



National Board for Education Sciences Meeting

October 14, 2011
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

Jonathan Baron, J.D., Chair
Bridget Terry Long, Ph.D., Vice Chair
Deborah Loewenberg Ball, Ph.D. (via telephone)
Anthony Bryk, Ed.D.
Adam Gamoran, Ph.D.
Robert Granger, Ed.D.
Kris D. Gutierrez, Ph.D.
Margaret (Peggy) R. McLeod, Ed.D.
Sally E. Shaywitz, M.D.
Robert A. Underwood, Ed.D.

NBES Members Absent

Frank Philip Handy, M.B.A.

Ex-Officio Members Present

John Q. Easton, Ph.D., Director, Institute of Education Sciences (IES)
Elizabeth Albro, Ph.D., Acting Commissioner, National Center for Education Research (NCER)
Alison Aughinbaugh, Ph.D., Research Economist, Office of Employment and Unemployment Statistics, Division of National Longitudinal Surveys, Bureau of Labor Statistics (BLS)
Sean P. "Jack" Buckley, Ph.D., Commissioner, National Center for Education Statistics (NCES)
Rebecca Maynard, Ph.D., Commissioner, National Center for Education Evaluation and Regional Assistance (NCEE)
Peggy McCardle, Ph.D. M.P.H., Branch Chief, Child Development & Behavior Branch (CDBB), Center for Research on Mothers and Children (CRM), Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)
Joan Ferrini-Mundy, Ph.D., National Science Foundation (NSF)
Deborah Speece, Ph.D., Commissioner, National Center for Special Education Research (NCSER)

NBES Staff

Monica Herk, Ph.D., Executive Director
Mary Grace Lucier, Designated Federal Official

IES Staff

Lisa Bridges, Ph.D.

Sue Betka

Allison Orechwa, Ph.D.

Anne Ricciuti, Ph.D.

Allen Ruby, Ph.D.

Invited Presenters

Gilbert Botvin, Ph.D., Weill Cornell Medical College

Deborah Gorman-Smith, Ph.D., University of Chicago, Society for Prevention Research

Saskia Levy Thompson, New York City Department of Education

Kathy Stack, Office of Management and Budget (OMB)

Carl Wieman, Ph.D., White House Office of Science and Technology Policy, Committee on Science, Technology, Engineering and Math Education (CoSTEM)

Members of the Public

Melanie Bateman, American Association for the Advancement of Science (AAAS)

Kimberly Broman, Coalition for Evidence-Based Policy

Teresa Duncan, ICF International

Adam Fine, Society for Research in Child Development

Steven Glazerman, Mathematica

Carla Jacobs, Lewis-Burke Associates

Judy Johnson, Reading Recovery Council

Jim Kohlmoos, Knowledge Alliance

Augustus Mays, WestEd

Eileen Parsons, AAAS

Gerald Sroufe, American Education Research Association (AERA)

Meeting Summary

Call to Order, Approval of Agenda, Chair's Remarks

Jon Baron, J.D., NBES Chair

Mr. Baron called the meeting to order at 8:35 a.m., and Ms. Lucier, Designated Federal Official for NBES, called the roll. NBES members unanimously approved the summary of the June 29, 2011 NBES meeting with no changes. NBES members also unanimously approved the agenda for this meeting.

Mr. Baron congratulated Dr. Bryk on being reconfirmed by the Senate as an NBES member and welcomed Dr. Granger, who returns to NBES for a second term after having served as NBES's first chair.

Mr. Baron said he had been involved in several excellent meetings with congressional staff and other federal agencies, including the Office of Management and Budget (OMB). The Board's unanimous recommendation to advance the credible use of evidence in federal education decision-making has been particularly helpful in those discussions, providing a good introduction for explaining why rigorous research and evaluation are not luxuries, but rather key ingredients for progress. It is often eye-opening when staff members learn that many measures of educational progress over the past 30–35 years do not show a lot of improvement. For example, high school graduation rates are actually slightly lower now than they were in 1969.

In discussions with federal decision-makers, Mr. Baron points to examples of how IES supports rigorous research that offers a way forward in education. For example, a study by Dr. Long, funded by IES and the Gates Foundation, of a large, multi-site experimental program, found that when H&R Block tax professionals assisted low- and middle-income families with completing college financial aid applications college enrollment among those families increased by 25 to 35 percent, a result that persisted over a 3-year follow-up period. Such findings speak volumes to non-researchers in a way that 1,000 experts testifying on Capitol Hill cannot, said Mr. Baron. They give a sense of the credible path forward that differs from the past 50 years, which has been a path that moves from one educational strategy to the next, depending on who is president or who is in control of congress, without significant progress.

Senators Harkin and Enzi recently introduced a bill to reauthorize the Elementary and Secondary Education Act (ESEA). It has bipartisan support and will be marked up soon; it reflects a strong appreciation of IES and education research and includes NBES recommendations. If the bill passes, it will represent the first time that major education legislation explicitly recognizes IES as the lead agency for program evaluation. The inclusion of the term "evidence-based" in the current bill reflects high standards for evidence. The bill would substantially increase funding set-asides for evaluation. For example, funding for evaluation in the No Child Left Behind (NCLB) program would increase from 0.5 percent to 3 percent, and Title I would have a new 1 percent set-aside for

evaluation. Mr. Baron thought the bill would likely be passed by the Senate HELP (Health, Education, Labor, and Pensions) Committee with bipartisan support, but whether it would be enacted in the near term is an open question. If it were not enacted, the focus on evaluation would probably serve as a basis for future legislation.

Regarding annual appropriations, IES is doing quite well compared with the rest of the federal government. The Senate authorized level funding, but allotted 5 percent out of funds from Race to the Top and Investing in Innovation (i3) for evaluation and technical assistance (TA) (in which IES is involved) to support good impact evaluation of programs. The House appropriations bill includes a small increase for IES of \$12 million for the Regional Education Laboratories (RELs), which is in contrast to fiscal year (FY) 2011 when Congress proposed eliminating funding for the RELs.

Mr. Baron said he and Dr. Long wrote a letter to the U.S. Department of Education (ED) Secretary Arne Duncan on July 19, 2011, on behalf of NBES regarding the Secretary's authority to grant waivers to states and districts releasing them from some of the NCLB adequate yearly progress accountability provisions. The letter draws an analogy to welfare program waivers granted in the 1980s and 1990s that allowed states to establish experimental programs if they included rigorous field evaluation. The results of those experimental programs had a major impact on federal policy, particularly the 1996 welfare reform legislation. The letter makes the case for incorporating similar evaluation requirements into the NCLB waiver policy.

Dr. Gamoran asked about the status of the Education Sciences Reform Act (ESRA). Mr. Baron said he and Dr. Herk had a good meeting with Republican House members, who asked some follow-up questions about ESRA. Questions covered, for example, whether NBES recommends defining "scientifically based research" or "scientifically valid research." While there seems to be interest in ESRA, Mr. Baron said representatives seemed to be focused on reauthorization of NCLB, not IES.

Update: Recent Developments at IES

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

Dr. Easton noted that the work of IES takes place in four centers and two deputy offices, and representatives of most of those entities would provide updates to the Board.

Introduction of the NCSER Commissioner

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

Dr. Easton introduced the recently confirmed National Center for Special Education Research (NCSER) commissioner, Dr. Deborah Speece, who is NCSER's first commissioner in 5 years. Dr. Speece comes to IES from the University of Maryland, where she spent nearly 3 decades studying learning disabilities. Dr. Easton said Dr. Speece is highly regarded by her peers, and he looked forward to the leadership and guidance she would provide to the community.

Dr. Speece said she has been working in the field of special education for over 35 years, first as a teacher of children with learning disabilities and behavioral disorders, then in higher education—spending the bulk of her career at the University of Maryland as a teacher and researcher. Her primary focus has been the identification of children with learning disabilities, specifically reading disabilities. Her research over the past 10 years focuses on response to intervention as both an identification and intervention process.

In the 38 days since she began at NCSER, Dr. Speece said she has learned a lot and has come to understand the breadth of the IES research portfolio. She is currently working with Dr. Albro on documents that synthesize and highlight the contributions of IES research on reading in early childhood and among children and youth with (or at risk for) disabilities. Dr. Speece concluded that there is a lot to do, but she appreciates the challenge.

Swearing-in of New NBES Members

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

Dr. Easton said it was a pleasure to bring back Drs. Bryk and Granger. He said Dr. Bryk had been a marvelous mentor to him for many years and Dr. Granger has a long history with NBES and IES. Dr. Easton then swore in Drs. Bryk and Granger to NBES.

IES Commissioners' Update

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

Dr. Maynard encouraged Board members to visit the *What Works Clearinghouse (WWC)* website, which has been substantially reorganized and now allows users to search by topic, not just product. The *Find What Works* tool is more visible and functional for users. The site now features a fully searchable, downloadable database of all the studies reviewed, with links to reports that cite those studies. It has many new features that improve use of the content in ways the database could not previously support.

In keeping with federal and IES interests, NCEE is emphasizing small-business contracting. NCEE has identified small businesses that can do the work required and is making an effort to structure its work to accommodate them. NCEE made two small-business awards in the past few months and expects to award more funds in the next 3 to 6 months.

NCEE continues to strengthen its staff, recently hiring three new staff members who bring great experience. The center is in the final stages of REL competition, and staff members are working hard to review 10 proposals that are each 800 pages long. New RELs will launch in January. A parallel effort is underway to ensure analytic and technical support is in place for the RELs.

Dr. Maynard said NCEE is reaffirming, with the RELs, its commitment to well-designed causal inference studies—the center will not do studies that it cannot stand behind. A big push is underway to take translational research and local capacity-building efforts to the next level.

