

National Board for Education Sciences Meeting

June 29, 2011
Washington, DC

Meeting Summary



Location

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

Participants

National Board for Education Sciences (NBES) Members Present

Jonathan Baron, Chair
Bridget Terry Long, Vice Chair
Deborah Loewenberg Ball (via telephone)
Adam Gamoran
Kris D. Gutierrez
Margaret R. McLeod
Robert A. Underwood

NBES Members Absent

Anthony Bryk
Frank Philip Handy
Sally E. Shaywitz

Ex-Officio Members Present

John Q. Easton, Director, Institute of Education Sciences (IES)
Jack Buckley, Commissioner, National Center for Education Statistics (NCES)
Rebecca Maynard, Commissioner, National Center for Education Evaluation and Regional Assistance (NCEE)
Peggy McCardle, National Institute of Child Health and Human Development (NICHD)
Lynn Okagaki, Commissioner, National Center for Education Research (NCER) and Acting Commissioner, National Center for Special Education Research (NCSER)
Dixie Sommers, Bureau of Labor Statistics (BLS)
Joan Ferrini-Mundy, National Science Foundation (NSF)
Robert Kominski, U.S. Census Bureau

NBES Staff

Monica Herk, Executive Director
Mary Grace Lucier, Designated Federal Official

IES Staff

Elizabeth Albro
Sue E. Betka
Chris Chapman
Matt Devine
Erica Johnson

Laura LoGerfo
Ellie McCutcheon
Audrey Pendleton
Anne Ricciuti
Allen Ruby
Stephanie Schmidt
Marilyn M. Seastrom
Marsha Silverberg

Invited Presenters

John W. Wallace, former Vice President for External Affairs, MDRC
Eric Bettinger, Stanford University School of Education
Robert Slavin, Johns Hopkins University and the Success for All Foundation

Members of the Public

Christine Talbot, American Education Research Association (AERA)
Lauren Gibbs, Success for All
Virginia L. Neale, New Mexico State University
Judy L. Johnson, Reading Recovery Council
Jean Gossman, Education Daily
James B. Williams, Department of Education
Sarah N. Hutcheon, Society for Research in Child Development
Jim Kohlmoos, Knowledge Alliance
Sarah D. Sparks, Education Week
Kim Bromann, Coalition for Evidence-Based Policy
Karen Studwell, American Psychological Association (APA)
LaTosha Plavnik, Consortium of Social Science Associations (COSSA)
Michael Ross, NCGS
Shukurat O. Adamoh-Faniyan, Knowledge Alliance
Gerald Sroufe, AERA
Philip Chalker, American Association for the Advancement of Science (AAAS)

Wednesday, June 29, 2011

Call to Order, Approval of Minutes, Swearing in of New Members, Chair Remarks

Mr. Jon Baron, Chair, NBES

Mr. Baron called the meeting to order at 8:32 a.m. Ms. Lucier took a roll call of Board members and ex-officio members and noted that a quorum was present. Other meeting participants and attendees also introduced themselves. Dr. Gutierrez, who was not present at the previous meeting, introduced herself to the Board. She was sworn in as a new member.

Mr. Baron gave an update on items discussed at the last meeting, including a presentation on the proposed Advanced Research Projects Agency for Education (ARPA-ED). ARPA-ED is still in the President's FY 2012 budget, and several members of Congress are working on the authorizing legislation. The Administration's draft authorizing language is being held up by the Office of Management and Budget (OMB) due to questions from IES. In the current budget climate, it is uncertain whether the proposal will go forward.

The Regional Education Laboratories (REL) received funding, but not full funding, in the FY 2011 appropriations bill, thereby allowing the recompetition for the next round of REL contracts to go forward. The Board's March 23 recommendation regarding the RELs received positive responses from Congress. The extension covered by the contract recompetition is 5 years.

The Board's March 23 recommendation to advance the development and use of credible evidence and to ensure program effectiveness in federal education programs has been well received. The recommendation led to a number of meetings with key congressional staff and with the Department regarding the recommendation's application to specific programs, such as School Improvement Grants (SIGs), formula grants, competitive grants, and No Child Left Behind. These discussions continue, and meetings with members of Congress may follow.

Mr. Baron called for a motion on the previous meeting's minutes. Dr. Gamoran moved approval. Dr. Gutierrez seconded, and the motion carried unanimously.

Update: Recent Developments at the IES, Including the National Center for Education Evaluation and Regional Assistance (NCEE), the National Center for Education Research (NCER), and the National Center for Special Education Research (NCSER)

Dr. John Q. Easton, Director, IES

Dr. Easton announced that Dr. Okagaki is leaving IES to become the Dean of the College of Education and Human Development at the University of Delaware. Dr. Elizabeth Albro will become the Acting Commissioner of NCEE. Dr. Joan McLaughlin will be the Acting Commissioner of NCSER. Dr. Easton will soon be able to announce the new permanent NCSER Commissioner.

Applicants' proposals for the REL recompetition were due the day of the meeting.

The fiscal year (FY) 2011 budget brought some reductions for IES. The FY 2011 IES budget is approximately \$50 million lower than the FY 2010 budget. In response, IES reduced budgets for REL activities, NCSER, and the State Longitudinal Data System (SLDS) grant program. Dr. Easton expressed disappointment but voiced the expectation that over the long run IES would be able to recover from the FY 2011 reduction.

Dr. Easton reported that since the last NBES meeting, five National Assessment of Education Progress (NAEP) reports had been released: a high school transcript study, a report on civics and history, a Hispanic/White achievement gap report, a geography study, and a study of how state tests map to NAEP. The results of the three social studies NAEP reports – on civics, history, and geography – showed a strong pattern. This year, there will be 15 NAEP reports, which may be a record. There will be fewer NAEP reports next year.

Dr. Easton said he had visited an IES-sponsored summer institute on designing and conducting randomized field control studies as well as one on single case design. Dr. Easton remarked on how impressed and proud he was of what he observed at these institutes, which are developing the kind of broad expertise necessary for effective educational research.

Mr. Baron opened the floor for discussion and asked for further detail about the summer institutes. Dr. Easton said the workshops had about 35 people each. Most participants are new Ph.D.s, but there are some senior participants who are there to upgrade their skills. Dr. Okagaki said that this is the fifth year of the summer workshops. Of the first cohort of attendees, one third subsequently received awards from IES to conduct efficacy or scale-up trials. From the second cohort of participants, 20 percent have subsequently received new IES grants or are working on IES-funded randomized controlled trials (RCTs). Dr. Okagaki stated that senior institute attendees are likely to return to their home institutions and teach their students what they've learned in the workshops. The institute sessions have been videotaped and are available online. New NCER staff view these tapes as part of their training. However, the videos do not include the small group projects, which are an intensive part of the actual institute. The fact that this year's institutes include some junior researchers as instructors indicates that the field is developing strong junior methodologists. Dr. Easton concluded the discussion by describing his message to the participants of these institutes: that they should put their training to practice in working on real-world problems and that they should work in partnership with practitioners and policymakers.

Dr. Okagaki gave a brief update on NCER and NCSER. Both centers are working on syntheses of the research they have funded to see what has been learned and what gaps the centers should be addressing. A synthesis on struggling readers is in the review process, and one on early childhood is ready to be sent to the Standards and Review Office. In both centers, staff have been working on improving processes and procedures, standardizing grant monitoring and budget reviews, and ensuring efficiency in main activities.

Dr. Gamoran commented that he cited the Hispanic Achievement Gap report in a forthcoming journal article. NAEP reports have proven very useful, and they show that NCES is accomplishing its mission.

