TECHNICAL WORKING GROUP MEETING

ies INSTITUTE OF EDUCATION SCIENCES

Researching College- and Career-Ready Standards to Improve Student Outcomes

Meeting Summary

August 19–20, 2013 | Washington, DC

Lyndon Baines Johnson Department of Education Building

Participants

Technical Working Group Members

Jennifer Barrett, M.A.
Middle/High School Math Consultant
Kenton County (KY) School District

James Benson, Ph.D.
Associate Research Scientist
National Center for Education Research
Institute of Education Sciences

Thomas Brock, Ph.D.
Commissioner
National Center for Education Research
Institute of Education Sciences

Jack Buckley, Ph.D.
Commissioner
National Center for Education Statistics
Institute of Education Sciences

Michele Cahill, M.A.
Vice President, National Programs
Carnegie Corporation of New York

Jere Confrey, Ph.D.
Chief Mathematics Officer, Amplify Learning
Joseph D. Moore Distinguished University Professor
The Friday Institute for Educational Innovation
College of Education
North Carolina State University

Carol McDonald Connor, Ph.D.
Professor
Department of Psychology
Learning Sciences Institute
Arizona State University

Michael Coyne, Ph.D.
Associate Professor
Department of Educational Psychology
Neag School of Education
University of Connecticut

Marielena DeSanctis, Ph.D.
Executive Director, Instruction and Interventions
Broward County (FL) Public School System

Janice Earle, Ph.D.
Senior Program Director
K-12 Science, Technology, Engineering, and Math Education
National Science Foundation

John Q. Easton, Ph.D.
Director
Institute of Education Sciences

Stephen N. Elliott, Ph.D.
Mickelson Foundation Professor
The Sanford School
Arizona State University

Maria Ferguson, B.F.A.
Executive Director
Center on Education Policy
Graduate School of Education and Human Development
The George Washington University

Joan Ferrini-Mundy, Ph.D.
Assistant Director
Directorate for Education and Human Resources
National Science Foundation

Lynn Fuchs, Ph.D.
Professor
Peabody College
Vanderbilt University

Beth C. Gamse, Ed.D.
Principal Associate
Abt Associates

Ray Hart, Ph.D.
Director of Research
Council of the Great City Schools

Daniel Heck, Ph.D.
Senior Researcher
Horizon Research, Inc.

Heather Hill, Ph.D.
Professor
Harvard Graduate School of Education

Karen King, Ph.D.
Program Director
Division of Research on Learning in Formal and Informal Settings
National Science Foundation

Paul LeMahieu, Ph.D.
Senior Vice President, Programs and Administration
Carnegie Foundation for the Advancement of Teaching

Nonie K. Lesaux, Ph.D.
Professor of Education
Harvard Graduate School of Education

Gary A. McCormick, M.A.
Secondary Literacy and Curriculum Consultant
Kenton County (KY) Board of Education

Rebecca Kang McGill-Wilkinson, Ph.D.
Associate Research Scientist
National Center for Education Research
Institute of Education Sciences

Joan McLaughlin, Ph.D.
Acting Commissioner
National Center for Special Education Research
Institute of Education Sciences

Barbara Means, Ph.D.
Director
Center for Technology in Learning
SRI International

James Middleton, Ph.D.
Director and Professor
Center for Research on Education in Science, Mathematics, Engineering, and Technology
Arizona State University

Brett Miller, Ph.D.
Health Scientist Administrator
Child Development and Behavior Branch
Eunice Kennedy Shriver National Institute of Child Health and Human Development
National Institutes of Health

Anne Ricciuti, Ph.D.
Deputy Director for Science
Institute of Education Sciences

Barbara Schallau, M.A.
Curriculum Coordinator for Mathematics
East Side Union High School District (CA)

Catherine Schmidt, M.A.T.
Striving Readers Elementary Training Coordinator
Washoe County (NV) School District

Catherine Snow, Ph.D.
Patricia Albjerg Graham Professor of Education
Harvard Graduate School of Education