NCEE is part of an interagency working group to establish common standards that would be used to judge how much evidence supports moving a policy forward, as well as to determine the standards that should be expected for new programs or policies when there is no applicable evidence. There is a lot of confusion about whether evidence is needed at the outset of programs, on completion, or both. In many areas of education there is no evidence going into a program, so there is a push toward agreeing to implement programs even when there is no evidence upfront, but to identify the needed evaluations to ensure the evidence from the completed program is strong.

NCEE has been assertive and persistent in pushing for the ESEA flexibility (waiver) package, and Dr. Maynard believes progress is being made. The language of the package includes a strong invitation to waiver recipients to work with ED to add evaluation. Three areas were singled out that are most pressing in terms of data: college and career-ready standards, state-differentiated recognition and accountability systems, and support for effective leadership and instruction. NCEE wrote a model strategic plan describing how a state or district could take advantage of an evaluation opportunity to improve the success of its implementation and how a state could implement most but not all components of a plan and then evaluate progress. Dr. Maynard said IES program staff needed to see that evaluation will not up-end the programs in schools. More must be done to get NCEE's approach included in the waiver.

Roundtable Discussion

In response to Dr. Granger, Dr. Maynard explained that language in the waiver encourages evaluation but does not require it, because there is no quid pro quo expectation of cooperation and no explicit funding for evaluation. If a lot of institutions offered to institute evaluation efforts, NCEE would be scrambling to support them, although the center does have some mechanisms in place. In contrast, welfare programs in the 1980s and 1990s had program dollars, grants, and matching funds. Dr. Granger asked whether resistance is related to giving states flexibility to implement programs or the possibility of an unfunded mandate to evaluate programs. Dr. Maynard said the discussion has not reached that level yet. Rather, she sees the issue as the normal tension between program implementation and evaluation, and NCEE must convince program officers that evaluation can be conducted in a sensitive, non-intrusive way that provides information useful to program staff. For program officers, admitting that evaluating programs is a good idea may give the impression that the program is not already proven. Historically, Dr. Maynard continued, evaluation has been intrusive, results have not been packaged in a way that is meaningful to the users, or results were not provided to those on the ground. However, IES staff has become more sophisticated and now understands that the purpose of evaluation is to help the program and the policy community.

Responding to Dr. Gamoran, Dr. Maynard said she hoped RELs would reach out to states that are considering evaluation. Technical support and guidance for RELs will codify IES' expectations for effectiveness evaluation and highlight issues that should be studied in real-world settings and that matter to the regions served. Dr. Gamoran said that, as a REL board member for 6 years, he has come to believe that state agencies will probably not take on evaluation unless pressed to do so, and he asked what persuasive power ED or IES could bring to bear, because they have more leverage than RELs. Dr. Maynard said RELs have a different leadership structure now that will facilitate more expertise on the ground, which she hopes will help in gaining a foothold for evaluation.

Mr. Baron congratulated Dr. Maynard on getting evaluation language included in the waiver policy. However, the language describes the procedure without giving the rationale. He suggested developing a brief (e.g., two-page) outline that explains what evaluation efforts involve (emphasizing that they need not be an administrative nightmare) and how programs benefit from evaluation. The outline should include some concrete examples of valuable knowledge gained from evaluation that states have used to improve practice. The language in the waiver should be demystified so that it is clear to non-researchers and should include a rationale to motivate evaluation.

Dr. Maynard responded that such a document would be in line with the TA, guidance, and documentation that NCEE hopes to have in place; the concept is also consistent with the planning underway for the new RELs, as well as with the background documents NCEE wrote to explain how evaluation would work. She said that when a program first rolls out, studies tend to be narrow in focus, which helps get at the critical question of what works. It's important not to nail things down too firmly, she added, and to protect the integrity of a study while allowing researchers to tailor implementation to the site.

Dr. Long asked whether the Board could play a role in bolstering Dr. Maynard's position. Dr. Maynard replied that it doesn't hurt to have multiple voices supporting the same issue, so an easy-to-read, two- or four-page summary written by NBES would be a nice addition to the arsenal. She added that the culture must change to make evaluation easier and more useful, and NCEE is moving in that direction. The RELs' launch meeting takes place in late January. Mr. Baron reiterated that it would be helpful for NBES to develop an explanation and rationale for evaluation and to do so in close consultation with NCEE to ensure everyone is on the same page.

Dr. Granger pointed out that when the evaluation component for welfare waivers was created, federal funds covered 90 percent of the costs and states covered only 10 percent. Welfare program staff had the same resistance to evaluation as education staff do now, but money changes behavior, said Dr. Granger. No one could imagine tying teachers' salaries to student achievement until Race to the Top provided the money, he added. Dr. Maynard agreed that states have no money for evaluation right now, so NCEE would have to come up with resources using its current structure. She stated clearly that no additional money is allotted to support state

evaluation efforts. However, NCEE can ensure that evaluation is built into initiatives and provide TA that will yield some good research.

National Center for Education Statistics (NCES)

Jack Buckley, Ph.D., Commissioner, NCES

NCES posted a new request for applications (RFA) for the next round of Statewide Longitudinal Data Systems grants. Applications are due December 15. NCES is seeking applications from state education agencies and their partners who propose to design or improve K–12 data systems or to link existing systems with early childhood, postsecondary, or labor force data systems. Agencies that have received grants under the American Recovery and Reinvestment Act of 2009 are not eligible.

The Assessment Division recently reported on the latest iteration of its state mapping project, which uses National Assessment of Educational Progress (NAEP) findings as a common metric to put NCLB proficiency scores on a common scale. As usual, the report found wide variations in states' proficiency standards, and many of those standards fall between NAEP's basic level and proficient level. The most interesting new finding, said Dr. Buckley, is that most states that had a change in cut scores had moved in the direction of greater rigor, setting the bar higher. South Carolina, however, set less rigorous standards, arguing that the state's previous standard had been too high.

The National Assessment Governing Board postponed the release of the 2011 state and national NAEP results for reading and math to November 1. Dr. Buckley explained that the tone of the report regarding exclusion rates will differ from previous reports because of a policy change. The Trial Urban District Assessment Report will be released about a month after the national and state results.

NCES has also been looking at exclusion/inclusion policies, independently of the Governing Board, in response to efforts by the Government Accountability Office to measure changes in states' exclusion/inclusion rates. NCES seeks to look at state demographics and predict inclusion rates. Recent reports showed no decrease in inclusiveness for grade 4 or grade 8 for either subject in the years compared.

The Elementary/Secondary Division is moving ahead rapidly on a new project to identify all public K–12 catchment zones through geocoding and geomapping; as NCES currently does at the school district level with its partners at the U.S. Census Bureau, it will use the data to run tabulations at the school level for the American Community Survey. As a result, NCES will be able to report demographic data at the school/neighborhood level in the catchment zone for schools. Thus, said Dr. Buckley, NCES will have real data about the neighborhoods surrounding a school instead of proxy data, such as eligibility for free and reduced lunch programs.

The 2008 reauthorization of the Higher Education Opportunity Act included many provisions for expanding the availability of information to parents and students about higher education. As part of that effort, the Postsecondary Division assisted ED in developing a methodology for colleges and universities to implement a required net price calculator that all Title IV-eligible institutions will be required to put on their websites this month. NCES collected institutional net price data and disseminated it through College Navigator (<http://nces.ed.gov/collegenavigator/>) and FAFSA.com, the website of the ED's Federal Student Aid office. In addition, NCES launched the College Affordability and Transparency Center (<http://collegecost.ed.gov/>), which ranks institutions by tuition and net price. College Navigator includes a link to the College Affordability and Transparency Center.

NCES is releasing data from the *Baccalaureate and Beyond* study on the 2007–2008 cohort and will soon release transcript data. The data reveal some new and interesting information about students' trajectories through college, said Dr. Buckley.

The Early Childhood, International and Crosscutting Studies Division is collecting data for the U.S. component of the Programme for the International Assessment of Adult Competencies (PIAAC), an adult literacy study sponsored by the Organization for Economic Cooperation and Development (OECD) that includes the United States and 25 other countries. It is designed to allow for cross-national comparisons and to link with other OECD studies, such as the International Adult Literacy Survey and the Adult Literacy and Lifeskills Survey. Data collection will be completed in March 2012, but findings won't be released until October 2013. International data take a long time to coordinate, said Dr. Buckley.

The PIAAC is completely computer-administered, Dr. Buckley said, and adaptive at the test level, well below the usual two-stage adaptive assessment. It focuses on assessing components of reading and problem-solving in a technology-rich environment. As such, it looks at 21st-century technology skills.

Finally, the same division is redesigning the National Household Education Survey to move from a telephone-administered to a mail-administered survey. Experiments so far indicate that mailing the survey will boost the response rate measurably.

National Center for Education Research (NCER)

Elizabeth Albro, Ph.D., Associate Commissioner, Acting Commissioner, NCER

Dr. Albro noted that IES has administered 40 awards across the 2 research centers (NCER and NCSER)—specifically, 38 research grants, 1 evaluation of state and local programs, and a new Research and Development Center focused on postsecondary education and employment. Fourteen of the new research awards focus on math and science across different topic areas:

- Exploratory research to understand features of interactive science learning and motivation in elementary school students
- An efficacy study on interleaving math practice problems in which teachers will evaluate implementation in a typical classroom
- Development of curriculum resources for teachers based on successful Japanese materials

NCER also established the Center for Analysis of Postsecondary Education and Employment to identify links between postsecondary education and the labor market. It has 12 projects to carry out over 5 years. They address, for example, whether short-term occupational degrees improve labor market outcomes, the effects of non-credit workforce programs, the role of for-profit educational institutions on employment outcomes, and the trajectory of earnings growth after college. The Center will use data from North Carolina, Michigan, Ohio, Virginia, and Florida.