Dr. Maynard gave an update on NCEE. NCEE has added staff since the last NBES Board meeting: one person to manage the *What Works Clearinghouse*, another senior staff member to work with the REL program, two administrative and mid-level support staff for NCEE's research programs, and a senior staff member who will be joining NCEE in October for analytic and technical support and methodology.

NCEE has 47 active projects in its evaluation portfolio, half of which are major RCTs being conducted by the RELs. Three new, large projects that NCEE is in charge of are evaluating Race to the Top, the Investing in Innovation (i3) Technical Assistance contract, and School Improvement Grants. Four new reports on evaluations of education interventions have been released since the last NBES meeting. Significant NCEE staff time was devoted to crisis management when the funding situation disrupted REL work. Now that the RELs have restarted their work, many work products will be coming out by the end of the year.

The *What Works Clearinghouse's* production has slowed down because more effort has been devoted to planning the next wave of improvements: a major website redesign, including the Find What Works button, and a sorted list of every study that NCEE has fully reviewed for one of its intervention studies.

Dr. Maynard stated that strategic planning for the National Library of Education had not proceeded as quickly as she had hoped and that it will be one of her priorities in the next 6 months. Libraries are changing dramatically, and it is not yet known what kinds of services and staff will be needed.

Mr. Baron asked how well it is working to have the i3 grantees either do their own evaluations or hire their own evaluator. Dr. Maynard said the quality of the evaluations will vary. The i3 evaluations are funded at reasonable levels and most of them have credible, professional evaluators attached to them. Dr. Easton commented that he had recently attended a session for i3 evaluators run by ACT, Inc., NCEE's technical assistance contractor for the project, and at that session he saw strong evaluation designs from strong contractors. Dr. Maynard said the technical assistance to i3 evaluators should be responsive to the evaluators' needs. Mr. Baron asked whether the i3 model – where a grantee has sufficient resources to hire a good evaluator and IES provides technical assistance – is a viable model. Dr. Maynard said that it depends on the quality of the evaluator and the strength of the partnership between the program and the evaluator.

Mr. Baron introduced a resolution thanking Dr. Okagaki for her service to the nation, IES, and the Department in light of her dedicated service and many contributions to NCER, NCSE, and IES. Dr. Gamoran seconded the motion and it carried unanimously.

National Center for Education Statistics (NCES): Linking NCES and State Data, and Other Initiatives to Create a More Comprehensive Portrait of U.S. Students and Schools

Dr. Jack Buckley, Commissioner, NCES

Mr. Baron explained that beginning with this meeting, each subsequent meeting will include an in-depth look at one of the IES centers. This meeting's report was from NCES. Dr. Gamoran and Dr. Buckley had developed a set of questions to frame the session:

- What steps is NCES undertaking to explore possible linkages of national survey data with state administrative data systems at both the school and student levels?
- What is the long-term plan for the longitudinal study series, and how well will this plan provide a consistent, ongoing portrait of U.S. schooling? Can the various grade-level longitudinal studies ever be integrated into a continuous look at a cohort?

To begin addressing these questions, Dr. Buckley discussed:

- current longitudinal studies (K–12);
- challenges to the current longitudinal sample surveys;
- current studies including the High School Longitudinal Study of 2009 (HSL:09), the Early Childhood Longitudinal Study and Kindergarten Class of 2010–2011 (ECLS-K:2011); and
- possible future innovations to address the issues mentioned in the framing questions (e.g., sequencing the studies, use of administrative data, and a potential middle school study).

Dr. Buckley commented that the existence of vastly more administrative data than there was a decade ago changes the nature of NCES's work.

HSL:09 started in 2009 with data from 21,000 ninth graders in 944 schools. ECLS-K:2011 started in the fall of 2011 and is the second kindergarten cohort study. Follow-up data collection is planned each year (in first through fifth grades) for ECLS-K:2011. For HSL:09, the first follow-up will be in eleventh grade, with a twelfth grade update on the participants' transition to postsecondary education or to the workforce. HSL:09 also has a follow-up 2 years after graduation, and transcripts are collected in between. ECLS is ambitious, with two data collections per year in first and second grades and annual collections through fifth grade.

Dr. Buckley stated that NCES longitudinal surveys have advantages and disadvantages:

- Advantages
 - Rich, deep data for each student, including behavioral, attitudinal, and cognitive assessments
 - Multiple data sources for each student, including the student, parents, teachers, school administrators, and school counselors
 - A large and nationally representative sample
 - Subsamples of sufficient size to be able to generalize regarding significant subgroups in the population

- A focus on key decisionmaking and transition points
- Disadvantages
 - A sample, not the universe, so generalizations may not be possible about some subgroups
 - Cost is expensive and rising
 - Limited coverage of the cohort
 - Progressively harder to conduct

Recruiting schools is a major challenge. Schools' response rate has decreased significantly. Schools that enter the study tend to stay in the study, and student response rates are good, but the current assessment and data collection burden on schools makes it hard to secure participation. NCES conducts a non-response bias analysis to ensure a nationally representative sample and sufficient subsamples, but there is always the possibility of bias in key covariates of interest in the studies. Schools' reluctance to participate affects the cost of conducting these surveys and also contributes to rising rates of attrition.

Because there is no single listing of U.S. students, NCES uses stratified random samples and cluster-sampling of schools to conduct its longitudinal surveys (i.e., schools are randomly selected and then students within the selected schools are randomly selected). Although this approach is cost efficient, it does not create ideal variance in the sample, as compared to a direct random sample of all U.S. students. However, the cost advantages of this approach are lost when students who were initially clustered in the same school transition to multiple schools (e.g., in the transition from middle to high school). This creates a tradeoff between study cost and ideal research design for these studies. This tradeoff, in turn, has led to shifts over the years in the starting points of these studies, especially at the secondary level.

There has not been a longitudinal study covering the middle school years. Given the recruitment and attrition issues, it is not economically feasible to study the same survey panel for its entire education and beyond. Instead, NCES can line up longitudinal studies (e.g., K–5, 6–8, and 9–12) so that the same cohort is covered through multiple panels. Because there already are K–5 and 9–12 panels planned, the next step is to bridge the middle school gap by timing a middle school panel so that it would cover the same educational cohort as the elementary and high school panels. The high school panel, in turn, would then be linked as closely as possible to the National Postsecondary Student Aid Study (NPSAS).

NCES studied the relative effectiveness of integrating sequential panels of the same cohort through using a continuation sample (i.e., enrolling a subsample of respondents from the first panel into the subsequent panel) versus using retrospective surveys, administrative data, and imputation. It found that using retrospective data plus imputation is as effective as fielding a 20 percent continuation sample, but at a much lower cost.

Dr. McLeod asked a follow-up question about cost. Dr. Buckley said the cost is low because IES already sponsored four waves of the SLDS grant program, which helped states build the kind of

data NCES would collect. The issue is navigating the practical, legal, and policy barriers to sharing the data—in particular, ensuring compliance with the Federal Education Rights and Privacy Act (FERPA).

Roundtable Discussion

Mr. Baron opened the floor for discussion. Dr. McCardle urged NCES to follow through on the middle school proposal, because girls move away from science between fourth and eighth grade, and we need middle school data to understand why. She endorsed NCES's desire to link to administrative data and noted the parallels between the challenges that NCES faces with FERPA and that the National Institutes for Child Health and Development (NICHD) faces with HIPAA. The ideal would be to have data from the real-time cohort and the retrospective administrative data, and the medical records. Dr. McCardle emphasized the importance of collecting data during the early adolescent period.