Meghan Snow, M.Ed., Ed.M.
Director
Office of Science, Technology, Engineering, and Math Education
New Jersey Department of Education

James Spillane, Ph.D.
Professor
Northwestern University

Jonathan Supovitz, Ed.D.
Associate Professor and Director
Consortium for Policy Research in Education
Graduate School of Education
University of Pennsylvania

Mark Thomas, M.A.
Principal
Northview High School (MI)

Bill Tucker, M.B.A., M.A.
Deputy Director, U.S. Program Policy and Advocacy
Bill & Melinda Gates Foundation

Mary Visher, Ph.D.
Senior Associate
Young Adults and Postsecondary Education Policy Area
MDRC

Priscilla Wohlstetter, Ph.D.
Distinguished Research Professor
Education Policy and Social Analysis
Teachers College, Columbia University

Tonya Wolford, Ph.D.
Deputy, Research and Evaluation
The School District of Philadelphia

Observers

Corinne Alfeld, Ph.D.
Standards and Review Office
Institute of Education Sciences

Lisa Bridges, Ph.D.
Standards and Review Office
Institute of Education Sciences

Janis Brown, Ph.D.
National Center for Education Statistics
Institute of Education Sciences

Jackie Buckley, Ph.D.
National Center for Special Education Research
Institute of Education Sciences

Christina Chhin, Ph.D.
National Center for Education Research
Institute of Education Sciences

Wai-Ying Chow, Ph.D.
National Center for Education Research
Institute of Education Sciences

Elise Christopher, Ph.D.
National Center for Education Statistics
Institute of Education Sciences

Karen Douglas, Ph.D.
National Center for Education Research
Institute of Education Sciences

Robin Hall, D.A.H.
Council of the Great City Schools

Amanda Hoffman, Ph.D.
National Center for Special Education Research
Institute of Education Sciences

Ronna Spacone, M.S.
Office of Vocational & Adult Education
U.S. Department of Education

Kathy Mann Koepke, Ph.D.
Eunice Kennedy Shriver National Institute of Child Health and Human Development
National Institutes of Health

Mark Leddy, Ph.D.
Directorate for Education and Human Resources
National Science Foundation

Benjamin Nickerson
National Center for Education Research
Institute of Education Sciences

Rob Ochsendorf, Ed.D.
National Center for Special Education Research
Institute of Education Sciences

Kristen Rhoads, Ph.D.
National Center for Special Education Research
Institute of Education Sciences

Diane Stark Rentner, B.S.
Center on Education Policy

Allen Ruby, Ph.D.
National Center for Education Research
Institute of Education Sciences

Kim Sprague, M.Ed.
Standards and Review Office
Institute of Education Sciences

Amy Sussman, Ph.D.
National Center for Special Education Research
Institute of Education Sciences

Katina Stapleton, Ph.D.
National Center for Education Research
Institute of Education Sciences

Meeting Summary

Day One: August 19, 2013

Welcome

Thomas Brock, National Center for Education Research (NCER), Institute of Education Sciences (IES)
Joan Ferrini-Mundi, National Science Foundation (NSF)
Dr. Brock welcomed the members of the Technical Working Group (TWG). He noted that the Common Core State Standards (CCSS) and related state initiatives have generated a great deal of interest and scrutiny. On the one hand, proponents believe such standards have the potential to transform American education and improve student readiness for careers and postsecondary education. On the other hand, skeptics may support the goals but worry about implementation and unintended consequences, particularly for students with special needs and those struggling academically. Therefore, high-quality research and evaluation is needed to understand what is happening on the ground and how schools, teachers, and students are affected.

Dr. Brock said the TWG is not expected to reach consensus on the merits of the CCSS but rather to provide input on what research is needed, including possible frameworks and methodologies for assessing the CCSS in the short-, medium-, and long-term. He hoped the meeting would help federal research funding agencies, including IES, NSF, and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), identify funding priorities. He also expressed hope that private philanthropies and others interested in studying the CCSS and related initiatives would benefit from the TWG discussion. For this reason, a summary would be prepared and made available to the public. Dr. Brock thanked Janice Earle of NSF and Brett Miller of NICHD for their collaboration in planning this meeting. He also acknowledged the work of James Benson and Rebecca Kang McGill-Wilkinson of IES, who played lead roles in conceiving and organizing the meeting.