Dr. Albro said her tenure at IES allows her to see long-term research programs deliver results. She cited Doug Clements' work on the Building Blocks math curriculum, which was originally developed with National Science Foundation (NSF) funding and supported by small efficacy trials funded by both NSF and IES. An evaluation of the program is currently scaling up. The program includes an early childhood curriculum to support math knowledge in pre-K children. Researchers are studying students who took part in the program through kindergarten and into the first and second grades. They've found a sizeable, positive effect size (0.76) in pre-K children that persists with reinforcement in the classroom. Dr. Albro noted that effects seen in early childhood usually fall away, so she was pleased to share results of an intervention where the effects are maintained over time.

Dr. Albro noted that a teacher quality project recently published a study in the *Journal of Research on Educational Effectiveness* looking at professional development for teachers of mainstreamed Latino English language learners that used a common strategic approach to analytical writing instructions. Through a randomized study, the program found effects not just on the literary analysis but also on the English language arts subtest of California's standardized test. James Kim and Carol Olson authored the study.¹

Finally, Dr. Albro noted, ED and NCER are participating for the fifth year in the Presidential Early Career Award in Science and Engineering. Roy Levy, a young scholar who received an award through NCER, is being recognized for his psychometric work.

Roundtable Discussion

Dr. McCardle asked whether any of the 12 studies by the Center for Analysis of Postsecondary Education and Employment would address English language learners. Dr. Albro did not know, but said it probably depends on how well coded the state data are.

¹ Kim, J. S., Olson, C.B., Scarcella, R., Kramer, J.S., Pearson, M., van Dyk, D., Collins, P., & Land, R. (2011). A Randomized Experiment of a Cognitive Strategies Approach to Text-Based Analytical Writing for Mainstreamed Latino English Language Learners in Grades 6-12. *Journal of Research on Educational Effectiveness*, vol. 4, no. 3, 231-263.

Peer Review of Research Proposals: The IES Approach and Possible Refinements to Increase Findings of Policy Importance

Introduction

Jon Baron, J.D., NBES Chair

Mr. Baron noted that while grant making and journal publishing rely heavily on peer review to assess the quality of applications and submissions, there is little evidence about the effect of peer review on the quality of funded research or published papers. Mr. Baron pointed out that peer review is highly respected, but it may not guarantee a good product, citing examples of education research articles published in *Science* that would not have met the standards of the *WWC*.

Opening Remarks

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

Dr. Ricciuti agreed that peer review is not a perfect process at IES or anywhere else. She limited her remarks to the review process for research grant competitions run by NCER and NCSEER, which does not apply to the evaluations conducted by NCEE. The Standards and Review Office, which she runs, handles receipt of proposals and peer review; the Research and Development Centers generate the RFAs that guide applicants as well as the peer review process, and the Centers make the funding decisions. Background materials provided to the Board include procedures for peer review (approved by the Board in 2006), information about updates and improvements since 2006, reviewer materials and guidance, a list of reviewers for FY 2011, and a sample RFA.

The Standards and Review Office received more than 1,400 applications last year, about twice as many as in 2006. It has expanded use of the online scoring system to handle compliance and responsiveness screening and conflict-of-interest identification and documentation. It has also sought to improve instruction and guidance for reviewers. A “highlights” document for reviewers summarizes the major requirements (as described in the RFAs) of each of the research goals of the primary research grant competitions.

In response to applicants’ request to receive feedback more quickly after panel meetings, IES is exploring an online applicant notification system, similar to that used by the National Institutes of Health (NIH) and NSF, through which applicants could log in and get their summary statements. Dr. Ricciuti welcomed feedback and ideas.

Roundtable Discussion

Dr. Gamoran commended Dr. Ricciuti for the quality of the IES peer review process, stating that the RFAs do a “tremendous” job of giving clear instructions and describing what ED seeks to fund. In addition, the reviews provided to applicants, particularly for unsuccessful proposals, are extremely helpful, constructive, and timely. As an individual who has received funding from both

IES and the National Institute of Child Health and Human Development (NICHD), Dr. Gamoran said he has insight into both agencies' review processes. NIH program officers bring significant expertise to the process—perhaps because NIH has been funding grants for so long, NIH program officers can anticipate what peer reviewers are looking for and they apply standards more effectively than IES program officers. He urged Dr. Ricciuti to encourage IES program officers to listen carefully to the deliberations of the peer review panels and to explicitly state that program officers should understand what reviewers are looking for.

Dr. Gamoran said that, in his experience, IES program officers do not have enough experience to help applicants determine before submission whether their proposals are a good fit for a given competition. By contrast, there is an expectation that NICHD program officers have that kind of expertise.

Dr. Gamoran posed several other questions to Dr. Ricciuti:

- How many rejected proposals are resubmitted?
- How many different topics does each panel review?
- How stable are panels over time?
- Because serving on a scientific review panel is a burden, has IES considered working with SREE or another organization to enhance the prestige of serving on a standing review panel, which would promote stability and continuity over time?

Dr. McCardle said that Dr. Gamoran may be overgeneralizing the issue related to the experience of program officers. She suggested that applicants should develop a relationship with at least one program officer and that applicants should always talk to program officers before they submit proposals. Bad advice from a program officer should be reported to the officer's boss, and the applicant should ask the question again, ask someone else, or consult a little more about the question.

Dr. Bryk said he was struck by the focus on "development and innovation" in IES peer reviews, terms likely to elicit enormous variability among reviewers and panels. Traditionally, innovation involves a lot of risk, and 90 percent of good ideas don't pan out. Dr. Bryk said the bulk of IES proposals now fall under the category of "development and innovation." Dr. Albro responded that the funding percentages closely reflect the percentage of applications that come in; about half of the projects funded fit under development. Dr. Bryk questioned the reliability of the review process, because the category of development and innovation appears to require more subjective judgment than others.

Dr. McLeod commended IES on the rigor of its peer review process but noted that the list of IES peer reviewers provided to the Board lacked Hispanic surnames. She acknowledged that surnames tell little, but she was nonetheless concerned. She pointed to well-known pipeline issues (e.g., only 2 percent of doctorates go to Hispanics each year) and said that Hispanic-serving institutions may not be heavily involved in research. However, she asked IES to consider how to increase the number of Latino peer reviewers.

Opening Remarks

Deborah Gorman-Smith, Ph.D., Senior Research Fellow, Chapin Hall at the University of Chicago, President, Society for Prevention Research

Dr. Gorman-Smith provided her perspective on recent significant changes in the NIH review process. NIH undertook the changes to (1) decrease the burden on reviewers and (2) encourage reviewers to focus more on the significance of the proposed work and less on minor details about the approach and methodology. Application page limits were cut in half—but reviewers still expect to see the same level of detail. The shift from a narrative to a bullet-point format for reviews decreases some of the burden on reviewers, but also provides less helpful information that applicants can use to resubmit their proposals.

New investigators in particular don't understand the importance of developing relationships with program officers. Program officers are critical to the process; they listen to panel members during the reviews and understand their expectations, so they can work to ensure that an application goes to the most appropriate panel for review.

On the basis of her experience as a reviewer, findings from focus groups, and issues raised at a recent National Institute on Drug Abuse meeting on the review process, Dr. Gorman-Smith described some of the key ingredients for a high-quality review.

- Scientific review officers (SROs) oversee panels, set the tone for the quality of the review, and signal to the members the important aspects of the review. Some SROs carefully select members by taking into account not only individual qualifications but also personality and communication style.
- Even more important than having a range of expertise across substantive and methodological areas is having a mix of seniority in the room.
- Many qualified potential reviewers opt not to serve on panels because it undermines their own funding prospects. (Applications by an impaneled reviewer go to a conflict panel, where scoring is not as well calibrated and the reviews are not as good.)
- Panels sometimes suffer from a lot of turnover (which seems to be more pronounced among IES than NIH panels). It can take two to three rounds of review before panelists reach a comfortable working relationship.
- The panel chair sets the tone for the review by keeping an eye on the time that panelists spend talking about an application and shaping the summary.
- Discussion limits of 10–15 minutes per application may seem insufficient for review, but scores are very similar to those that resulted from longer discussions. In fact, the longer discussion continues, the worse scores seem to get, because reviewers find more details to criticize.
- It takes reviewers several rounds of review to begin scoring for impact and significance—not just evaluating the methods and approach and calling that the impact. Over time, reviewers learn to balance their considerations, partly through training and partly through continuing discussion.

Roundtable Discussion

Dr. Long said she appreciated recent efforts to clarify RFAs for both applicants and reviewers and to incorporate technology to make the review process more efficient. She agreed that reviewers can get hung up on, for example, how a very detailed regulation might apply to a given proposal, overlooking the larger picture. She asked how the centers get feedback on how reviewers and applicants are interpreting regulations. Dr. Long noted that IES panels are more interdisciplinary than, for example, NSF panels, so applicants are challenged to provide enough technical detail to satisfy the experts in the room and enough background and explanation to reach those from other disciplines. She noted that panel chairs get advice on procedures but not on managing the discussion.

Terms such as “significance,” “impact,” and “innovation” speak to what Dr. Long believes is the need to highlight and support work with the potential to fill a major gap or push a boundary in the field. At present, many studies are designed to make small contributions or don’t provide information about outcomes of interest. In addition, Dr. Long suggested that reviews could give further consideration to areas in which federal government funding provides a comparative advantage on the basis of size and resources.

Dr. Shaywitz proposed evaluating the review process by assessing the impact of funded proposals on improving education or changing policies and practices. She suggested identifying those studies that have had a significant impact and looking at the emphasis and scores of the applications. She acknowledged that it would be a difficult and complicated process but an important one. Dr. Easton said IES is currently summarizing the findings from IES-funded studies in three different topic areas. IES could look at the summary reviews of the applications for those studies. Mr. Baron suggested also looking back at funded studies that failed to produce positive findings, but Dr. Shaywitz hoped the emphasis would be on work that made a difference and moved the field forward.