Dr. Gamoran noted that NCES's series of longitudinal strategies is one of the main components of the Center's strategy. Dr. Gamoran's first comment was to urge NCES to revise its college update plan (for HSLS:09) so that rather than skipping twelfth grade it includes a twelfth grade survey that is consistent with the preceding four waves of high school longitudinal studies. Understanding student postsecondary trajectories relies on capturing the high school data. It does not make sense to skip twelfth grade. First, it is the last year of high school, and second, surveys going back 40 years have captured twelfth grade data.

Dr. Gamoran's second point was that because the federal government has invested significant funds in the states for them to develop administrative data sets, it is important to devote as much effort as possible to making the data sets useful for more than just reporting state test scores. There has been progress in making these administrative data more available to researchers. Dr. Gamoran stated that consistent pressure and leverage from NCES will help in overcoming the remaining data access barriers that researchers face. Dr. Gamoran noted possible uses of the data, such as obtaining indicators of where schools are in their state's accountability system. The best scenario would be to have year-to-year achievement data cross-matched with data from the Early Childhood Longitudinal study. That would open a new world of analysis and should be the ultimate goal.

Dr. Gamoran's third point was to ask whether NCES had considered incorporating new data-gathering technologies such as video or beepers into its data collection.

Dr. Buckley responded to Dr. Gamoran's second point by stating that it remains an IES-wide mission to improve the use of NCES's large sets of longitudinal data and to get these data into the hands of state education agencies, school districts, and schools. For example, the designs developed by both Common Core Data Standards consortia include formative assessments, collected using improved technology. These data should be fed back to the classroom level to help teachers target students' weak points. And this information should be housed in the data

system the Department helped the states to build. RELs are building these sorts of activities into the core of their mission.

In response to Dr. Gamoran's third point, Dr. Buckley stated that the Trends in Mathematics and Science Study (TIMSS) did a video study that yielded interesting research. Video studies are expensive due to the coding requirements. However, the Gates Foundation recently funded work on teaching (that included video studies), and since technology has improved, NCES is looking at the possibility of automated coding of such studies.

Dr. Long noted difficulties that researchers have accessing state administrative data. Different states interpret FERPA differently. The Department has tried to clarify FERPA. Dr. Long suggested that it may serve the public good to invest the funds necessary to settle the legal issue.

Dr. Buckley responded by first summarizing the FERPA-related challenges to accessing the data. State attorneys general and other actors have adopted varying interpretations of the Act. The Department has issued clarifying regulations, most recently in 2008. The Department is currently working on the latest clarification of FERPA. IES is involved in helping respond to the latest round of public comments. The re-regulation will be completed this year.

Dr. Long said funding allocated for building databases in the states should be conditional upon states providing access to the data. IES should look at how they can help states reduce costs associated with providing access to administrative data. Possible approaches include understanding the processes specific states use and who is already receiving the data. Dr. Long described Harvard's strategic data project, which works with school districts to help with capacity and translation of the materials. Dr. Long urged NCES to consider cost-sharing and partnerships with outside organizations.

Dr. Long supported the use of administrative data to reduce the cost of other data collection (such as for the longitudinal surveys). Regarding survey panels that are not followed into high school, Dr. Long suggested accessing the National Student Clearinghouse for potential matches. Dr. Buckley responded that NCES is already utilizing the National Student Clearinghouse data in that way.

Dr. Kominski said all samples in any longitudinal study suffer attrition. He suspected that state administrative records would have fairly standard types of educational data, but the quality of data collection will vary across states, which will lead to gaps in the data system. More importantly, administrative data will not capture interesting questions such as how students get along with their teachers. A larger problem is students who disappear from state data systems (because they move or otherwise are not located). Given these types of issues, Dr. Kominski wondered at what point it becomes more cost-effective to simply conduct a third follow-up interview, as opposed to trying to reconstruct what happened from incomplete state administrative data.

Dr. Kominski also raised the question of how many levels a survey can be representative of: for example, a national survey may be able to be representative of gender, racial, and ethnic subgroups, but probably could not generate state-level conclusions for each of the 51 states.

Dr. Kominski's final point was that when data are collected, they are used to answer questions not initially built into the survey's design. He urged NCES to think about linking its data to non-educational administrative data, such as Bureau of Labor Statistics (BLS), Department of Labor, and Census data to help understand the relationship between individuals' education and their adult outcomes.

Dr. McCardle said linking the education administrative datasets and other records will enrich the longitudinal studies, creating a more complete picture of the developmental processes. Dr. Buckley said his intent is to collect every legally available piece of administrative data, but it is impossible to collect enough information for every research goal. Dr. Kominski agreed that, with a limited budget, trying to meet everyone's desires would be dysfunctional. Sample attrition is a serious problem, and NCES should exploit existing records to address the problem. He emphasized the importance of keeping the sample from disappearing due to attrition.

Dr. Okagaki noted that adding a year to the data collection may not be economically feasible and asked whether using technology reduces the data collection costs. Dr. Buckley responded that technology development takes time to pay for itself through reduced data collection costs. However, technology advances so quickly that it quickly becomes obsolete. With regard to using administrative data to reduce costs, NCES hopes to realize cost savings over the long run, but administrative data also have costs associated with checking, cleaning, and inputting the data.

Dr. Gamoran noted that High School and Beyond and the National Education Longitudinal Study (NELS) had high sample retention rates, although dropouts were difficult to follow. High-quality preservation of samples is a hallmark of NCES. Responding to Dr. Kominski, Dr. Gamoran stated that administrative records, such as the transcript studies that are part of HSLS, are valuable in answering why some people succeed and others fail. With respect to adding another wave to the HSLS, Dr. Gamoran noted that there is already a plan to do limited summer follow-up after twelfth grade, so funding is already set aside for a twelfth grade follow-up. Dr. Gamoran noted that given finite funding there will be tradeoffs regarding data collection, and that what is collected will ultimately be a resource-based decision.

Dr. Easton asked Dr. Buckley to speak on the Data Forum in relation to state data use and research access. Dr. Buckley said NCES has been bringing state and local education agencies together annually since the 1970s to help build the National Education Data Model. More recently, the activities around the SLDS have expanded to include technical assistance. The Privacy Technical Assistance Center (PTAC) provides help on the FERPA question. The Educational Technical Assistance Program (EduTap) assists in data structuring issues. At annual meetings, NCES learns from the communities, and the communities learn from each other.

Dr. Long noted that the state-level people building the data and supporting research are not the people making the decisions on linking data sources. Dr. Buckley said they are starting to be the same people. Part of the mismatch of priorities is due to FERPA regulatory obstacles.

Dr. Long asked whether it will still be possible to match HSL:09 data to state administrative data for the 10 states listed in Dr. Buckley's slides. Dr. Buckley indicated that because the matching depends on retrospective state administrative data, the matching can still be done later despite current delays.

Dr. Gutierrez emphasized the need for a middle school survey. In addition, she emphasized the continued importance, as the population grows more diverse, to continue oversampling subgroups. For example, the data may give a better understanding of the trajectories of English language learners (ELLs). The meeting recessed from 10:32 to 10:47 a.m.

The "Big Picture" Discussion

Mr. Baron introduced this topic, which was to be a roundtable discussion on two questions:

- Within the overall Board-approved IES research priorities, what are the most important and compelling research questions and topics to address?
- Among these, where are the gaps in knowledge the greatest and most serious?

Each of the members had been invited to prepare informal remarks.

