, there are meaningful differences between states. Dr. Ladd said data must be gathered within states, but models are also needed to facilitate comparison across states.

Dr. Gamoran pointed out that unexpected negative correlations of teacher credentials with student performance were interpreted by Dr. Ladd as the result of selection bias, but positive correlations were attributed to treatment (i.e., credentials). He wondered whether the data showed effects across or within schools. Dr. Ladd responded that there is lot of variation across schools. She noted that since the end of the court-ordered desegregation agreement, schools in Charlotte, North Carolina, are more segregated than ever, and it takes a concerted effort by principals to counteract the effects. She also noted that she and her colleagues aim to control for problems related to potential correlations between unobservable teacher and student characteristics, but it remains difficult to identify why teachers either leave the field or choose to pursue credentials.

Dr. Ball commented that in Dr. Ladd's work certification seems to be a proxy for teacher professional development. She asked why Dr. Ladd had not made any recommendations concerning improving professional training. Dr. Ladd said the next step is to focus on states that do not have good training programs and emphasize that credentials do matter.

Regarding the use of student performance data, Dr. Staiger emphasized that the measures selected must correspond with the area of interest. If the focus is on teachers' impact on state test scores, then measures of value-added are appropriate. If the focus is on student perceptions of teacher effectiveness, student survey results are the best measure. Dr. Staiger advocated creating composite scores from each year of data for a given teacher and using those data to predict performance over the next year. Adding in new data each year shows the progression of performance and improves the predictive power of the findings.

Dr. Bryk said it is important to distinguish the use of value-added scores for policy analysis versus decisions about individual teachers. The lack of reliability within the measures suggests that only at the extreme ends of the spectrum can value-added scores reliably be associated with changes in student performance. Clearly, a small percentage of teachers are either very good, or very poor, and most fall somewhere in the middle. Within that middle ground, individual performance is indistinguishable. Dr. Staiger agreed but also noted that the data do have the power to predict performance on average and they are the best data available at the moment. He emphasized that if policymakers' only lever is laying off teachers, then using this rule would lead to massive numbers of firings, which is not helpful. Because the cost to students of keeping an ineffective teacher in the classroom is high, however, an aggressive approach to removing ineffective teachers is optimal, regardless of the method used to assess teacher performance. The positive corollary is that the payoff for investing in keeping good teachers is also very high. The question is how to identify good and poor future teachers before individuals invest heavily in pursuing teaching as a career, said Dr. Staiger.

Dr. Singer said neither presenter mentioned tenure, which underlies the policy discussion. For teachers without tenure, administrators must make high-stakes decisions early in a teacher's career when few data are available and based on criteria that seem to prohibit using other relevant information. In universities, tenure decisions are not made on the basis of a single measure of effectiveness (e.g., citation counts), but that seems to be the goal for some proponents of value-added measures to assess K–12 teachers, said Dr. Singer. She noted that she is skeptical about the utility of value-added as the sole criterion for decisionmaking, especially early in teachers' careers. Dr. Staiger responded that one of the surprises of his research was that decisionmakers do have enough information early on to distinguish the best and worst teachers, and if decisionmakers are actually going to decide to remove some teachers, they should make that call pretty early.

Dr. Staiger said that after 1 year, enough data are available to identify low-performing teachers and consider dismissing the worst, even if that decision were wrong 5 to 10 percent of the time. Dr. Bryk objected, saying the data suggest that the assessment is wrong 40 percent of the time; he asked whether anyone would invest in pursuing a career if they had a 40-percent chance of being fired in the first year. Dr. Staiger said that in most careers turnover within the first 6 months is very high. The problem is that people invest as much as $30,000 in being trained to be a teacher even before they start teaching.

Dr. Granger suggested bringing more practitioner or local policymaker perspectives into the discussion of how to design systems to improve teacher performance. Many states and districts are investing a lot in changing their human resource practices regarding teachers, and it would be good to understand what they believe they have learned.