Dr. Ferrini-Mundy said she was delighted that NSF has partnered with IES and NICHD, and she thanked the "phenomenal" group assembled for the meeting. She said the CCSS and related activities pose a significant opportunity and a challenge for research to inform the continuing evolution of education standards. She hoped the TWG would address, among other topics, the meaning of such terms as "interpretation," "implementation," and "alignment of standards"—what it means to align standards with instruction or materials.

Currently, NSF is reviewing the CCSS and the Next Generation Science Standards, said Dr. Ferrini-Mundy. In the past, NSF funded the development of instructional materials and professional development aligned with new standards. She hoped the TWG's input would help NSF shape its future funding decisions. She also hoped that research would address the needs of practitioners, federal agencies, researchers, and policymakers.

Opening Plenary: The National Picture

Presenter: Maria Ferguson, Center on Education Policy (CEP)
Ms. Ferguson provided some key findings from CEP's recent state surveys on CCSS implementation, professional development, and teaching. Despite some vocal critics of the CCSS in the media, nearly all of the states that responded to CEP surveys indicated that they are unlikely to reverse, change, or limit their decision to implement the CCSS, and all said that the CCSS are more rigorous than current state standards. Most acknowledge that implementing the CCSS requires a major change in curricula and instruction, and indicated that they are phasing in the CCSS gradually.

Most states recognize the need to help low-performing schools make the transition to the CCSS. Nearly all states are collaborating with other states on implementation. Ms. Ferguson noted that the CCSS enable states to communicate using a common language for the first time and to consider economies of scale. Collaborations are happening in several areas, especially professional development. Most states are working

with schools and districts by hosting informational meetings, providing technical assistance, and developing materials and guides to prepare for the CCSS. At this time, however, only a few states have reported a large portion of teachers trained in the CCSS.

Ms. Ferguson said that political resistance to the CCSS has turned out to be a relatively minor issue for states. In states facing spending cuts, the implementation requirements for the CCSS are in the line of fire, however. More than half of the survey respondents are concerned about their ability to provide high-quality professional development. Ms. Ferguson said assessing the fidelity of implementation of the CCSS within and across states remains a challenge.

Most states reported that they wanted federal funding to help teachers, provide professional development, and assist with the cost of implementing and updating assessments over time. Only two states indicated that they would prefer to have no federal involvement. Overall, Ms. Ferguson observed, the lack of high-quality professional development opportunities for teachers is a huge problem in all states. There are also concerns about (1) the adequacy of resources and expertise to implement the CCSS with fidelity; (2) the need for a smooth transition to the CCSS, especially for English-language learners (ELLs) and students with special needs; and (3) the need to develop fair systems for teacher evaluation and accountability with the CCSS in play.

Presenter: Ray Hart, Council of the Great City Schools (CGCS)
Dr. Hart said CGCS represents 66 of the country's largest urban school districts and believes that the CCSS are very important for students around the country. He provided preliminary results from a 2013 CGCS survey of chief academic officers (CAOs) on implementation of college- and career-ready standards across school districts (the CCSS in most cases). Many CAOs said they felt well prepared for implementation but indicated that principals, teachers, and others in their districts are less prepared. Dr. Hart said that as part of CGCS's mission to translate standards into classroom practice, the possibility of collecting case studies on implementation to better understand what is happening on the ground has been discussed with school districts.

In 2012, CGCS found that 23 percent of school districts said they had modified formal or informal teacher observation protocols to align with standards, and that practice has risen to 50 percent in 2013. Asked to rate their districts' progress in using technology to facilitate implementation of standards (e.g., adopting computer-based assessments), most responded positively. While few classify their progress as "excellent," many have made significant steps since 2012, particularly in implementing technology to support new math and English language arts (ELA) standards in the classroom. Dr. Hart noted that many urban school districts have large proportions of ELLs, and CGCS is especially focused on the effects of implementation of the CCSS for students with special needs and ELLs.