Mr. Baron started with his own comments. He stated that increasing the relevance of IES research should be a priority. IES must be able to offer school districts and officials programs and strategies that they can implement with confidence that the strategies and programs will have a meaningful impact. For IES to be able to do that, there must be well-conducted, preferably randomized, impact evaluations that have been implemented in real-world school settings and shown sustained, sizable replicable impacts on important outcomes. Increasing the number of credible examples of programs with this type of demonstrated effectiveness should be a key organizing goal for IES. Some suggestions of ways IES can think about building these examples are noted below:

- Improve the chance that a chosen model or strategy has a strong chance of working, NCEE could look for credible pre-trial evidence of effectiveness before launching a larger experiment.
- NCER and NCSEER could encourage researchers to partner with entrepreneurs on research grants in order to facilitate scale-up and implementation.
- Allow ideas that lack scientific "glamour," such as book fairs or tutoring, to compete in the research process, if there is a likelihood that the approach will have meaningful impacts.
- Substantially increase the number of good experiments and quasi-experiments that schools, districts, and states undertake through the use of low-cost administrative data.

Dr. Gamoran responded to the questions as follows:

- IES should recognize that not all interventions work in all contexts and should emphasize greater nuance in intervention research, understanding the complexity of variation in context and participants. IES should be more systematic about promoting such diversity in its intervention research.
- IES should emphasize implementation research, by giving more scientific scrutiny in the study design to organizational conditions that support or impede implementation. Education researchers should be as deliberate about designing the scale-up and implementation of interventions as they are about designing the interventions themselves.
- Among possible topic priorities for IES research, Dr. Gamoran said that if he were forced to choose one it would be greater understanding of teacher quality.

Dr. McLeod responded to the questions from a practitioner perspective:

- Regarding process-
 - Entities receiving IES funding should have a proven track record of having worked with a district or state, so that they know what the issues are.
 - Research must address the needs of the practitioners, administrators, and students.
 - Dr. McLeod agreed with Dr. Gamoran that replication research designs should aim to determine the administrative or organizational capacity elements that are necessary for successful scale-up to make it possible for programs to be applied elsewhere.
- Regarding topic priorities-
 - Develop the What Works button that Dr. Maynard described, including the ability to search existing research in a topic area, such as special education, by entering just a few keywords.
 - Conduct more research on program models, responses, and interventions that work with different types of ELLs, including those with interrupted schooling and long-term ELLs.
 - Assess issues related to ELLs, including looking at the trigger point where proficiency levels start yielding meaningful results on statewide assessments.
 - Require grantees to talk to the people who consume their research at the state and local level in order to produce findings that are more relevant.

Dr. Ball agreed with the previous comments. She emphasized the following thoughts:

- IES should focus on studying and understanding more about implementation and should understand implementation as a problem of practice in its own right.
- If forced to choose one topical priority it would be teacher quality, or “instructional quality” as she prefers to refer to it. Dr. Ball recommended systematically looking at how instruction is handled in other countries.
 - Examining studies on what it takes to have systemically high instructional quality across learners and settings, including aspects of school organization that allow teachers to do high-quality work
 - Looking at curricular resources and the way they are structured
 - Staffing, and what it takes to organize so that each student receives high-quality instruction

- Helping to broaden the field's understanding of "instruction" and instructional quality so that it encompasses more than just the individual teacher, his or her qualifications, and his or her practices and instead focuses on how the learning "transaction" happens for children

Dr. Long's stated that the most important question is, point blank, how do we help students learn:

- It is important to test the many potential ideas in education and see what works in order to establish the truth.
- Going beyond overall effects, we need to get inside "the black box" and understand mechanisms.
- What are the incentives for implementation of effective practices, and what are the capacities of different organizations and people to implement such changes?
- We need to look at context and effects on subgroups (particularly the disadvantaged, minorities, and ELLs), because not everything works everywhere.
- Nuances and interactions matter; researchers like to isolate one input to a situation, but in fact the effect of that one input probably depends on the other inputs it is interacting with, both inside and outside the school.
- IES should think about the entire education pipeline, including students showing up for kindergarten school-ready and higher education.

Dr. Long stated that IES's most important criteria in its peer reviews and funding decisions should be the proposal's relevance to increasing the public and private benefits of education and the research's potential impact. In addition, IES's role should be to help fill the holes in education research that the "free market" in education research has not addressed. One way that IES might do this is to encourage researchers to coordinate and create partnerships, especially across disciplines.

Dr. Underwood named three priorities:

- Emphasizing teacher quality and preparation, because the intervener is an important part of the intervention
 - Not only what works but with whom it will work
- Improving the education pipeline
 - Preparing eighth graders for the challenge of obtaining a bachelor's degree, which will be a requirement for 50 percent of jobs in 2020
- Learning more about what it means to say that an educational approach is "culturally appropriate," and also including culturally appropriate options for educators in the *What Works Clearinghouse*

Dr. Buckley reiterated the need to not lose track of postsecondary education, adult education, and early childhood as parts of the education pipeline.

Dr. Herk raised the following priorities:

- She reiterated others' emphasis on teacher quality: how do we recognize who the good teachers are, and once we know how to identify good teachers, how do we recruit, train, and retain them?
- She stated that a similar set of questions apply to leaders of school transformation, such as principals and district leadership.
 - What do good leaders do that distinguishes them from others?
 - To what extent can leaders be trained?
 - How do we recognize, attract, and retain people with the potential to be good leaders?

Roundtable Discussion

Dr. Gutierrez said there is a dearth of studies on dual-language learners from birth to age 4. This population is understudied and sometimes excluded from studies. What works for mainstream children may not work for this population. More research is needed on how early language and literacy unfolds from birth to age 4, looking at home-language support, second-language learning in English, and how early bi-literacy supports learning in formal school environments. She recommended funding more design-based research on two levels. First, IES should consider more design-based research to fit in the space between developmental studies and efficacy studies, and to do so particularly with regard to professional development and instructional quality, so that people start thinking about professional development as an integral part of program design. Organizational supports necessary for program implementation also fall in this category. Second, IES should focus on developing more cross-sector interventions as well as research that integrates data from different sectors (e.g., education, health, justice, and housing) to increase our understanding of the developmental needs and assets of young people. Within special education research, there should be more systematic attention to issues of culture and context in the study of disabilities and interventions.

Dr. Easton described school improvement as more than the implementation of interventions and programs. Good schools are learning organizations that can identify the appropriate interventions and improve upon them. He was glad to hear the Board talking about implementation and support. His question is: how do we as researchers better understand how to promote the kinds of practices that allow the ideas embodied in programs and interventions to play out successfully in schools? What is the research role in understanding how these kinds of practices occur?

Mr. Baron asked about the timing of different types of research. He suggested that after an intervention is shown to have a large impact, then that might be the time to figure out for whom the intervention works and how. When an evaluation finds no overall impact, but there is impact for a subgroup, then it is not clear that that effect is valid or whether it is simply a matter of chance. In light of this, Mr. Baron asked whether it makes sense to do both types of studies—impact evaluation and implementation research—simultaneously.

Dr. Gamoran responded that he thought that the preceding discussion was not so much about *implementation research* (i.e., research about the implementation of a particular intervention) as it was about *research on implementation*—that is research on systems of support within schools that support successful interventions. Dr. Gamoran stated that he had experience conducting a scale-up in which there was pretrial evidence of effectiveness of the intervention but the scale-up was ultimately unsuccessful. It is very important for implementers to understand the conditions they will be going into when they scale-up an intervention. IES can develop a knowledge base that will help implementers understand that and be more cognizant of it going in.