Dr. Long asked what IES could do to translate and communicate the findings of Dr. Ladd's research. Dr. Ladd said the Center for Analysis for Longitudinal Data in Education Research (CALDER) and Duke University's communication staff are working to disseminate her findings, and she works closely with the press office at Duke to ensure that her studies are described accurately. She noted that Duke does not have much direct contact with legislators. Dr. Long said IES invests a lot in research, but opportunities to promote the findings are easily lost. Dr. Staiger said the IES grant to study teacher effectiveness included a requirement for dissemination; thus, he and his colleagues have had meetings with multiple school districts. The meetings often focus on methodology, but there is some discussion about interpretation and application. Dr. Long added that communication and dissemination both require money and staff with special expertise.

Dr. Ball wondered how to frame the findings of research in ways that help policymakers make good, informed decisions. She noted that researchers can sometimes say too little about the implications of their findings, and the message gets lost. Learning the boundary between going too far and saying too little and how to occupy the ground in between would be a smart approach, she said. Dr. Ladd said she does try to show how data can inform policy decisions, especially in highly controversial contexts; for example, one paper concluded that the more racially segregated a school is, the higher the salary needed to attract and retain teachers.

Dr. Staiger said combining multiple measures to assess teacher effectiveness may be one mechanism that policymakers can use. When good teacher performance data are not available, policymakers should look closely at what they are trying to achieve and evaluate data that are pertinent to their concerns, he concluded.

Lunch

Participants adjourned for lunch at 12:16 p.m., and the meeting resumed at 2:16 p.m. During the lunch break, NBES members traveled to the office of the ED Secretary, and new members were sworn in.

IES's Peer Review Process: Review Panel Criteria, Recruitment, and Training

Anne Ricciuti, Ph.D., IES Deputy Director for Science

Dr. Long framed the presentation by saying that the work that IES funds reflects the quality of the processes for receiving, reviewing, and funding grants. As the Board considers how IES can improve its research portfolio, Board members should consider those processes.

Dr. Ricciuti noted that IES recently completed the FY 2012 review cycle, processing more than 1,300 grant applications across the Centers. The applications were reviewed by about 370 reviewers across 30 panels. IES currently has 8 standing review panels that have principal members, rotating members, and occasionally ad hoc members.

Planning for grant review is an ongoing effort, said Dr. Ricciuti. IES holds two review cycles every fiscal year (in October and March). After every review cycle, IES conducts in-depth debriefings about the review panels to assess the process and substance of the review. Staff members assess the need for reviewers by considering who is rotating off panels, what competitions are coming up, and what other needs may arise during the year.

For standing panels, principal members are appointed to 3-year terms following approval by Dr. Ricciuti and the IES Director. Rotating and ad hoc reviewers only require approval by Dr. Ricciuti. IES staff are continuously assessing and refining how many reviewers are needed. Most principal members on standing panels have participated on an IES review panel in the past. IES considers not only a potential member's research expertise but also the person's ability to work well with others in the context of a review panel, to make considered judgments, and to listen to others' opinions. Panel chairs are chosen from among the panel members and play an important role. Dr. Ricciuti said that individuals who are invited to serve as panel chairs typically have experience with IES review panels and experience in leadership roles.

In response to previous Board suggestions about raising the level of prestige associated with serving on an IES review panel, the invitation to principal members and the appointment letters are now signed by the IES Director. Dr. Ricciuti welcomed other suggestions on recruiting top-quality reviewers. She provided the Board with a list of the FY 2012 reviewers, their affiliations, and the fields in which they received their graduate training in order to help Board members get a broad sense of the areas of expertise represented.

Dr. Ricciuti said IES recognizes the need to improve panel diversity. She is looking into ways to consider race/ethnicity as one aspect of demographic diversity, as well as exploring the ways in which other agencies do so. NIH collects race/ethnicity data from its reviewers on a voluntary basis but does not make that information publicly available.

Reviewers, including panel chairs, receive a reviewer handbook from IES. Chairs meet with IES staff the evening before a review begins and with all reviewers the morning of the review to answer questions and discuss information specific to a given review. Reviewers are instructed to evaluate applications according to the relevant RFA(s). Dr. Ricciuti noted that IES includes a review criterion related to significance, and within that criterion, reviewers can weigh the pros and cons of more risky research versus safe, conventional topics.