Dr. McCardle returned to the issue of assessing proposals for innovative potential. Programs must balance their funding decisions to allow for some risk (and potentially high payoffs) while also ensuring stability. She noted that to entice more reviewers to participate in standing panels, NIH now holds one of the three review panel meetings each year outside of the Washington, DC, area, but travel funds are not available for all NIH program staff. Dr. McCardle said that when staff members participate over the “terrible” phone systems they do not gain the full experience that reviewers in the room experience, so program officers may not be able to provide as much insight into the review process.

To address concerns about error and bias, NIH has an appeals process for applicants, said Dr. McCardle. Also, NIH has an administrative review process to assess the eligibility and appropriateness of a proposal for a given RFA. Reviewers are told that the applications they are reviewing have been deemed acceptable, so they don’t need to spend time discussing how eligibility requirements might apply, which can save a lot of time.

Speaking to Dr. Gorman-Smith's comment about the importance of seniority, Dr. McCardle said that junior reviewers are more likely to focus on finding mistakes in the methodology of an application at the expense of the bigger picture because they feel compelled to establish their credibility among their more senior peers. Dr. Long noted that a significant incentive for junior faculty to participate in review panels is the opportunity to network with senior faculty and demonstrate their skills. Dr. McCardle said reviewers also get an inside look at the review process and what constitutes a well written or a poorly written application.

Dr. Granger pointed out that in this discussion, the term "significance" was used to refer to the impact of research on the literature, but also to the impact on policy and practice. He believed that program officials, not reviewers, should determine what would significantly impact policy and practice—and that their determination should be reflected in the RFAs. He added that it's difficult to acculturate practitioners to the review of empirical work.

Dr. Ricciuti appreciated the positive comments of the Board members and acknowledged the work of her staff. She said IES struggles with balancing considerations of an application's significance with the importance of ensuring that methodology is sound. She noted that each RFA addresses significance in a way that is tailored to the goals of that competition. Dr. Ricciuti said new guidance to reviewers emphasizes that input from those directly involved in policy and practice can be part of a reviewer's rationale about the significance of an application.

Regarding the roles of program officers in providing advice to applicants, Dr. Ricciuti indicated that the Commissioners are the more appropriate individuals to address those questions because IES program officers are separate from the Standards and Review Office. In terms of resubmission, IES tries to send resubmitted applications to the same reviewer and to encourage reviewers to consider responsiveness to the previous reviews. IES is receiving resubmissions at a higher rate than in earlier years.

Dr. Ricciuti said she and her staff identify potential reviewers, and she approves those who serve on a rotating or ad hoc basis. She and the Director, both, must approve members invited to serve a 3-year panel term (principal members). Dr. Ricciuti said efforts are made to avoid excessive turnover. She added that ensuring diversity within panels is a goal, and she appreciated any suggestions to improve the diversity of panels. Most panel members fulfill their term obligations. IES often invites reviewers to serve on a rotating basis as a trial run and to ensure that reviewers understand the burden involved if they are invited to serve as principal panel members. Dr. Ricciuti said a number of people do see the service as an honor, but she welcomed suggestions on boosting the stature of the job. IES may look into honoring its best reviewers, as NIH does now. Dr. Ricciuti assured the Board that IES takes into account interpersonal skills when considering inviting someone to become a principal panel member.

Mr. Baron reiterated his support for identifying funded studies that produced important results and seeing how they fared in the peer review process. Dr. Ricciuti noted that, because they were funded, such applications had to have received high scores. Dr. Long suggested identifying items

in the *WWC* and evaluating whether they were funded by IES or rejected and funded by another entity. Dr. Granger warned against placing too much emphasis on the results or impact of a single study, which Dr. Maynard strongly supported. Dr. Shaywitz suggested further consideration about how the Board or IES could look more closely at the work that has made a difference. Dr. Maynard added that scientists can make a major contribution by highlighting that some things accepted as good practice are not. Dr. Shaywitz agreed, noting that the *WWC* demonstrated that many accepted practices are not effective, yet negative results often are not published. Dr. Maynard said the new website does publish negative results.

Dr. Gutierrez said that in 1999, the Office of Educational Research and Improvement (OERI), IES's predecessor, engaged an external review of its peer review process, and many of the resulting recommendations were incorporated into IES's peer review process. In 2006, the Board took on the role of evaluating the peer review process, and she suggested the Board consider whether to review the process again. Mr. Baron said such review is part of the Board's statutory responsibilities, and he suggested the next Board chair consider whether to undertake such a review.

Dr. Gutierrez noted that having a diverse review panel not only speaks to demographic representation but also brings to the table new questions, as well as new problems, in practice. There's so much to be learned from people on the ground, who grapple with issues every day. Dr. McCardle said efforts at diversity also respond to the demographic makeup of the nation and the children served.

The Board adjourned at 10:54 a.m. and reconvened at 11:07 a.m.

The Administration's "Tiered" Evidence Initiatives in Education and Other Areas: New Approach to Stimulating Development and Use of Rigorous Evidence

Introduction

Jon Baron, J.D., NBES Chair

Mr. Baron said that under President Obama, a number of initiatives have been enacted into law, embodying a new approach to stimulating the development and use of rigorous evidence, such as i3. Introducing the guest speakers, he said Ms. Stack has thought creatively about how to embed evidence-based concepts into the structure of federal education programs, and Ms. Thompson is a school official who has prioritized research in practice.

Opening Remarks

Kathy Stack, Deputy Associate Director for Education and Human Resources, OMB

Ms. Stack said OMB's role is to make government more accountable, efficient, and effective, but most federal programs are funded on the basis of hypotheses. When those programs are evaluated rigorously, they often show no impact, but the results are not used to inform future

program efforts. The Obama Administration provided a unique opportunity to address the problem. Ms. Stack explained that OMB leadership recognized that programs were competing for resources against research and evaluation efforts, instead of partnering toward the same goal. The concept of tiered evidence was invented to create incentives for researchers and practitioners to work together, build on existing evidence, evaluate their efforts, and produce new evidence to support the growth of best practices and to phase out efforts that were not effective.

Ms. Stack directed the Board to Table 1 of *Building the Connection Between Policy and Evidence: The Obama Evidence-Based Initiatives*², which summarized six evidence-based federal programs across several departments. The funding was structured so that large programs supported by high-quality evidence received more money, but funding remained available for developing new ideas and validating what worked. Taking this approach to two divisive topics—teen pregnancy and home health visits—the U.S. Department of Health and Human Services (HHS) calmed the political waters by sending the message that the approaches with the strongest evidence to support the best outcomes would get the most money, regardless of their philosophical foundation.

The same tiered structure was applied to i3, which sought to narrow the field of applications by tying i3 to the evidence framework devised by IES and the WWC. i3 inspired other funding agencies; through tiered grants, they could encourage new and innovative ideas at the lower tier of developmental grants, provide evaluation tools, and offer larger grants to scale up effective initiatives. i3 was the model for the Social Innovation Fund of the Corporation for National and Community Service. Similarly, the Department of Labor’s Community College and Career Training Program seeks to develop strong evaluations and move effective strategies up the ladder. Its Workforce Innovation Fund will support partnerships with ED and HHS on systemic reform and better interventions to improve workforce outcomes.

Ms. Stack said these programs were launched when resources were available; she worried that tiered evidence initiatives, particularly research and evaluation dollars, are at risk if the investigators and practitioners and program officers retreat to their silos. However, she also saw opportunities for bipartisan consensus around making smarter spending choices. Research and evaluation professionals can make the case that they are both relevant and necessary for effective programs. Ms. Stack suggested more partnering and more awareness of the role of IES in evaluation and research. She also recommended more focus on cost-effectiveness. Thanks to the tiered evidence structure, she noted, state and local governments now have strong evidence about effective programs to prevent teen pregnancy—but they still don’t have correlating cost-effectiveness data that would allow them to make smarter choices.

² http://www.brookings.edu/reports/2011/0907_evidence_based_policy_haskins.aspx

Opening Remarks

Saskia Levy Thompson, Chief Executive Officer, Office of School Support, New York City Department of Education

Ms. Thompson explained that her background in social policy research gives her an appetite for, and understanding of, evidence-based practice that she applies in her role as a school official. She pointed out that the New York City school district is much larger than most others (1.1 million students, 1,700 schools, \$23 billion budget), but lessons learned are still applicable elsewhere. She offered the following framework for incorporating research into practice:

- Building internal capacity to use data to solve complex problems: informed by lessons learned from the Chicago Consortium³ and partnership with the Research Alliance in New York City
- Evaluating district-initiated strategies through partnerships: looking at homegrown efforts
- Using research to develop strategies in high-need areas: identifying persistent problems and using research to fill gaps

Building Internal Capacity

To illustrate the goal, Ms. Thompson pointed to the New York City Department of Education's agreement with the City University of New York (CUNY) to share data about how local high school graduates fared in college (funded by the Gates Foundation). She noted that the city had made remarkable progress increasing high school graduation rates over the past 10 years, and almost 40 percent of those graduates attend CUNY. Data showed that half of the city's high school graduates required remedial education at CUNY, and similar problems were noted around college graduation rates. The findings were packaged into reports that went to every high school, showing how their graduates fared on various measures of academic preparedness and success, which was incredibly meaningful for the schools. The data also informed a set of college readiness metrics that are included on each school's annual progress reports.

Evaluation

One way to marry research with practice is to evaluate homegrown efforts, such as the Innovation Zone. Schools in the program implement innovative classroom models, ranging from incremental approaches (such as online coursework to recover credit) to whole-school redesign around blended learning. With private funding, the city forged a partnership with EdLabs to conduct a randomized study this year in 30 elementary schools, evaluating the impact of three personalized learning systems of varying intensity.

Research-Based Strategies

Two areas of particularly high need in New York are middle school student literacy and postsecondary school readiness among young men of color. Ms. Thompson explained the persistent literacy problem and said the school chancellor wants to create 50 new middle schools that approach literacy differently. Ms. Thompson hopes to ensure that the effort reflects literacy

³ <http://ccsr.uchicago.edu/content/index.php>

initiatives for adolescents that have actually worked. Also, young men of color in New York are graduating at lower rates than their peers, and those who do graduate are not necessarily well prepared, so research is needed to understand interventions that could benefit this population.