Ensuring that students are familiar with, and can use, technological tools is another priority for CGCS. The survey indicated that districts are not providing many professional development opportunities on integrating technology into classroom instruction. Dr. Hart said the survey data will be combined with other, more detailed information to paint a clearer picture of what standards implementation looks like, how standards are being translated into the classroom, and how the relationship between standards and classroom practices is progressing.

Discussion

Michele Cahill asked whether states are documenting or learning from the variations in implementation at the school and district level. Ms. Ferguson did not think states are learning from the districts, and she thought CEP should ask that question pointedly in an upcoming survey of school districts. Diane Stark Rentner of CEP said that such an exchange of information depends on the history of the state's involvement or control. The school district survey will show whether school districts want more state involvement, Ms. Rentner noted.

Dr. Hart said implementation of the CCSS by school districts varies greatly, and many are phasing them in gradually. All of the school districts are working to meet state deadlines for implementation. A lot of implementation also depends on the extent and reach of professional development efforts.

Heather Hill said her research, which involves a lot of classroom observation of math instruction, suggests that schools may not be nearly as far along in implementing the CCSS as states are reporting. From what she has seen, it is happening in only a very small subset of classrooms. Implementation takes a lot of time and requires professional development and learning communities for teachers. Dr. Hill said she is skeptical about the states' responses to the CEP survey.

Catherine Snow added that she was "amazed" that 30 states said that curricula and teaching are already aligned at some districts and grade levels, because she believes there are no aligned curricula available for purchase. Some districts and groups of teachers have put together curricula to support their implementation, but these are not truly aligned with the CCSS. Dr. Snow said it is important to be realistic about how far away we are from understanding the CCSS and from implementing them substantially across K–12. Ms. Ferguson said the CEP reports highlight the disconnect: states think they are implementing standards. No one who has worked at the school district or state level would argue that what the states say is happening and what is actually happening are the same.

Tonya Wolford said her interactions with school districts, teachers, principals, and staff support the CGCS finding that they are not really prepared for implementation. In some cases, people said they were trained, but in practice, it is clear they do not fully understand how to implement the standards.

Priscilla Wohlstetter said there seems to be an idea that implementation starts at the national level and then moves stepwise down into schools. However, some school districts or consortia of school districts have taken the lead in implementation in place of the state (and sometimes the lead entity changes over time). Therefore, implementation is not progressing as we thought it would, said Dr. Wohlstetter. She believes that those school districts or consortia that had infrastructure in place before implementation of the CCSS were in a good position to take a leadership role.

Barbara Schallau wondered whether those involved in content areas other than ELA and math have been involved in implementation. She pointed out that instructors may not be prepared for implementation because their current content knowledge is at odds with new content, new pedagogies, and new assessment approaches. Dr. Hart said CGCS has survey data from other personnel that may speak to that question. He added that the CGCS survey results reflect the CAOs' perception of preparedness of teachers and others. Several school districts have tailored their curricula to align with the CCSS and are focused on providing professional development on teaching the standards, said Dr. Hart.

Beth Gamse said that, many years ago, she was involved in a study of preparedness for new standards. In the first year, principals claimed to have made substantial progress, but by the second year, they realized just how much more they needed to do and were more modest in their claims. She predicted the same would be true in relation to the CCSS. She thought it would be important to include states and districts that are *not* implementing the CCSS in research. All states face issues around curriculum development and alignment, access to content, and capacity of state education agencies. Therefore, studying what happens in states and school districts that are not implementing the CCSS can reveal which challenges are unique to CCSS implementation and which are universal. Dr. Hart said the CGCS survey includes states not implementing the CCSS (e.g., Texas, Virginia, and Alaska), although the sample size is not large enough to break out their responses.