The reviewer handbook is revised annually and incorporates suggestions and lessons learned from the previous review cycles. In response to a suggestion from Dr. Long, the next handbook will include more tips for panel chairs on how to manage panel discussion.

Dr. Ricciuti emphasized that the review process undergoes continuous improvement, and she is open to feedback. She added that IES is working on a system that would allow applicants to log into an online system to see the status of their application. She noted that IES is watching the new NIH approach in which applicants are limited to one resubmission of a given application, and whether NIH likes this new approach.

Discussion

Dr. McCardle said that at NIH there is disgruntlement about the limitation on resubmissions. In her opinion, the process seemed to work better when NIH applicants were limited to two resubmissions. She feels that the move to one resubmission for budget reasons has made the process more difficult for researchers, reviewers, and program staff. Many at NIH feel the move was a mistake, said Dr. McCardle; she advised IES not to clamp down too tightly at first on resubmissions.

Dr. Gutierrez said there is a growing body of design-based intervention research, some of which focuses on special populations. She asked how the review process can be nimble enough to consider new fields by recruiting and training reviewers who understand them. Dr. Ricciuti said IES tries to include reviewers with relevant expertise on its panels; recommendations from Board members and others for potential expert reviewers are very helpful. Board members and their colleagues can also help by encouraging their peers to participate in panels when invited.

Dr. Bryk emphasized that for grants that encourage innovation in particular, reviewers must have enough understanding of the field to assess whether the proposal has merit. Dr. Maynard said recent improvements in the review process should help.

In response to several questions posed by Dr. Gamoran, Dr. Ricciuti said it is difficult to evaluate the stability of panels, because one panel may have multiple sections. Between 50 percent and 75 percent of reviewers participated on the same panel in the previous year, and there is a tradeoff between including a mix of new reviewers and maintaining stability. Dr. Easton added that IES sometimes moves reviewers to different panels, which may contribute to the appearance of instability. IES does consider the characteristics of revised and resubmitted proposals, said Dr. Ricciuti.

Dr. Gamoran suggested that publishing the list of reviewers publicly might raise prestige. Dr. Ricciuti said IES publishes a list of all the reviewers after the review cycle; the list includes reviewers' affiliations, but does not identify the panels on which they served. Dr. Singer suggested at least organizing the list by field or panel. She also urged IES to collect demographic data as well as more detail about potential reviewers' work and reputation in their fields.

Dr. Singer asked whether IES promotes the value of serving as a reviewer as a contribution to the field that is recognized by peers. Dr. Ricciuti noted that IES does work hard to recruit people at the top of their fields. She added that sometimes people who are not as well known or well established turn out to be very good as reviewers. She reiterated her call for recommendations from the Board of people who should be invited to serve on panels.

Dr. McCardle pointed out that everyone who serves on NIH Special Emphasis Panels, which tend to be small, is listed on a single roster, which prevents anyone from attributing comments directly to a specific reviewer. NIH assures applicants that the panel reviewing their application includes the appropriate expertise. Dr. McCardle said that NIH reviewers only make recommendations on the basis of an application's scientific and technical merits. NIH incorporates those recommendations into its funding decisions, which occur later in the process than they do for IES. Finally, Dr. McCardle requested that the Board review the process for an applicant to appeal when an application is rejected.

Dr. Long noted that in many other academic peer review processes, the names of the review panel members are made public. While there are always tradeoffs, Dr. Long believed that identifying the reviewers does contribute to prestige. Without knowing who serves on which panel, the Board and others cannot identify the gaps in expertise that should be filled. She suggested that IES can find ways to make the names of the reviewers and the panels on which they serve public without damaging confidentiality. Dr. Long also recommended collecting not only basic demographic data on panel members but also, for IES internal evaluation purposes, information that provides some insight into a member's career and standing in the field.

Dr. Chard noted that review panels tend to be highly critical, and Dr. Long added that junior faculty sometimes try to demonstrate their knowledge by being overly critical. Moreover, Dr. Chard said, panel members sometimes do not work well together. He asked whether reviewers get any kind of feedback about their performance. Dr. Chard noted that reviewers should come to the table prepared to contribute, not act as an obstacle to the process.