Ms. Thompson echoed Ms. Stack in saying that cost-effectiveness is widely under-addressed. Not many school districts are paying close attention to the findings of the *WWC*, but cost-effectiveness results would garner real attention.

Roundtable Discussion

Mr. Baron said the tiered-evidence initiative requires agencies to look beyond statistical significance and assess whether effects are meaningful in policy and practice. i3 did a particularly good job identifying programs backed by strong study designs that had meaningful effects on outcomes, such as reading comprehension, sustained over time. However, some of the teen pregnancy programs, for example, are backed by evidence of impact on factors that may not translate to reducing teen pregnancy, such as increased condom use or decreased number of sexual partners. Some of the home health visitation programs met minimum standards of evidence by showing statistically significant results, but the absolute numbers were small and the outcomes of questionable impact.

Ms. Stack pointed out that i3 revolved around issues that were not politically sensitive, unlike teen pregnancy interventions. Carefully negotiated legislative positions have limited what HHS can do to assert that effects are meaningful. Education focuses on student achievement and is far less partisan. Ms. Thompson said there's a dynamic tension in a school system between measures that have real accountability consequences and meaningful long-range measures. Managing a school system requires directing precious resources to address the most pressing and granular issues; the federal government can counterbalance that focus by looking at long-term impacts.

In response to Mr. Baron, Ms. Thompson described the elementary school evaluation of personal learning systems in more detail. Some data will demonstrate what students learn through technology and what they learn from teachers, which will inform a larger-scale approach in the near term. Data from standardized tests can show the effects of the intervention over time.

Dr. Granger noted that the two presenters represented two approaches to changing services: one driven by the research community, which develops, deploys, and evaluates programs; the other by practitioners who look to research to evaluate and refine their practices over time. Neither approach has had a "slam dunk" success so far, he said, but the goal is to keep learning from the efforts. Looking at the teen pregnancy outcome measures allows researchers an opportunity to revisit the outcomes they thought were meaningful at the time.

Dr. Granger said i3 does not have a strong political constituency, and if it is seen only as a funding stream, it will eventually run into the same political barriers that other funding efforts have. He asked how IES will learn from i3. Dr. Easton said the funded i3 applications all include rigorous

evaluation, for which IES provides technical support. Dr. Maynard said the evaluation expectations vary by tier, just as the funding does. Some of the issues that have arisen are as follows:

- Is it as important to have strong evidence on the front end as it is on the back end? IES can support a well-conceived idea based on sound logic (in the absence of strong evidence).
- Some effective changes can be implemented at low cost (e.g., administrative changes), while others (e.g., interventions for children with special needs) may be very costly.

Dr. Maynard emphasized that the goal of i3 is to stimulate good investments in education, partly by taking some risks and supporting development of new ideas and then challenging grantees to evaluate those new approaches to learn what worked, why it worked, and how to scale up.

Ms. Thompson said New York City only received two low-tier i3 grants despite numerous applications, and the funded proposals are not central to the administration's strategic priorities (although innovative). i3 did not fund the city's large-scale proposal to study the effect of closing large, failing schools and opening smaller schools, which would have provided unprecedented insight into reform strategies but did not lend itself to a quick research design. Mr. Baron said a goal of the tiered-evidence initiatives was to build knowledge about new approaches and, for approaches that already have good evidence of meaningful impact, to scale up those approaches to affect outcomes.

Dr. Gamoran asked what lessons were learned from the strategies ED applied to review 1,600 proposals in a short time frame (e.g., the use of specialized reviewers who focused only on the strength of the existing evidence and evaluation design). Dr. Maynard said the approach allowed for a more focused, consistent review process, because a smaller number of people reviewed applications for a very specific purpose, and ED was able to triage the applications. The emphasis of i3 was to get the resources out to stimulate innovative development and testing; so reviewers focused first on those aspects of applications, then on the evidence evaluation criteria. Dr. Maynard said there had not yet been discussion about whether the approach could be adapted to the general peer review process.

Dr. Maynard added that IES has a list of "certified reviewers" that other federal agencies can use, and it expanded the training program within the WWC substantially this year. These efforts are intended to build capacity. The Department of Labor, for example, recruited reviewers from the IES list and received technical assistance from IES to set up their review process. Certification requires 2 days of classroom training followed by a screening test and an exercise in which participants score applications. Dr. Maynard said 25 ED staff members have also been trained.

Returning to the observation by Dr. Granger of two approaches—one researcher-driven and one practitioner-driven—Ms. Stack described the Workforce Innovation Fund, which involves a two-phase process and takes into account the limited amount of funding available. The first phase brings researchers and practitioners together to better understand efforts underway, initiatives that failed, and approaches with promise. With that initial assessment, funders can focus on a small

number of promising programs and build in strong evaluation strategies. Ms. Stack said Congress was very supportive of the idea.

Ms. Stack said Congress also strongly supported the Supplemental Security Income (SSI) Promise Program, a partnership across the Social Security Administration, the Department of Labor, HHS, and ED to provide intensive services to teenagers with disabilities. Programs aim to prepare teenagers with disabilities for postsecondary education and employment so they can transition off SSI. As with the Workforce Innovation Fund, the partnership provides a mechanism to evaluate all the current programs and strategies, consider the evidence, and develop a funding strategy. The program received a 1-year appropriation for a robust planning and evaluation effort, to be followed by more funding for implementation. Mr. Baron identified the recurring theme of drawing on the results of previous research to focus on and evaluate particularly promising strategies.

Mr. Baron asked how New York City had managed to sustain such strong interest in impact evaluation (of education and other programs). Ms. Thompson replied that stable leadership with a philosophical commitment to measuring impact resulted in buy-in from every level of municipal government. Some of New York's efforts would not have been possible without an enormous amount of private funding paired with a willingness to take a political risk (on the part of the mayor in particular). For example, such risks may not be feasible for a district superintendent facing a school board that could turn over in 6 months. Ms. Thompson said the political environment in New York City played a large role.

The Board adjourned for lunch at 12:13 p.m. and reconvened at 1:03 p.m.

The Congressionally Established CoSTEM: Developing a Strategic Plan for Federal STEM Education

Carl Wieman, Ph.D., Associate Director for Science at the White House Office of Science and Technology Policy, CoSTEM Co-Chair

The reauthorization of the America COMPETES Act mandated the National Science and Technology Council under the White House Office of Science and Technology Policy to establish the Committee on Science, Technology, Engineering, and Mathematics (CoSTEM) and charged the committee with creating a 5-year strategic plan for the federal government for STEM education. That effort included developing an inventory of existing federal initiatives to identify duplication and overlap and to inform the strategic plan.

Dr. Wieman explained that most such previous inventories are of quite limited usefulness because every agency interprets the definitions of what should be included in the inventory differently. As a result, predefined categories contain programs that can be dramatically different in focus and scope. To avoid such problems, CoSTEM engaged representatives from 13 agencies at the outset to come up with a standard set of definitions (including a definition of STEM education) to create a database of useful information. Dr. Wieman noted that CoSTEM's report on the status of the inventory has been completed but is under review.

Each program in the inventory is identified by a single primary goal and secondary goal(s), the audience served, the levels at which the program operates, the activities and products of the program, the evaluation process, the area of STEM targeted, etc. CoSTEM did not include programs that resulted from earmarks or those that receive less than \$300,000 per year.

Across the 13 agencies, the inventory identified 252 programs that receive a total of about \$3.5 billion per year. About one third of those are highly agency-specific or focused workforce training programs, such as the Nuclear Regulatory Commission's graduate student fellowships to encourage graduate studies in nuclear regulatory science. ED and NSF dominate the remaining two thirds of programs, accounting for about 10 times more money than those programs not run by ED or NSF. Thus, a broad look at STEM education should focus on ED and NSF; others such as NIH fund a relatively small number of programs.

In creating the strategic plan, CoSTEM is focusing on the primary goals of the two thirds of programs that were not agency-specific/workforce training initiatives. About a third of the programs target groups that are under-represented in STEM as their primary goal—and all of the other programs include under-represented groups as a secondary goal. Most programs served one of two audiences: teachers or graduate students. Dr. Wieman said the analysis of the inventory painted a much different picture than most people expected.

While CoSTEM was charged with identifying duplication across programs, close analysis of the inventory revealed that programs that may seem to overlap are, in reality, completely distinct. For example, it would seem at first that two programs on educational research and development would overlap, but one is an NSF program focused only on gender in STEM education, while the other is an ED program that covers educational research and development broadly.

The analysis of the inventory demonstrated that current STEM programs are all different, and Dr. Wieman said that's not so surprising, because programs labeled as STEM occupy an enormous space—from field trips for kindergarten students to support for doctoral studies. The analysis also identifies gaps so large that trying to fill them is a meaningless effort, said Dr. Wieman. We should stop worrying about filling gaps or eliminating duplication, and focus on coordinating programs to make a real impact. Where overlap exists—for example, teacher professional development programs that use different approaches—efforts can focus on developing common standards of evidence and common evaluation practices.

The inventory and analysis will inform development of the 5-year strategic plan. Agencies will come together to discuss where and how to focus programs across agencies. One step is to create basic criteria for an effective program—factors that CoSTEM believes go beyond best practices to elements that are essential for a program to work well. Everybody recognizes that there is room for improvement in evaluation, said Dr. Wieman. The challenge is to identify criteria that are specific enough to be useful and meaningful but broad enough to apply to a wide range of programs across multiple agencies.

To address the cost and complexity of evaluation, especially evaluation of programs across so many different agencies run by people with widely varying backgrounds, CoSTEM is considering proposing a centralized consulting service for educational evaluation. For example, the service would provide advice and guidance to programs on what kind of evidence to collect.