Barbara Means suggested that rather than ask general questions about whether standards are being implemented, researchers could try to identify markers that indicate implementation. For example, the writing element of the CCSS represents a big change and requires professional development on how to teach

argumentative and explanatory writing in content areas like science and social studies. Comparing specific markers of CCSS implementation in states that are and are not using the CCSS would provide a better sense of what is happening than asking about the intention to provide professional development or money spent on professional development. Dr. Hart said the CGCS survey does address writing, and he was surprised by how many school districts are focused on writing across the curricula. Robin Hall of CGCS added that her organization has been working on multidisciplinary professional development around writing. Ms. Ferguson supported the idea of identifying markers. State-level data usually are not very specific, but at the district level, it would be very important to get the markers right, she said.

Gary McCormick expressed skepticism about the perception that teachers are prepared for implementation. To determine whether students are learning, he suggested looking at the "artifacts" (e.g., tests, papers, and presentations) and using them to determine whether teachers are shifting their practices. It would be very helpful to provide teachers with new strategies and education about practices they can use to implement the new standards, said Mr. McCormick.

James Middleton said it appears that those involved in implementation are not all talking about the same thing. Thus, he supported the need to gather evidence and clarify terms. For example, references to "standards" should clearly indicate whether the CCSS or other college- and career-ready standards are meant. An operational definition of implementation that includes the degree and type of implementation is needed. He asked whether efforts are underway to develop or elicit interpretations of what implementation means to school districts, states, and teachers.

Mark Thomas pointed to challenges in teacher capacity. For example, very few high school teachers have extensive training around reading, while few elementary school teachers have training in math teaching skills. Mr. Thomas called for more of a focus on providing information to analyze and establish systems that will address implementation needs.

Transforming Classroom Practice in Mathematics

Moderator: Janice Earle, NSF
Presenters:
Jere Confrey, North Carolina State University (NCSU)
James Middleton, Arizona State University
Barbara Schallau, East Side Union High School District

Dr. Earle posed questions and topics to the presenters throughout the session (identified here in boldface and italics).

Discuss the development of a theory around learning progressions (following increasing levels of sophistication around a concept over time) for curricula and instruction.

Dr. Middleton said he takes an engineering approach to design, in which design occurs in cycles as theory and practice inform each other. He pointed out that we do not know what motivates an individual to persist in a progressive system of math education, particularly the role of students' perceptions of the usefulness of math skills outside of school. We need to know how math is learned naturally so we can create systems to allow "unnatural" learning. We should begin developing methods to study how math learning affects students—for example, whether they recognize math-based problems and have the inclination to solve them once they see they have the skills to do so. Dr. Middleton suggested focusing on translating data into information and on the role of computational methods in comprehension.

Dr. Middleton called for better understanding of how teachers interact with learning progressions and how teachers manage motivation for math throughout K–12. Starting around sixth grade, said Dr. Middleton, motivation drops significantly.

Ms. Schallau suggested embracing our understanding of development across K–12, pointing out that students require time to learn and need to see things in various ways and at different levels. However, we need to identify early indicators that show a student is missing something that will be key to progression over time. Poor/incomplete teaching in the early years can have a long-term effect. Ms. Schallau hoped we could not only prevent negative practices that hamper learning but also identify them when they occur and intervene.

Ms. Schallau said elementary school districts in her region are pushing for accelerated learning. She would like to see research addressing the value of acceleration for students, especially over the long term. She questioned what content and problem-solving experiences such students may be missing. Does the benefit of acceleration outweigh the benefit of sticking to one's grade level?

Comment on the implications for research of digitally delivered curricula.

Dr. Confrey outlined the kinds of changes that could be expected with digital curricula, including engaging problems and challenges, uses of rich media, opportunities for differentiation based on prior data, and just-in-time teacher support. Digital curricula can be highly interactive. There is a critical role for educational research in developing good digital curricula, but often developers are not well-versed in the literature, so supporting public/private partnership may be key to realizing this potential in the era of CCSS.
Dr. Confrey summarized some high-priority research areas:

- Understanding readiness (e.g., do students understand multiplication, division, and fractions before they are taught proportions? Are they prepared for harder math vocabulary?)
- Assessing how schools teach accelerated content
- Determining whether grade levels can be sustained in a digital environment and what alternatives exist
- Understanding what motivates a student to engage in a "productive struggle" to master a topic
- Defining terms such as "individualization," "adaptation," "customization," and "personalization"