Dr. Long said a researcher has a responsibility to track how an experiment is done so it can be replicated if it works and learned from if it does not work. When a researcher works with secondary data, the difficulty is that the researcher often does not know the details of how the experiment was implemented. The broader question is how IES should invest in research. Should IES follow up projects with successful outcomes with more research on implementation? Or, should IES revisit interventions that were not successful for everyone but might be successful in certain contexts or for certain subgroups? Dr. Long said that the question of priorities does not have an easy answer but depends on the problem, the research literature, and other factors.

Dr. McCardle suggested that an intervention that works only in a particular subgroup may be the beginning of learning what works for whom. Some things work only for a subgroup, and that is very important to know. Not everything that is valuable scales up.

Addressing Mr. Baron's point regarding possibly invalid subgroup findings, Dr. Buckley said that researchers need to internally police themselves to use the appropriate statistical tests that reduce the likelihood of spurious subgroup findings. Secondly, if an educational intervention is found to work, it is not always followed by research into why or how it works—the “active ingredients” question. Especially in natural experiments where a positive effect is found, IES should think carefully about what the next step should be. Is there a way to sharpen the methodology around answering the “active ingredients” question?

Responding to Dr. Gutierrez's comments, Dr. Gamoran wondered to what extent the ECLS birth cohort could help in studying ELLs from birth to age 4. Because it is a nationally representative sample, ECLS captures some of the geographic and linguistic diversity of ELLs. Dr. Buckley agreed that he thought ECLS would be useful for hypothesis development in this area. Unfortunately, the ECLS cohort is not followed into grade school.

Reviewing other themes from the discussion, Mr. Baron commented that in educational research and evaluation, a finding of true effectiveness is rare. Many interventions that seem to work on a small scale fail in larger-scale implementation. Given this history, Mr. Baron suggested that it makes the most sense to conduct implementation research and determine why an intervention works only in those cases where there is a strong expectation that the intervention does work, at least for some people in some conditions.

Dr. Long pointed out that policymakers want silver bullet solutions that will work for everyone. However, they may not exist. How do we communicate to external audiences, such as teachers, the nuances and complexities of interventions that work for particular subgroups under particular conditions? Dr. Long made a second point that even if the Board agreed on a handful of research priorities, the bulk of IES-funded research is generated through competitive grant processes, so IES would have to send very clear signals to communicate these research priorities to the researchers, which is another aspect of the communication process.

Dr. Gamoran agreed that scale-ups sometimes fail because the intervention in question was not sufficiently vetted in the pretrial phase. However, Dr. Gamoran emphasized that sometimes interventions fail because conditions for successful implementation are not present. He gave a hypothetical example of implementing a professional development program only to have half the teachers leave, or half the principals turn over, or four superintendents in 5 years, or the district's entire science leadership team is dismissed and the research office is closed. Dr. Gamoran stated that this shades the implementation issue somewhat differently.

In response to a request by Dr. Easton, Dr. Gutierrez described design-based implementation research. She stated that this new approach to research takes into account the supports, the context, and the ecology of the intervention. The implementation questions are built into the research design.

Dr. Maynard said that study design depends on what question someone is trying to answer. Sometimes people want to answer what is the effect of making one simple policy change, such as giving performance incentives to teachers. The more holistic approach described by Dr. Gutierrez assumes that policymakers have more policy levers to pull than they typically do. Both approaches have merit, but they are answering different questions.

Dr. Gamoran commented that even if a particular study indicates that performance incentives did not work in a particular context, another question might be under what conditions do incentives help teachers perform better. Answering that question would require a more holistic design or going back to small-scale studies and design-based research to develop hypotheses, which would ultimately build to a large-scale study.

Dr. McCardle noted that although NICHD and IES have many overlapping areas of interest, many Board members thought that teacher quality, teacher training, and instructional quality were important, and only IES works on those issues. Mr. Baron agreed IES is unique in working on those issues, which is a reason to continue focusing on them.

Dr. Long said postsecondary education should be part of that research, because we know very little about how to teach college students effectively. Many professors were not taught pedagogy.

Dr. Ball said anything affecting student learning opportunities in an organized setting will depend on *instruction*, which is still an imperfect word choice. “Teaching” takes one back to the individual teacher. We need a word which encompasses the entire instructional delivery system and all transactions that raise the probability of student learning. Any effort that attempts to understand or improve what young people learn needs to make *instruction* the center of the enterprise. It is a difficult dynamic to study, so it has been, unfortunately, avoided.

Dr. McLeod suggested broadening the concept to “educator quality,” because schools must have the conditions in place for even an excellent teacher to make important changes at a school or district level. The administrators are part of that. The focus should be broadened beyond the individual classroom.

Mr. Baron suggested broadening the focus further by including research on interventions in which teachers are hired on a provisional basis and given tenure based on their track records in improving student achievement. There may be other ways besides this value-added approach of identifying who will be a good or bad teacher or administrator at the outset. In addition to professional development and coaching, a range of interventions to improve teacher quality could be tested.

Dr. Gamoran commented that based on recent findings, he was skeptical that an incentive-based approach based on a teacher hitting a certain mark within a certain timeframe would work. The larger point is that there is a broad range of approaches to improving student experiences in the classroom. There are two schools of thought on how to get better teachers: one is to pick the right people and the other is that the way to get better teachers is to help existing teachers get better.

Dr. Long commented that when she brought up incentives, she did not just mean pay-for-performance. She meant any incentives teachers might have to change what they are doing. These incentives are not always monetary, and non-monetary ones can be more important. Dr. Long continued that there are areas such as teacher recruitment and pre-service training where more research could be done. She stated that there seemed to be general agreement that improving educator quality is a fruitful area for research.

Dr. Ball commented that the issue is not about individual teachers but about growing the capacity of a large pool of people who can help students learn. Every time one looks for evidence to help settle this debate about recruitment, training, and systems, then the evidence is very scant. IES could make an enormous difference if it could find a way to support and encourage work of this type.

Dr. McLeod said part of the problem is that most researchers do not have recent teaching experience.

Mr. Baron raised the question of how to move good research findings into practice and towards something that is actually going to make a difference in schools.

Dr. McCardle commented that requiring that researchers have proven track records will harm new researchers. The peer review system can serve to verify that a researcher has a commitment from a specific school system to collaborate in the proposed research. Dr. Easton noted that there is a distinction between access to data and subjects on the one hand, and sensitivity to the needs of the school system and the collaborative spirit of the research. Dr. McCardle responded that there should be a feedback loop in which schools won't agree to work with researchers unless the researchers agree to come back to the school to discuss their findings and continue to do professional development around their intervention.

Dr. McLeod said that she would like to see researchers engaging stakeholders at the beginning of a project to help define the areas of research. She would like it if researchers came to practitioners to ask, "What do you need answered?" Dr. Okagaki commented that we have 21st century expectations of science and what science can do, but we are just getting to the point that we have researchers who are addressing issues that really matter to schools and districts. Some people are concerned that we don't have enough interventions with good efficacy data, but we haven't been doing this work that long, and science takes time.

Dr. Gutierrez suggested that future discussions look at the peer review process. The meeting recessed from 12:20 p.m. to 1:19 p.m.

Communications

Dr. John W. Wallace, former Vice President for External Affairs, MDRC

Mr. Baron introduced the session, whose purpose was to look at effective ways for communicating key research findings to inform federal, state, and local education policy and educational practice at the school or classroom level.