Roundtable Discussion

Mr. Baron pointed to the importance of prioritizing evaluation and research as a central component of the strategic plan. He noted that IES and the WWC are demonstrating that many widely used practices in STEM education are not supported by strong evidence that demonstrates good outcomes. He cited as examples two well-designed IES studies (of a prize-winning math software program and a math professional development program) that had few or no effects over time. He said STEM education lacks a body of validated knowledge about what works at scale. Marshaling government resources around that concept would be helpful, he said. Dr. Shaywitz agreed, saying many practices that are intuitively appealing and well intended don't work, yet that message does not get promoted and ineffective practices are used over and over. She asked how to better disseminate the knowledge about what does not work.

Dr. Bryk pointed to the medical field, noting that despite more than 3 decades of investment in research—far more than education will likely ever see—about 10 percent of practice is based on good supportive evidence, and about 10 percent of the knowledge base represents good evidence about practices that are harmful. The other 80 percent is contested territory. Dr. Bryk said current methods of research and evaluation are too cumbersome, expensive, and slow; we should question our assumptions about getting proven practices implemented in educational and social institutions.

Dr. Wieman agreed and pointed out that some programs in the inventory have no reasonable measures of impact at all. He explained the struggle to encourage agencies to consider the literature base and to build evaluation into programs without killing their enthusiasm for the programs, which requires that agencies build capacity for research and evaluation. He added that over the past 4 years, many programs have added evaluation components, although about half of those lack control groups. Dr. Wieman said CoSTEM hoped to move things further in the right direction.

Dr. Wieman noted that the country spends a large amount of money on graduate fellowship programs, mostly through NIH and NSF, presumably to entice students to pursue STEM education. However, in most cases, STEM graduate students receive funding from other sources and the fellowships provide only a fraction of support. For example, NIH research grants support three times as many graduate students as NIH graduate fellowships. Thus, said Dr. Wieman, one could ask whether graduate fellowships accomplish anything other than allowing researchers to hire more graduate students from China. He emphasized that the United States spends billions of dollars, but nobody really knows what impact many of those dollars have.

Dr. McCardle asked whether CoSTEM seeks to evaluate the programs that fund research or the findings of the funded research. Dr. Wieman said evaluation is needed for both, but the capacity required to do so differs, so developing capacity is an area of consideration.

Dr. Granger pointed to the growing perception that many publicly funded programs are so well entrenched and protected by political interests that evaluation is a waste of time, because negative findings are almost always dismissed. Instead of asking whether graduate fellowships should be discontinued, policymakers should consider how to improve their impact. As Dr. Bryk suggested, it may be necessary to come up with faster models to assess impact and evaluate improvements, said Dr. Granger. Mr. Baron noted that program offices may be more likely to buy into evaluation if it's aimed at improvement and not a threat to the program's existence.

Dr. Wieman agreed that the strategic planning process involves high-level representatives from multiple federal agencies. Therefore, to move forward, it must focus on improving programs, not getting rid of them.

Dr. Gamoran asked whether CoSTEM's report would address whether federal funding should support relatively small-scale programs that have a narrow focus (or primarily serve to generate good publicity) or should concentrate on efforts that have a wider impact. Dr. Wieman said the strategic plan will seek a middle ground because of political considerations, but also will identify areas of opportunity for real change. While CoSTEM hopes to recommend areas for consolidation, better evidence standards, and more coordination, Dr. Wieman said many agencies feel they do not have the freedom to change their programs because they are at the mercy of congressional directives on how appropriations can be used.

Dr. Ferrini-Mundy, who co-chairs the CoSTEM Strategic Planning subcommittee, explained that the strategic plan may direct agencies to refer to a list of non-negotiable best practices, identified by CoSTEM, to be used as the basis for any new investments or programs, which would prevent programs from repeatedly implementing practices that don't work. She said the strategy should require programs to accumulate relevant information through appropriate evaluation techniques to assess best practices and build a foundation for future efforts.

Dr. Ferrini-Mundy said that because NSF funds graduate fellowships, training initiatives, and research grants, there may be a natural model for studying how to improve programs and gain more clarity about the goals of funding graduate fellowships. She hoped the strategic plan would offer concrete proposals for coordination across programs to improve impact.

Dr. Herk asked whether the Board could assist CoSTEM in any way. Dr. Wieman said CoSTEM operates under some legal constraints that limit with whom it can consult and under what circumstances.

Dr. Wieman agreed with Dr. Ferrini-Mundy that CoSTEM should look closely at best practices, evaluation, and evidence so that it can recommend an approach that is reasonably good and

can be implemented. Dr. Ferrini-Mundy noted that, as Dr. McCardle pointed out, evaluating investment in research and development is different from evaluating programs (such as Space Camp). Mr. Baron noted that the interagency approach to building peer review capacity described earlier by Dr. Maynard may be a model for such evaluation.

Acknowledgment of Outgoing Board Members

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

Dr. Easton thanked outgoing members Mr. Baron, Dr. Shaywitz, and Mr. Handy for their service and presented them with plaques of recognition. He said Mr. Baron has been a tireless crusader for evidence-based policy, Dr. Shaywitz has provided both expertise and perspective from the field, and Mr. Handy had been very helpful in Dr. Easton's transition to the role of IES director.

Dr. Shaywitz said she has observed the maturation of the Board over time. She was proud to serve on the Board and particularly proud of the WWC. Mr. Baron expressed how much he enjoyed serving on the Board, noting that he appreciates Dr. Easton's receptiveness to the Board's input.

Communication of Research Findings

Jon Baron, J.D., NBES Chair, and Sally Shaywitz, M.D., NBES Member

Mr. Baron noted that Dr. Shaywitz has repeatedly focused on the importance of communicating research findings, and he asked that she describe a related project in which she is involved. Dr. Shaywitz emphasized the importance of alerting people in the field about what is effective and what lacks evidence, particularly evidence that the intervention results in a meaningful outcome. A member of her staff is gathering information from the WWC with the intention of compiling it in a manner that's easy to use. She said she would present the work to the Board for consideration if it would be helpful. Mr. Baron suggested the Board consider whether it would like to take up Dr. Shaywitz's work at a future meeting.

Continuous Improvement Research: Is It a Path for Achieving Program Effectiveness in Large-Scale Implementation?

Opening Remarks

Gilbert Botvin, Ph.D., Chief, Division of Prevention and Health Behavior, Weill Cornell Medical College

Dr. Botvin said that despite good evidence based on well-designed trials gathered over 30 years of intervention research, schools continue to make decisions about drug abuse and tobacco prevention policies without considering the evidence, sometimes implementing interventions known to be ineffective. He lamented that evidence alone is not enough to persuade people to use the interventions that work.

Research in prevention science relies on testing hypotheses, which can be considered to be an approach to continuous improvement. The prevention research cycle involves the following:

- Evaluate the epidemiology and etiology of a problem
- Develop an intervention
- Test the intervention on a small scale for feasibility, acceptability, and preliminary effectiveness
- Revise the intervention as needed
- Test the intervention in large-scale, randomized trials
- Evaluate use of the intervention
- Modify as needed
- Disseminate the intervention broadly
- Assess the impact

Dr. Botvin said the cycle can restart with new information or different problems, with the goal of continuous refinement and improvement. For example, research on the Life Skills Training (LST) substance abuse prevention program focused on the generalizability of the program and the effectiveness of the accompanying training programs. The populations studied changed over time. As the program gained a foothold, research focused on delivery, or implementation fidelity, and investigators are now looking closely at dissemination, also called diffusion.

LST is a 3-year curriculum for middle/junior high school students, with 15 class sessions in the first year and booster sessions in subsequent years to reinforce the material. It teaches general social and self-management skills to resist media and peer pressure to smoke, drink, or use drugs. LST has been provided successfully by teachers, peer leaders, and outside health professionals coming into the classroom. The program has been tested in more than a dozen randomized trials, resulting in 30 peer-reviewed publications and demonstrating that it produces sizable, sustained effects on preventing or decreasing use of tobacco, alcohol, and illicit drugs. It also has a positive effect in decreasing violence and other risky health behaviors. LST's effectiveness has been replicated by other groups.

LST is not only effective, it's cost-effective. Every dollar invested in the LST program yields a \$25 benefit, Dr. Botvin stated. Numerous professional organizations and federal agencies promote LST as a high-quality, evidence-based prevention program, yet not enough schools use it. Dr. Botvin said researchers paid particular attention to methodological issues to bolster the quality of the evidence. Early studies demonstrated positive effects, but they eroded over time, so researchers added booster sessions. The long-term effects are not as strong as the immediate effects of the program but are still significant.

A large randomized controlled trial (RCT) found that providers were equally effective whether they received in-person training or learned the program from a videotape. It also found that implementation fidelity—how the program was delivered—made a big difference: The strongest effects came from teachers that followed the protocol and teacher's manual. However, LST had positive effects on all students, regardless of fidelity. Another research team found that LST alone was as effective as LST plus a family-based prevention component.

Additional research has tested various scheduling formats and efforts to emphasize and promote program fidelity. Dr. Botvin said investigators are now studying obstacles to dissemination, adoption, implementation, and sustainability of evidence-based prevention. The results will be used to design randomized trials to enhance implementation and effectiveness. Researchers will also compare the LST program as designed with adaptations proposed by practitioners.

In response to Dr. Shaywitz, Dr. Botvin clarified that the videotaped training for teachers was a videotaped version of a live provider training session. In the past year, training has shifted more toward Web-based options, which seem to be much better than the videotaped training but have not yet been tested.

Opening Remarks

Anthony Bryk, Ed.D., President, The Carnegie Foundation for the Advancement of Teaching, NBES Member

Dr. Bryk introduced the premise that research for policy and research for practice are very different phenomena. At the heart of IES is the translational model, in which an issue is identified and an intervention is designed, tested, and disseminated. Research studies take a long time, and what is labeled as proven effectiveness often is an indication that an intervention can work—not that it will work when used by different people in different circumstances. “Adaptive integration” is the effort to understand how to make an intervention work for different populations and settings.