Research on CCSS-aligned curricula should draw on the finding from the National Research Council (NRC) report, "On Evaluating Curricular Effectiveness." It addresses the challenges of validly comparing the effectiveness of different curricula. For instance, it stressed the challenge of using appropriate outcome measures that are consistent with the chosen standards and are fair measures, relative to the opportunity to learn, that examine rigor and coherence. She cited the work on the COSMIC project as representing progress in evaluating curricula. Dr. Confrey also commented that if the new assessments are reporting at the level of claims and not targets for assessment, their use to inform instructional decisionmaking by teachers may be limited. She emphasized that effectiveness is probably not best understood as a global characteristic, but should be examined in terms of what works, for whom, and under what conditions. Finally, she recommended that research take into account how to use class time effectively, including tying into out-of-school time, use of digital resources, games, and informal science partners.

Dr. Confrey offered three suggestions related to methodology:

- Build on the progress made by the NRC
- Anticipate the possibility of using analytics gathered from digitally-delivered curricula to conduct experiments in the digital environment and create and test theories of action
- Partner with developers to provide independent review and research, e.g., setting up demonstration schools and tracking effects over time

Do practitioners see new ways of delivering curricula and personalization of learning through digital curricula on the horizon?

Ms. Schallau said teachers are excited to see technology moving into the classroom, and many are taking advantage of it. Digital whiteboards give teachers more freedom to move around the room, she said, and some are creating short, didactic videos that students watch on their own time. Access to technology outside of school is increasing, but access remains an important component to ensure equity, said Ms. Schallau. She added that she is interested in seeing a comparison between students who complete the traditional CCSS sequence and students who complete the integrated CCSS sequence. Do they have the same depth of mathematical knowledge? Do they continue taking math beyond 3 years at the same rates? Do they have the same facility of finding math in regular problems?

Ms. Schallau said she would like to see more education that promotes student exploration and innovation. She applauded the CCSS problem-solving approach but believes that we should "think further outside the box." Digital access to textbooks makes sense in the current era, said Ms. Schallau, and she wondered whether learning outcomes differ when teachers and students are involved in creating the text for their classes. She said that when she has a question, she uses Google to find the answer, and she asked why we do not promote that kind of learning in class. Such an approach could affect a student's confidence in his or her ability to find information, provided students are taught about finding credible sources.

Discussion

Ms. Cahill said insufficient readiness and the compensatory thinking that students do at the middle school level that interferes with learning later are high-impact problems for practitioners that require research. Dr. Confrey said that there are developers who are designing games and interventions to improve readiness. Agreeing with Ms. Cahill, she cited the example of multiplication and division and the number of students who lack competence in it. This is a known area that affects students' progress in learning the complex topics in middle school, and suggests the value of attention to flexibility and fluency in readiness topics.

Dr. Middleton said that some of the supports provided to help students in the middle grades understand ratios, proportions, etc., can lead to an arrested development of understanding. That is, they use the supports and do not learn the fundamentals of arithmetic, which prevents them from engaging in higher-level math. Ms. Schallau pointed out that when a student has problems with spelling, we suggest using a dictionary. But if a student has a problem with addition, subtraction, multiplication, or division, we do not encourage using a calculator. She suggested focusing on teaching the easier facets of basic math to facilitate understanding of the principles and permitting students to use calculators for the harder problems. Ms. Schallau said online mental math games aimed at the high-school level offer one option. She suggested more creativity in teaching math.

Marielena DeSanctis asked whether any research has addressed the college model, in which classes have a master teacher and graduate assistant(s). She pointed out that some school administrators lack the knowledge to assess and critique teachers. Dr. DeSanctis said that digital curricula could not take the place of the teacher, but we need more efforts to avoid having poorly prepared teachers in classrooms. Dr. Confrey stressed that digital curricula are intended to support teachers, not replace them. She noted that some schools are experimenting with different approaches (such as large, multi-grade classes) and differentiating teachers by their talents. Dr. Confrey said she is not convinced that flipping the classroom (delivering didactic information digitally to allow for more discussion and engagement during the lesson) is worthwhile without careful examination of how the classroom time is used and the quality of the flipped "lectures."