Dr. Easton opened his remarks by briefly describing his experience at the Consortium on Chicago School Research, which saw communication and outreach as a key part of its mission. He commented that many of the same principles apply to IES. IES's audiences must trust IES for being objective, fair, and unbiased. He went on to describe three areas that IES staff are actively working to improve. First, in written materials, staff are trying new formats, developing clear expectations of what constitutes a well-written report, and working to produce written material that is easily and quickly understood. Second, staff are working to improve IES's electronic communications, including its website, search tools, and sites like the *What Works Clearinghouse*. Finally, there is in-person communication. IES still conducts numerous workshops and webinars and makes presentations to professional organizations. Dr. Easton finished by raising the question of the appropriate roles of the Board, himself, and IES senior staff and Commissioners in fostering better communication.

Dr. Wallace began his presentation by stating that the purpose of social policy effectiveness research is to produce credible, reliable results and to communicate these findings to policymakers and practitioners. Its central goal should be to put definitive research results into action by informing policy development and improving on-the-ground practice. He focused on communicating research findings to policymakers and practitioners, because this is an area that makes social policy effectiveness research different from other areas of research that communicate findings primarily to other researchers.

In education, the core audience is non-researchers, which in education is a very large population, especially compared to other policy areas. This imposes an especially important responsibility for managers of education research to communicate results so that they are easily understood by policymakers and practitioners. This statement leads to a number of questions.

The first question is what are the key categories of information from effectiveness research that need to be communicated if the findings are to be acted upon. There are three main areas that need to be addressed if policymakers and practitioners are to move forward with research findings:

- Whether or not an intervention worked credibly, for whom, in what environment, and by how much (impact findings)
- What the intervention was and what the management practices and service strategies were, so practitioners know what to adopt or adapt and what to avoid (implementation research findings)
- The costs and benefits of the programs (benefit/cost results)

The second question is what does it take to successfully communicate this information, and what changes in the current research culture and resource allocation decisions are needed. Producing research findings in ways that practitioners and policymakers can understand is as important as conducting the research itself. Without clear and effective communication of the results, there is no reason to have done the research. A number of lessons have been learned about communicating social policy results:

- Communication staff are needed who understand that the primary audience is non-researchers, and they must know how to speak to that audience. Communication staff must be able to translate complicated technical findings in language that is simple and easy to understand but true to the research.
- Communication staff need to build relationships with their primary audiences (congressional and agency staff, public interest groups, and other groups) to make the audiences open, eager, and receptive to the research.
- Communication staff must focus not only on developing clear executive summaries but also on translating the full report and making it easy to read.
- Taking communication seriously affects resource allocation decisions. Most agencies and organizations underfund communication. Federal research contracts usually do not include resources for communicating the results.

The third question is: What does this mean for the vehicles by which we communicate research findings: Internet-based tools, websites, e-mail, social networks, and the written reports themselves?

- Creating a new website or revamping an old one should be guided by the recognition that the primary audience is non-researchers. The site should be attractive, inviting, user-friendly, and easy to navigate, before the site is marketed.
- If a legislative staffer or state superintendent does not quickly and easily understand the information on the site, he or she will not return.

Dr. Wallace offered two examples of scientific research in which the results were clearly conveyed: a recent HIV finding that early treatment after exposure to the virus reduces the incidence of AIDS and the finding that HDL cholesterol boosters do not reduce heart attacks or mortality. The examples were understandable, actionable, and led to concrete changes in policy and practice. Education research should strive for those same attributes. He concluded by reasserting that recognizing the importance of communications would represent a critical shift in thinking and culture in social policy research.

Roundtable Discussion

Mr. Baron commented that IES has shown substantial progress in communications over the past few years. He has sometimes been frustrated that Board discussions of communications have usually been held in the abstract without concrete examples. So he offered two examples of IES reports which clearly explained what was evaluated in plain language without statistical jargon: the “Impact of the Violence Prevention Program for Middle Schools: Findings After 3 Years of Implementation” report and the Middle School Mathematics Professional Development Study. As an example of a report in which communications could be improved he cited the “Creating Independence through School-Owned Strategies (CRISS)” summary on the *What Works Clearinghouse* website.

Dr. Gamoran said researchers are a part of the IES audience; thus, some written materials, especially synthesis studies, can be targeted for researchers and other materials targeted toward a general audience. Dr. Long agreed that the researchers are an important third audience, but clarity is still necessary to communicate across researchers in an interdisciplinary field. Dr. Long encouraged IES to consider that a tool should not try to communicate to researchers, policymakers, and practitioners at the same time. It is unclear whether the *What Works Clearinghouse* is supposed to target researchers or practitioners and policymakers: currently much of the language on the site seems more geared towards researchers.

Dr. Long said that researchers will want the full study when they go to a website. Practitioners can use shorter summaries with a more accessible writing style. Writing for different audiences requires very different skills; therefore, the intended audiences for each publication need to be clear from the outset.

Dr. McLeod pointed out that the Board members do not represent typical consumers of research. The IES website should be simplified, as should a good deal of the writing, such as the Project CRISS report and some of the four-page summaries. It is important to carefully consider a document's intended audience. Dr. McLeod also suggested that it might be helpful if users could sign up for automatic e-mail notifications when new research findings came out in specific areas. The writing should be understandable by someone who can understand the New York Times.

Dr. Maynard pointed out that the *What Works Clearinghouse* has a different set of purposes than most IES products. The *What Works Clearinghouse* is a database designed so that users can query the database in ways related to their particular interests. Dr. Maynard emphasized her openness to feedback but expressed caution about changing the *Clearinghouse* in a piecemeal way for one audience and thereby reducing its usefulness for other audiences.

The discussion turned to whether practitioners could or should make decisions about what to implement in their schools based on one document in the *Clearinghouse*. Dr. McCardle said that practitioners should collect evidence from multiple sources and then decide, and that while IES can provide data to guide such decisions, schools (district leaders, principals, and teachers) ultimately decide what to implement. Mr. Baron said congressional staffers and other users of research collect evidence from many sources, but for each study, they want to know whether there is valid evidence of a real effect. The *Clearinghouse* can be the first stop for those looking for more information about a particular approach or intervention.

Dr. Long asked if we know whether the *What Works* website is user-friendly for the average user. Dr. Maynard responded that there are ongoing usability studies. Dr. Long said another question concerns whether the information in the *Clearinghouse* is written in a way that gives a person enough information to determine if the program works. Dr. Long expressed concern that the *Clearinghouse* information may currently be unclear and confusing for practitioners.

Dr. Wallace said that if the primary audience of a website is practitioners and policymakers, then the website should emphasize detailed descriptions of the interventions and their costs.

Dr. Buckley commented that most of IES's communications are not about communicating evaluation impacts, and that a discussion of IES communications should also look at these other efforts. For example, 20 percent or more of NCES's communications budget is for simple NAEP descriptor statistics. NCES communications staff work on layers and levels of publications for different audiences. However, effective communication strategy goes beyond rewriting for different audiences. Because communications are mediated, Dr. Buckley devotes significant time to media outreach and explaining findings in clear terms to journalists and others.

Dr. Long expressed support for Dr. Buckley's efforts in this area and agreed that active outreach is critical. She stated that while the IES website is evolving, a continual feedback mechanism will be necessary, and she asked whether funding for such feedback is in the budget.

Dr. Easton said that improving the website is a high internal priority; there is a contract to improve the site, including data collection and outreach. Following up on Dr. Buckley's comments, Dr. Easton asked for input from the members on IES communications efforts other than the *Clearinghouse*.

Dr. McLeod stated that she liked the comparison tools on the NCES website. Dr. Long said NCER webinars have made the grant application process clearer.