Dr. Ricciuti was intrigued by the concept of providing feedback to reviewers, which IES does not currently do systematically. She asked the Board for suggestions on what kind of feedback would be useful and taken to heart by the panel members. Dr. Chard said reviewers have a difficult role: they should look at applications critically but also see the big picture.

Dr. Bryk said that in the case of resubmissions, applications should be judged according to their responsiveness to the initial criticisms and comments. Dr. Ricciuti said IES assigns resubmissions to the same reviewers whenever possible; subsequent reviewers always receive the comments of the previous panel review and, as part of their review, are asked to assess whether the applicant addressed the initial comments. However, she stated that the process is imperfect. Dr. Bryk said the success rate of resubmissions is an indicator of the quality of reviewers' comments. Dr. Albro noted that program and research officers have processes in place to work with applicants on resubmissions.

Discussion turned to early career awards. Dr. McCardle said NIH offers awards for new and early-career investigators. Dr. Albro said NCER programs include a nice mix of experienced and junior researchers.

Returning to the discussion of providing feedback to reviewers, Dr. McCardle said that in general presentations to potential reviewers, program officers address the problem of new reviewers being hypercritical and emphasize the need to keep an eye on the big picture. Dr. Ricciuti added that IES emphasizes the importance of a balanced review in its instructions to reviewers and plenary meetings for reviewers. Dr. Long wondered whether the criteria explicitly instruct reviewers to think about innovation; Dr. Ricciuti and Dr. Albro indicated that the development and innovation goal specifically states the focus on supporting innovation. Dr. Long was concerned about how reviewers look at proposals that truly push the frontier.

Dr. Singer pointed to literature on group decisionmaking, including research specifically on the peer review process. She noted a recent study in which the willingness of judges to grant parole correlated with lunch and snack breaks. Also, in that study the first applications receive the most discussion, and the agenda is often in the hands of the chair, said Dr. Singer. She and Dr. Long agreed that experiments could be designed to assess the review process, such as re-ordering the agenda. Dr. Singer suggested that Board members volunteer to observe and analyze the review panel process, offer constructive criticism, and give feedback to reviewers.

Dr. Ricciuti noted that staff of the Standards and Review Office observe panels and, unlike program officers, can speak up if the process seems to be amiss. Staff may speak to the chair or address the entire panel. Occasionally, the chair talks privately with an individual when an issue arises. However, IES has not established a formal system to let reviewers know how they performed. Dr. Long noted that only a small percentage of reviewers may need feedback, and IES could address these issues when reviewers are rotating or ad hoc members and before they are made principal members of standing panels.

The Board discussed its role in advising on the review process. Dr. Ricciuti said she and her staff would consider the Board's suggestions and concerns and work with Dr. Easton on how best to address them. Dr. Easton said some analysis of the review process would be helpful.

Dr. Ricciuti said that IES is considering whether and what kind of additional training for panel chairs might be helpful, including inviting strong, experienced chairs to provide advice and training sessions for new chairs. Dr. Ball said most universities offer some guidance to new chairs, and Michigan has some useful information that she can share. Dr. McCardle said NIH has a training process for chairs and can extend the terms of chairs when it wants to retain good leaders. NIH also appoints interim chairs, which presents a training opportunity. In addition to receiving written training materials, new chairs talk by phone with NIH staff and meet briefly with a review officer.

Dr. Ricciuti clarified that new chairs of IES panels do receive written materials that identify the duties of the chair and meet with her and other Standards and Review Office staff before the review begins. Dr. McCardle said it appears that IES has a good process in place, and codifying and standardizing the current process would probably be sufficient to address the questions raised by the Board. Dr. Long suggested that the agenda for the next NBES meeting (October 2012) include a follow-up discussion of reviewer recruitment and retention.