In contrast, the action research model begins with practitioners (e.g., individual teachers, small groups of teachers, small schools) creating interventions that are highly contextualized and locally specific. They can be very powerful, but they contain no mechanisms to accumulate or transfer knowledge, so the intervention dies with the practitioner.

The alternative to these two common approaches is quality improvement (QI) science. QI initiatives began in industry 5 decades ago and migrated to health and social services. QI takes a different approach to learning from practice through disciplined inquiry.

The foundation for continuous QI rests on a multi-level learning model in which each agent involved in practice is thinking about the nature of the work, conducting experiments, perhaps gathering data, and using data to make decisions to advance the efficacy of the institution. The big gains come when practitioners work together using common data to address common problems. Six components form the basis of continuous QI:

1. Problem- and User-Centered Work

We often approach problems by jumping to solutions without looking closely at the cause of the problem. “User-centered” refers to understanding the problem from the perspective of the user. Dr. Bryk said his organization is evaluating failure rates among community college students. The

problem- and user-centered approach requires investigators to understand the issues from the students' perspective.

2. Learning from Variation

Many interventions work in some places, and almost none of them work everywhere. Data can reveal what works where. So the central question, "What works?" is replaced by the question "How do I advance effectiveness among diverse teachers with various challenges working with different populations of students and in different school organizational contexts?" The goal is to achieve efficacy reliably at scale. Fidelity is a mechanism, not a goal, said Dr. Bryk.

3. Understand the System to Improve It

Understanding the problems means understanding the whole system. Many tools can be helpful, such as a program improvement map that lays out all the core processes that affect an outcome and organizes the processes from the most granular effects to the broadest level. Dr. Bryk noted that such mapping can help target an area for intervention; for the community college study investigators are focusing on high failure rates in developmental math classes. The map also draws attention to the interdependence among core processes, at all levels, that can hinder implementation.

Dr. Bryk presented a driver diagram, which illustrates how proposed solutions speak directly to specific causes and, if implemented, address the problem. For example, good descriptive data indicate that community colleges lose students during the transitions between courses, so one solution aims to consolidate courses to eliminate some of the transitions.

4. Accelerate Improvement with Small Tests of Change

Education researchers tend to push for quick implementation of interventions without understanding the capacity of schools to execute them efficiently at scale. The contrasting approach is to conduct rapid-cycle research and analysis of small-scale efforts, using the mantra "Learn fast, fail fast, improve fast." Over time, these small-scale efforts build up a database that provides insight into the mechanics of the process.

5. Learn From Evidence: You Can't Change What You Can't Measure at Scale

Put another way, "You can't fatten a cow just by weighing it," said Dr. Bryk. It's important to understand the process that produces the results and to develop measures to assess those processes, not just specific improvement targets. Because evidence suggests that many community college students in developmental math classes start disassociating from the instruction in the first 2 to 3 weeks of class, process measures were developed to assess student attitudes and behaviors before classes begin and 3 weeks in. This allowed investigators to assess the effectiveness of the intervention in its early stages.

Balance measures are an effort to anticipate unanticipated consequences. The system for accountability under NCLB resulted in so-called "bubble students" (those just above the

proficiency threshold who are ignored) and dramatic increases in cheating, consequences that could have been predicted.

A practical measurement framework, including balance measures, ensures that measures are integrated into the day-to-day work of a system and provide meaningful feedback. Increased use of technology by students and teachers provides an opportunity to capture abundant data unobtrusively. Researchers can take advantage of interactive technology by incorporating very short surveys into, for example, an online homework exercise, which forces the researcher to limit questions to only those that provide informative data.

Another element of measurement is rapid analytics. For the community college intervention, researchers aim for a 48-hour turnaround time. Data are collected at the end of a week of lessons, analyzed, and fed back within days to program developers, who implement the findings into the lesson plan for the next round of instruction 30 days later. Researchers can use the opportunity to learn from natural variations by asking what changed and why, instead of seeing the variations as noise.

6. Accelerate Improvement: Tap the Wisdom of Crowds

Technology spurs extraordinary capacity for innovation when large groups of people attack the same problem. Taking advantage of crowd wisdom requires some infrastructure with common definitions, tools, and measures, but it can become a powerful resource for learning to improve. Education is extremely well structured for crowd sourcing because it involves hundreds of thousands of classrooms full of people trying to improve every day. They can be harnessed to provide a framework for better understanding problems and systems, conducting rapid cycles of study and analysis, and mapping out the terrain in ways that a large-scale study cannot.

Roundtable Discussion

Mr. Baron observed that Dr. Bryk's approach involves convening learning communities around a problem to develop, test, and refine practices for continuous improvement, while Dr. Botvin's work reflects an effort to build on evidence of impact with successive research on variations to make the intervention more scalable. He asked Dr. Bryk to describe the relationship between the two approaches, and Dr. Bryk replied that they have different goals. Continuous QI encourages individuals to learn how to use evidence to improve what works, as opposed to the typical approach of implementing a program designed and tested elsewhere to improve practices in your own setting, which often fails. To achieve efficacy reliably at scale, the intervention must be adapted to the context, and that adaptation typically shows up as a failure to implement with fidelity.

Mr. Baron pointed out that it's much easier to measure successful improvements and implementation with products than it is with people. Dr. Bryk noted that continuous QI tracks progress in short increments. For example, baseline data show that 5 percent of community college students who take developmental math classes achieve their math credits within 1 year. The initial target of Dr. Bryk's community college intervention was to double that percentage, but

that goal can easily be revised upward. The ultimate goal is the ability to replicate the intervention. Continuous QI focuses both on building an effective intervention and improving the intervention at the same time.

In response to Dr. Botvin, Dr. Bryk said work such as the developmental math initiative draws on research funded by NSF and IES on, for example, good mathematics instruction and well-designed homework. In the sense that continuous QI pulls together research that focuses on specific pieces of the puzzle, it is translational. Dr. Botvin said the approach seems to presuppose that an intervention works. Dr. Bryk agreed, but said there's a reciprocal dynamic that allows research to learn from practice and to generate hypotheses to explain unexpected results.

Dr. Granger noted that the discussion highlights how much continuous QI differs from the traditional NIH/IES research paradigm. Each has strengths, and we have not yet learned how to combine them to capitalize on their mutual strengths. The traditional approach produces strong, internally valid estimates of the efficacy of various interventions over many years, but many interventions fall by the wayside before they get that far. The continuous QI approach begins with the practitioner and uses rapid research cycles so that the users are invested in the solutions. In the latter case, the various practitioners borrow from each others' work, and some measures must be in place to distinguish real change from random variation.

Dr. Granger said the challenge for IES is to recognize the opportunities and limitations of each approach and build on the strengths of each. To some extent, i3 relies on the traditional model but puts research into practice in different settings. The RELs may have an opportunity to incorporate the continuous QI approach.

Dr. McLeod said she led a district-level team in implementing the rapid plan-do-study-act model and it worked great, but it requires stable leadership and commitment to the process. Dr. Gutierrez said she appreciates the fact that continuous QI builds on previous work but adds there is often something missing: an understanding of the problems of practice in situ. It's also a hybrid that brings various pieces together in a powerful way, similar to a change laboratory. Dr. Bryk emphasized that continuous QI requires a big, long-term cultural change to get people in education, who have always worked in silos, to form networks, develop common measures and frameworks, and share information. Progress requires such a framework.

Dr. Maynard emphasized that the two approaches are not competitors but rather complement each other. She hoped they were not being pitted against each other, because each may have advantages for answering certain questions. Dr. Bryk agreed, noting that he is concerned about the perception associated with ED that randomized trials are the only way to learn, which Dr. Maynard said is not the case. Dr. Long said it's unlikely that the answer lies on the extreme ends of the spectrum; it is more likely that there's a middle ground, and she believes these two approaches overlap.

Dr. Long strongly agreed with the need to be user-centric and consider the student's perspective in education initiatives. She wondered whether faculty and school officials have an incentive to change their longstanding approaches to their work. In that respect, said Dr. Long, researchers should consider the perspectives of faculty and school officials as other users of a given program.

Dr. Easton felt the continuous QI approach could provide a good framework for assessing development and innovation.

Dr. Albro noted that some interventions don't show immediate results but do produce measurable outcomes over time, so she was concerned about relying on rapid-cycle data. She asked in what context the approach would be most effective and where it might be harmful. Dr. Bryk said the goal is to change complex systems, and every intervention will affect the system, sometimes in unintended ways. The timetables for measurement are arbitrary, but the goal is to measure something that moves over time.

Dr. Herk asked how well short-term and intermediate measures predict longer-term outcomes. If our metric is improvement on short-term measures that are in fact not linked to improvement in the longer-term desired outcomes, do we run the risk of promoting practices that are not truly effective in the long run? Dr. McLeod said the short-term measures are often disconnected from the long-term goals, which can be confusing for the practitioners involved in continuous QI. She emphasized that the approach requires adults—not children—to do their work in a very different way. Dr. Bryk said the approach pushes researchers and practitioners to learn faster; in education, many people take it for granted that processes cannot move any faster than they do now.

The Board adjourned at 3:25 p.m. and reconvened at 3:37 p.m.

Election of Board Officers

Jon Baron, J.D., NBES Chair

Mr. Baron said that, on the advice of ED's general counsel, the Board would elect new Board officers in an open session but still use secret ballots. He said 7 of the 11 current Board members would return at the next meeting. Mr. Baron, Mr. Handy, and Dr. Shaywitz will complete their second terms on the Board as of November 28, 2011. Dr. Gamoran will complete his first term on the Board on November 28, 2011. The White House has re-nominated him for a second term, so it is likely, but not guaranteed, that he will return to the Board. At its next meeting, the Board may have two or more new members. Mr. Baron raised the concern that a chair and vice chair elected at the current meeting may not represent the wishes of the future Board. Therefore, he proposed an interim election of a chair and vice chair with the understanding that when new Board members are confirmed, the Board will elect new officers.