Dr. Okagaki reported that IES is working on a one-stop search tool that will allow searching for reports across all of IES instead of having to search each Center. Responding to Mr. Baron's earlier comments, Dr. Okagaki stated that NCER's research reports include a two-paragraph summary written in plain language for policymakers and practitioners. But researchers need more than that level of detail, so the reports include a detailed, structured abstract for researchers. Principal investigators funded by IES (and other agencies) often have difficulty communicating their research in plain language to others outside their field. Researchers should learn to speak more plainly about what they do, but the fact remains that researchers and policymakers will need different levels of detail about a project. Mr. Baron agreed that different actors would need different levels of detail, but that all of it should be written in plain language.

Dr. Okagaki said it is necessary to strike a balance between plain language and precise language. Phrases like "potentially positive findings" have a precise meaning.

Dr. McCardle noted that the term "plain language" can be confused with the Plain Language effort in the government, so it might be better to say "clear and readable."

Dr. Maynard pointed out that the *Clearinghouse* must be consistent in its use of language, so if a term like "potentially positive" is redefined, staff will have to go back and apply the new definition to everything in the database.

Dr. Wallace said IES and rigorous education research have been around for a relatively short time; thus, now is the right time to focus on communication issues. Mr. Baron agreed that more attention has been paid to improving the way evaluation findings are reported and that the *Clearinghouse* search tool is an important improvement. He encouraged continued evolution over time.

Dr. Gamoran commented on two other areas of IES communication:

- The IES Requests for Proposals, which are clear and well-structured; and
- Results of proposal reviews that IES gives to proposers, which are effective and useful.

The meeting recessed from 2:31 to 2:50 p.m.

Low-Cost RCTs: Could They Play a Key Role in Building Knowledge About “What Works” in Education?

Dr. Eric Bettinger, Stanford University; Dr. Robert Slavin, Johns Hopkins University and the Success for All Foundation

Dr. Bettinger focused on the use of administrative data. He first gave the example of his work on educational vouchers in Bogota, Columbia, where he used local university students to collect data. There was a 55 percent response rate, the cost was \$300 to \$350 per observation, and the data collection took 1.5 years. Years later, he ran across administrative data that could be matched to the original subjects, resulting in 100 percent follow-up on the original subjects that cost \$6 per observation and only took 2 months.

He discussed a few lessons learned from the experience:

- Administrative data are key to reducing the cost of RCT evaluations.
- The quality of the data initially collected from the subjects influences the probability of whether subsequent data collection will be low-cost.
- When administrative data are available, the time to complete the study is much shorter.
- Using administrative data has some inherent tradeoffs.

In the United States, he has been able to use administrative data to track a number of outcomes: college enrollment, college retention, degree completion, college engagement, college choice, student achievement, student completion of the Free Application for Federal Student Aid (FAFSA), student utilization of loans, student receipt of state aid, student receipt of federal grant aid, student college entrance exam scores, and students' earnings 6 years after they initially enrolled in college. Others have used personnel records, welfare records, and Social Security earnings.

The two largest barriers to using administrative data are privacy and access to the appropriate data.

The time of initial data collection from the subjects is crucial because it determines whether or not it will be possible to use administrative data. Subjects' initial entry into the experiment is: the opportunity to obtain informed consent; the time to validate the randomization; the opportunity to collect background data for any sort of subgroup analysis; a time to gather and shed light on some of the potential mechanisms of the intervention; and the determinant of our ability to do subsequent matching with the administrative data. The data collection efforts of many organizations that implement educational interventions are sub-par at this important initial point of entry. Failure to talk to the researchers and formulate hypotheses before enrolling subjects in the study causes the study's costs to escalate.

There are tradeoffs inherent to the use of administrative data. Often, the key policy question or implications rely on the underlying mechanisms of the intervention. One of the weaknesses of using administrative data is the risk of losing contact with those mechanisms. This can be minimized by gathering data at baseline and using it effectively to shed light on the mechanisms.

Mr. Baron requested that Dr. Bettinger briefly describe the Inside Track College Coaching Study that Dr. Bettinger conducted using administrative data. This RCT at eight colleges sampled 13,000 students. The intervention was coaching aimed at keeping students in college. Students who were randomly assigned to receive the intervention received a call from a coach within 2 weeks of starting their freshman year. For the next 12 months, students who agreed received regular calls from their coaches in which the coaches encouraged them and discussed any problems the students might be having. At the 12 month point, there was a 5 percent improvement in student retention compared to the control group and 3 percent improvement up to 24 months. Among the subset of students that could be tracked to graduation, there was a 3 percent improvement in retention.

Dr. Bettinger's team only had access to the administrative data and was not allowed to contact the students, which was a barrier to determining how the intervention is effective. Conducting the study cost only 3 weeks of staff time.

In response to a question from Dr. Gamoran on what is known about promoting college retention, Dr. Bettinger cited financial aid literature indicating that for every \$1,000 spent on financial aid there is a 3 percent increase in the student's likelihood of staying in college, but there have been few rigorous studies supporting these figures. (The cost of the coaching intervention was also \$1,000 per student.)

Mr. Baron asked about the cost of conducting a FAFSA intervention that Dr. Bettinger had been involved with. Dr. Bettinger said they had completed follow-up on the participants for each of their first 2 years of college. Overall cost of data collection for those 2 years was \$38,000, or less than a dollar per year for each of the 20,000 students in the study. The initial set-up cost of the study—software and training—was more expensive, but it is a fixed cost. The actual evaluation is inexpensive.

Mr. Baron commented that these kinds of results in real-world settings at multiple sites represent the kind of studies that make the case for the value of education research.

Dr. Slavin addressed the Board about the use of federal funding streams to encourage school districts to participate in experiments. In education research, the gold standard should not be just the RCT but the *cluster* randomized trial, because students are taught in classrooms, and classrooms are in schools. But because cluster randomized trials are expensive, instead studies often randomly assign students to classes or conditions *within* one school, which is an artificial arrangement that does not exist in real schools. What researchers should do is randomly assign schools to implement an intervention, but that is extremely expensive. Power analysis indicates

that the ability to detect an effect size of 0.2 would require 40 schools and the ability to detect an effect size of 0.15 would require 50 schools. Plus we need a lot of these studies to evaluate many different interventions, and frequently the studies find that the intervention is not effective.

The challenge is how to conduct many of these studies in the real world. Often the government provides targeted funding to schools to conduct some category of activity. School Improvement Grants (SIGs) are a current example of such funding aimed at improving low-performing schools. Often there is an evaluation associated with the program, but it is expensive and findings often are not available until after the program has been abolished.

Dr. Slavin proposed embedding evaluations into targeted grant programs from the outset. The strategy would be to require similar districts or schools to apply for funding in pairs to implement the same intervention. The schools (or districts) would be randomized such that one school (or district) would implement the program in the first year, and the other in the second year, if the first year results were consistent with effectiveness. Study results would accumulate over time through aggregation across many pairs of schools or districts. Administrative data, such as state test scores or a group-administered test, could be used to measure outcomes. The cost of doing the study would be simply the cost of the assessment, data collection, and analyses. Several of these studies might be underway at any point in time, producing high-quality evaluations every year at a low cost. If grant funds were awarded based on a 100-point proposal score, then applying in this paired format might be worth 5 points.

The funding agency would need to develop standards to determine which interventions are worth evaluating in this way. Interventions might be worth evaluating either because they are in wide use, or they have promising initial evidence behind them, or there is a compelling government interest in knowing the intervention's impact.