Legislative and Policy Issues

NBES Roundtable Discussion of the ESRA Markup by the May 2008 Board and the Current Board's Consensus

Bridget Terry Long, Ph.D., NBES Chair

At the February 2012 NBES meeting, Dr. Granger explained that in 2008, the Board marked up the ESRA legislation that was eligible for reauthorization that year; that markup is available online (link at bottom of webpage). No action has been taken on ESRA reauthorization, and the legislation is still eligible for reauthorization. The current Board members agreed to review the earlier Board's mark-ups for ESRA from May 2008. The markups were summarized by Dr. Herk in a memo to members. The Board accepted most of the suggested revisions. Additional suggestions and the rationale behind them are as follows:

- Among the groups from whom the President must solicit suggestions when considering individuals to serve on the Board, add the National Academy of Education.
- To ensure that the Board always has a sufficient number of members, automatically extend by 1 year the terms of Board members whose successors have not yet been appointed. Dr. Herk will check whether the recommended change is consistent with the Federal Advisory Committee Act.
- Board members favored revising ESRA to give the Board hiring and evaluation authority over the NBES Executive Director. Dr. Herk will check with legal counsel whether an advisory board can supervise a federal employee. Dr. Buckley noted that National Assessment Governing Board (NAGB) does appear to hold hiring and evaluation authority over its executive director.
- The Board agreed to retain the language allowing it to accept charitable contributions when Dr. Buckley pointed out that federal agencies can no longer pay to provide coffee during advisory board meetings.
- In the context of adjusting the circumstances under which the IES Director or Board members may be removed, the Board agreed to delete the suggested language regarding the appointment of the NCES Commissioner because the Board could not reach consensus.
- While there was support for expanding the language of ESRA so that a broader range of educational records can be incorporated into the SLDS, Dr. Buckley asked for time to review the suggestion with legal staff to ensure that it does not result in unintended consequences. He said the Higher Education Act includes some restrictive provisions on researchers' access to data.

- Regarding authority over the content and release of NAEP reports, Dr. Buckley explained that he supports clearer demarcation of the roles of NAGB, the NCES Commissioner, and IES. Currently, the NCES Commissioner is responsible for releasing the NAEP data, but NAGB controls the format and timing of the release, which can affect the release, said Dr. Buckley. There was some discussion about giving the NCES Commissioner the authority to set the format and timing of release of NAEP reports, but ultimately the Board agreed that it could cause NAGB to resist the original recommendation.

The Board made no recommendation regarding the suggested change to the appointment process for the NCES Commissioner, because opinion was evenly divided among the members. Some agreed that the procedure for appointing the NCES Commissioner should be the same as that for the other IES Center commissioners to support smooth and efficient functioning of IES. Others felt strongly that the current requirement that the NCES Commissioner be appointed by the President and confirmed by the Senate reflects a hard-won acknowledgment that education statistics deserve national-level prestige.

Dr. Easton said that the relationship between IES and its Centers has not always been as smooth as it is now; equalizing the commissioner appointment process could help IES facilitate more work across Centers and minimize the silo effect. Dr. Buckley pointed out that, across the federal government, statistical agencies function in unique ways; he also noted that the current appointment process does not provide any guarantee of job security to the NCES Commissioner. Drs. Buckley and Easton both noted that when Congress is divided, appointments can languish for a long time while positions are filled by "acting" staff. Dr. Singer said the prestige of a Presidential appointment may be countered by the negative opinion generated when an acting director fills a position for a long time.

Dr. Long will revise the ESRA markup according to the Board's suggestions and circulate the revised version to Board members for consideration and additional comments. When the markup is completed, Dr. Long will present it along with the FY 2011–12 NBES Annual Report to ED Secretary Duncan and Congress.

NBES Roundtable Discussion of the Draft Scientific Integrity Policy for ED

John Q. Easton, Ph.D., IES Director

Dr. Easton noted that, by Executive Order, all agencies are required to have a scientific integrity policy. IES was tasked with writing the policy for ED. The draft policy is open for public comment. The Board voted unanimously in support of the draft Scientific Integrity Policy as written.

Closing Remarks and Adjournment

John Q. Easton, Ph.D., IES Director, and Bridget Terry Long, Ph.D., NBES Chair

Dr. Easton thanked the Board members for a provocative and productive discussion. Dr. Long looked forward to the October 5, 2012 Board meeting and invited members to provide input for its agenda. Dr. Herk thanked the staff of AFYA, Inc., the meeting contractor that handled the logistics of this meeting, and Wilma Greene, IES Management/Program Analyst, the liaison to the contractor, for all their hard work. Dr. Long adjourned the meeting at 4:35 p.m.