Dr. Gamoran proposed a minimum 1-year term but did not support a single-meeting term. He added that all but two of the current members are serving terms that will end in November 2012. Dr. Granger pointed out that it takes about two meetings to acclimate to the Board; a chair

serving an abbreviated term may not feel very productive. Dr. Underwood suggested a 1-year term with no term limit. Dr. Long said new Board members may not be ready to elect new officers at their first meeting. Dr. Gamoran added that by October 2012, the Board would know who had been reappointed.

Mr. Baron confirmed the consensus of the Board to elect a chair to a 1-year term with no term limit. He then pointed out that the Board is not required to have a vice chair. The position was created as a compromise to ensure broad representation of viewpoints across the Board when all of the members were new. Dr. Gamoran noted that a vice chair can fill in for the chair if needed, and Dr. Long said the two leaders can provide balance by ensuring the perspectives of both researchers and practitioners are represented among the leadership. Mr. Baron said having a vice chair has been helpful for him.

Dr. Gamoran moved that the Board continue to have a vice chair and that the next vice chair serve a 1-year term as well. Dr. Long seconded the motion, and the Board unanimously approved.

Dr. Underwood nominated Dr. Long for the position of chair, and Dr. Bryk seconded the nomination. In the absence of any other nominations, Mr. Baron declared that Dr. Long was elected chair.

Dr. Gamoran nominated Dr. Gutierrez to serve as vice chair. The nomination was seconded, and Dr. Gutierrez accepted. Mr. Baron noted that both Drs. Long and Gutierrez are researchers. Dr. Gamoran was not convinced that the chair and vice chair must represent both sides (research and practice), because the Board works smoothly and allows all voices to be heard. Mr. Baron noted that having both camps represented helps appearances. Dr. Granger nominated Dr. McLeod for vice chair, and she accepted. At Dr. Granger's request, the two nominees left the room for the ensuing discussion.

Dr. Long said that she and Mr. Baron had different perspectives and opinions, and they complement each other. They had very different opinions about the letter to the secretary concerning the NCLB waiver, which Dr. Long said ultimately strengthened the letter. Regardless of who is elected vice chair, said Dr. Long, she intends to reach out to various members to complement her own skill set.

Dr. Granger noted that Drs. Long and Gutierrez have distinctly different research approaches, qualitative vs. quantitative, which could be a strength, as could having a vice chair from the practitioner camp. He felt that Congress took the Board more seriously with both Mr. Baron and Dr. Long representing the Board. However, he suspected that would remain the case regardless of which candidate was elected vice chair. Dr. Easton said having one person from each camp broadens the perspective. Drs. Gamoran and Long said both vice chair candidates would bring complementary expertise.

Dr. Long said that Dr. Gutierrez, as a former president of AERA, would have influence among that constituency that Dr. Long did not. She emphasized that she and Dr. Gutierrez had very different areas of research. Dr. Gamoran said Dr. Gutierrez is highly accomplished, has done important work on bilingual education, and is highly respected in her field. He added that she was a very effective AERA president, so he thought she would be an effective partner for Dr. Long. He said he did not know Dr. McLeod outside of his interaction with her on the Board but felt she had been a good contributor.

Dr. Bryk said that from the perspective of the education research community, having Drs. Long and Gutierrez in the leadership positions is attractive, but he could not judge whether it would be politically sound for the agency. Dr. Easton said that within the agency, the practitioner/researcher division is not compelling, and the overall strength of the team would not be affected. Dr. Long praised Mr. Baron's inside-the-Beltway knowledge, and Dr. Gamoran thought Dr. Gutierrez to be very knowledgeable in the same area, having been a member of President Obama's transition team.

The nominees returned to the room and the Board undertook a secret ballot for the position of vice chair, with Dr. Ball calling in her vote by phone to Dr. Herk. While the results were tabulated, Mr. Baron thanked Ms. Lucier for her assistance with logistics and travel and also thanked Wilma Greene, who serves as liaison with the contractor for meeting logistics.

Mr. Baron confirmed that Dr. Gutierrez was elected vice chair and that her term would begin on November 29, 2011. He again thanked Ms. Lucier on behalf of the Board.

Dr. Gamoran said that in case he is not re-nominated, he wanted to thank Mr. Baron and Dr. Easton for the increased richness of the Board meetings over the past year, which has made the meetings more interesting and meaningful. He appreciated the openness of both Mr. Baron and Dr. Easton. Dr. Gamoran said IES remained stable despite the change of administrations, which speaks to the strength of IES leadership and staff. Mr. Baron added that under Dr. Easton, IES essentially created a new enterprise that has had an impact on policy and practice.

NBES Annual Reports: How to Ensure Their Independence and Usefulness, Consistent with Congressional Authorizing Language

Introduction

Jon Baron, J.D., NBES Chair

Mr. Baron said the Board is required by statute to write an annual report assessing IES. Previous reports have been largely descriptive and based on submissions from IES staff. He asked the Board to consider whether the annual report should be more substantial, noting that Dr. Herk could play a primary role in drafting the report.

Opening Remarks

Monica Herk, Ph.D., Executive Director, NBES

Dr. Herk asked the Board to consider what process it should use if it would like to have more input in writing the annual report for 2012. To meet the deadline of July 1, 2012, the final draft must be completed by June 1, 2012, to allow for review and printing. She proposed two potential processes:

Possibility A: The Board breaks into subcommittees that would respond by the February 2012 meeting to the four areas identified in the legislation: carrying out scientifically valid research, conducting unbiased evaluations, collecting and reporting accurate education statistics, and translating research into practice. Alternatively, the subcommittees could report on the work of each of the four centers: NCER, NCEE, NCES, and NCSE. To comply with the Federal Advisory Committee Act, subcommittee members can confer among themselves, but substantive discussions and decisions must take place in an open meeting, such as the February 2012 Board meeting. Subcommittees could present their findings, and IES staff could gather input, revise the documents, and combine them into a draft for Board review at the June 2012 meeting.

Possibility B: The Board writes the annual report as a committee of the whole, making decisions on the issues at hand at the February 2012 meeting. In that case, it may be necessary to have a meeting in May to make the June 1 deadline.

Roundtable Discussion

Dr. Long said that with the Board losing members, she worried about the subcommittee approach. She noted that the Board receives updates from the Commissioners at every meeting. She suggested convening working groups to review updates from the Commissioners, capture the Board discussion over the past year, and acknowledge issues about which the Board had strong feelings. The resulting draft could be the basis of the annual report. The Board would then gather feedback to the draft during a Board meeting (without taking up a lot of meeting time).

In response to Dr. Granger, Mr. Baron said the Board report has not been substantially different from the IES biannual report. Dr. Granger suggested that Dr. Herk draft a summary for the February 2012 meeting on the basis of Board discussions over the past year. Dr. Long agreed that Dr. Herk could draw from the summaries of past meetings.

Mr. Baron said the Board can advance the goals of the law that authorized it and the purposes of IES by making recommendations to Congress and the secretary that reflect voices outside of ED. The Board also brings up ideas during meetings to explore ways to improve, for example, the IES peer review process. Mr. Baron suggested the annual report include the Board's resolutions to Congress and capture some of the discussion around key issues of interest, such as peer review. Such a report may be useful to IES and others.

Mr. Baron clarified that the Board's annual report reflects its thinking about how IES is doing. Dr. Granger raised concerns that it would be difficult and labor-intensive to summarize the wisdom of the Board around peer review. Mr. Baron said the report could identify the main ideas raised by the Board, and the draft would be circulated to allow members to weigh in.

Dr. Easton said the annual report could provide backup for the Board's recommendations for IES reauthorization, but to date the report has been a perfunctory activity. He suggested that the Board take an incremental approach, perhaps focusing the upcoming report on reauthorization—evaluating the strengths and weaknesses of the agency. Dr. Granger said such an assessment would require an intensive approach, and Dr. Bryk said he would need a better sense of the perspective from the field for such an assessment. Dr. Herk noted that the Board has a budget that could support a survey of the field, if desired, to inform a longer term approach.

Dr. Long spoke in favor of an annual report that reflects the themes of the Board's discussions: how to support innovative research with meaningful outcomes and how to translate research into practice (including discussion about the WWC). Dr. McCardle noted that such information is reflected in the Board's minutes, which have already been reviewed and approved by Board members.

Dr. Bryk said he was uncomfortable characterizing today's discussion of peer review as a review of the process. Dr. Gamoran said it would be appropriate to say the Board provided input. Dr. Long said the discussion of peer review was one part of a continuing conversation about supporting innovative work that will have an impact—an issue that cannot be solved in one meeting. Dr. Gamoran said the Board is directed to assess the effectiveness of IES, but the mandate does not rise to the level of evaluating the operation. Dr. Granger felt the intent of the legislation was to position the Board to oversee and guide IES.

Dr. Granger reiterated that Dr. Herk, as Executive Director, should draft the annual report rather than IES staff and that she should provide an outline of the report to the Board for discussion at the February meeting. Dr. Herk noted that any Board member may email her, and it is not considered a deliberation. Because NBES is an advisory and not a governing board, said Dr. Underwood, he hoped the report would demonstrate that the Board has at least considered the issues and approved the priorities of IES.

Closing Remarks and Adjournment

John Q. Easton, Ph.D., IES Director, and Jon Baron, J.D., NBES Chair

Dr. Easton said he sees the Board as a resource for IES. The members do not need to agree unanimously but should represent different points of view, as they did during the discussion of research. Dr. Easton looks to the Board to integrate a broad sense of ideas within the agency.

Mr. Baron offered his continued service to the Board, for example, as a liaison with Congress, who understands the committee structure and can follow up on issues of concern. He adjourned the meeting at 4:31 p.m.