The grants would be at the national level, and states would be asked to include this competitive preference in their re-grants of the funds to participating schools. Any new models or interventions would be tested against "business as usual" in the school (or district) that does not receive the intervention in the first year. The basic notion is to accumulate a national sample from many of these experimental pairs. Participating in this paired evaluation process would not be a requirement for funding, but would simply provide competitive preference to schools (or districts) that applied in this way.

Dr. Gamoran noted that there is a connection between this idea and Dr. Bettinger's coaching study. The coaching study used internal randomization (randomizing students at a particular university) rather than cluster randomization and then aggregated the data across the experiments at multiple universities. What Dr. Slavin is describing would randomize schools (or districts) in pairs and then aggregate the data across multiple experiments testing the same intervention. This approach is most likely to work for narrowly-targeted interventions rather than a complex school turnaround strategy. This is consistent with the Department requirement that school districts accepting funds implement rigorous evaluations of the programs. For best results,

the interventions should be standardized. Dr. Gamoran commented that this approach offers a way to get more out of what is currently done with those evaluations.

Dr. Long echoed Dr. Gamoran's comments, saying that this approach is more likely to find positive results in the case of small, narrow, highly targeted interventions, because with larger, more complicated interventions, the inevitable compromises that arise in the course of implementation reduce the ultimate effect size. Dr. Long then asked Dr. Bettinger whether he was able to predict in advance that his coaching study would be low cost and would lead to a positive finding, because often research projects are unpredictable in both regards.

Mr. Baron commented that even if most studies do not result in a positive finding and it is hard to predict which ones will, if we conduct many of these studies then we will discover some successful approaches that we could not predict in advance.

Dr. Gamoran stated that he thought that approach would be too expensive because it would require so many evaluations. Dr. Gamoran's recommendation would be to invest more in development and careful testing and place more emphasis on discipline-based theories and theories of change.

Mr. Baron said the approaches are complementary, especially when there is the opportunity to obtain several quick, inexpensive impact evaluations through the means Dr. Slavin described.

Dr. Pendleton said IES supports Dr. Slavin's idea, but it works better with some federal grant programs than with others. For formula grant programs, where money goes out based on the number of disadvantaged students in states and districts, it would take an act of Congress to change how money is allocated. In contrast IES has worked with the program offices within the Department to insert evaluation requirements into the application for some discretionary grant programs where applicants apply directly to the Department. Agreeing to participate in an evaluation is a condition of the award, and these requirements are noted on the application notice to potential proposers. Striving Readers, drug testing, and the Teacher Incentive Fund are examples of programs where this has happened. It is a matter of IES working with the program offices early enough to get into the program's application notice. Ms. Silverberg said there are different models for incentivizing the evaluation, including one under which the grantee gets a regular award to implement the intervention and an extra award if the grantee agrees to participate in a random assignment evaluation.

Dr. Maynard gave the example of the Department of Health and Human Services working with states to conduct welfare reform experiments using administrative data for employment, earnings, and welfare receipt. Federal welfare rules were about to change, and the federal government offered states waivers from the new rules, but the waivers came with the requirement that states participate in an evaluation of their proposed alternative approach using an RCT with a large sample. The evaluations were conducted at low cost because they focused on outcomes

that could be measured with administrative data. It was a productive line of research that advanced the science of welfare policy.

Mr. Baron wondered whether a similar approach might work in education. Because so many schools will face No Child Left Behind sanctions, offering waivers conditioned on participation in experiments could incentivize participation in rigorous evaluation that builds knowledge. Dr. Slavin gave as a possible example setting up a policy that schools not meeting No Child Left Behind goals could continue to use the teachers already in the schools for Supplemental Education Services (SES) as long as they agreed to participate in the evaluation of specific replicable models that qualify for SES funding. This type of approach is in keeping with the intent of the funding stream but allows variation to be explored.

Returning to an earlier discussion, Dr. Long commented that because education funding is so limited, it will be too expensive to conduct large numbers of experiments in search of the few that show positive effects. Funders of research need to figure out how to place our money on the most promising bets. However, Dr. Long agreed with Mr. Baron that if evaluations of current interventions are possible at low cost they should be pursued.

Dr. Long also noted that working with for-profit companies (as in the college coaching and FAFSA studies) defrays costs, but there are tradeoffs, such as the company considering the intervention their intellectual property. Mr. Baron noted that there is also the option of working with nonprofits, as Ms. Silverberg did with the drug testing study. Ms. Silverberg noted that there is a long gestation period, including a period for public comment, before ED's program office puts out a discretionary grant notice. If IES wants to add an evaluation requirement to a grant notice, IES has to obtain a commitment from the program office a year before the grant notice goes out. But this delay helps IES develop a strong relationship with the program office.

Closing Remarks, Including Next Steps

Dr. John Q. Easton, IES Director; Mr. John Baron, NBES Chair

Mr. Baron asked the members what general qualities Dr. Easton should look for in a new NCER Commissioner and the process Dr. Easton might use in finding Dr. Okagaki's replacement.

Dr. Okagaki said the person must know education research broadly and understand different methodological approaches, be skilled in administrative management, and be able to mentor the new staff and young researchers.

Dr. Gamoran said the position has grown and changed during Dr. Okagaki's tenure. The new commissioner must continue to institutionalize a process and organization that is already established, fight for investment in educational research and to maintain opportunities for educational research, and provide a vision for the field while remaining responsive to the field. He hoped there would be further discussion on mentoring, especially as it relates to the proposal

review process. The next commissioner should also be a team player who is able to work with the IES senior management team, advance the interests of IES, and work for field-based research.

Dr. Gutierrez stated that the candidate should also understand the complexity of the problems in education today, both methodologically and theoretically, and should be able to think of a wide range of interventions, including interdisciplinary ones, to address these problems.

Mr. Baron added that the commissioner should not only have strong research credentials but also should understand what is of policy and practical importance. NCER's end goal should be developing strategies or knowledge to improve American education; thus, the commissioner should be open to evaluating programs that may be small or unglamorous but potentially could have a large impact. The focus should be not just on research but on generating policy of practical importance. He asked Board members to e-mail him to suggest possible candidates.

Dr. Easton added some closing remarks. IES has two areas where the Board's approval is required: IES's research priorities and the review process. Reviewing IES's review process will take longer than one meeting. He requested that the Board begin addressing that item and suggested keeping it on the agenda for the next year or longer.

Dr. Gamoran appreciated the discussion of NCES and hoped there would be a discussion of NCER at the next meeting. The roles of program officers are different in different federal research organizations, with different strengths and weaknesses. There is much that can be learned by looking at this. Dr. Gamoran continued that the Board should be alerted to education research initiatives as they arise. Dr. Herk listed the Science, Technology, Engineering, and Mathematics (STEM) initiative as one upcoming interagency initiative.

Dr. Gutierrez wanted to know more about the new synthesis reports and the role of field-initiated research in the Centers. Mr. Baron said the peer review process, the synthesis reports, and some of the initiatives in NCER can be future agenda topics.

Dr. Herk spoke about the NBES Annual Report, one of the Board's statutory duties. It must be submitted annually by July 1. The statutory requirements for the report were included in the Board notebooks. Mr. Baron said the report has historically been purely descriptive of IES activities. He wondered whether the report can provide more meaningful input and recommendations to IES or the policy world. Dr. Easton added that the report is currently written by IES staff, and it could be written by the Board. Dr. Herk noted that if the Board chooses to generate the report itself in the future, it will require addressing the content for the report at more than one Board meeting.

Mr. Baron announced that a new chair will have to be elected at the next meeting. Dr. Gamoran moved adjournment. Dr. Gutierrez seconded and the meeting adjourned at 4:14 p.m.