Jennifer Barrett asked whether any research addresses when technology should be employed, especially given that teachers are struggling to understand standards. Often, new technology is used by teachers to cover the same old material. Dr. Middleton responded that little of the research on technology is relevant to situations

in which access is poor or knowledge of the technology is limited. Various types of technology have different features that can be useful for different purposes. Dr. Middleton suggested that teachers adopt technology soon and determine which technology to use by considering what they want to do, for whom, and under what conditions.

Jonathan Supovitz said that field research makes a distinction between teaching programs and learning programs. Dr. Confrey said she does not make that distinction. She pointed out that once digital curricula are in place, researchers can gather empirical data that show how individual students progress and whether they all need to follow the same path and link these behaviors with the variety of teaching acts in such a way as to link teaching and learning fundamentally.

How does new access to resources (whether through software, websites, or human resources) affect the professionalization of teaching, and how can researchers assess the impact on teaching and on standards?

Dr. Confrey said we have dramatically underestimated the challenges that standards pose for teachers in terms of learning progressions, etc. Teachers know they need to drill deeper into the CCSS, but their schools are also facing budget cuts and other challenges, so they do not have adequate support. Dr. Confrey suggested we focus on professional development that includes time to learn and ways to share what you have learned as you go. She said teachers respond well to videos, and more good video exemplars are needed.

Dr. Confrey further suggested that research focus on more effective mechanisms for reaching teachers and leveraging resources across states. To effect real change, she said, we should tie professional development to the curriculum and ask what forms of planning and just-in-time resources support are needed and what else teachers need to move from interpretation to action. Dr. Confrey also suggested studying massive open online courses (MOOCs) for professional development, such as are being offered at the Friday Institute for Educational Innovation at NCSU, and assessing whether technical support is needed to ensure that professional development is fairly distributed.

Ms. Schallau noted that in California, some teachers take on summer internships so they see how math is used in the workplace; she suggested more such opportunities for teachers. There is a huge disconnect between what is happening in secondary-level math education and what jobs require, said Ms. Schallau. Technology could help establish learning communities across states and open doors for teachers.

How can assessments be designed to provide useful information for teachers?

Dr. Middleton said teachers are drowning in data. We need to consider how to assess how students take what they learn in class and apply it outside the classroom. Assessment should be ongoing and focused on milestones and interim progress instead of summative assessments. We also need to understand what products students use (e.g., online notes), said Dr. Middleton. We must determine how to gather data and assess it in ways that are individualized to the student and do so not just for teachers but also for administrators and policymakers. Dr. Middleton also called for identifying common systems of interpretation among actors, so that we can eventually link the results of assessments and translate data into useful information.

Ms. Schallau said one useful strategy for formative assessment has been "exit tickets," by which teachers can see how a student fared with a problem and use that information to decide what to do next. Useful assessments can have an impact on teachers' mindsets, said Ms. Schallau. She also asked how to make students more comfortable with self-assessment. Ms. Schallau suggested that some key problems could be revisited multiple times over the year, allowing time for reflection.

Discussion

"Math has a branding problem," Mr. Thomas observed. By middle school, algebra 2 is perceived as useless and painful. Mr. Thomas pointed out that grit, resilience, and persistence in learning are important tools for thinking and making decisions throughout one's life. He proposed emphasizing those qualities. Ms. Schallau said teachers across all levels, including college, should work together. She said that algebra 2, as it is taught traditionally, has no relevance in most of today's careers. Teachers should talk about the role of perseverance when learning and doing math. What do students (and teachers) do when an incorrect answer is given? What learning can happen from that incorrect answer?

Dr. Confrey pointed out that the early statistics standards in the CCSS in math at elementary school appear in the context of measurement and do not do justice to the research on statistical reasoning and how it develops over time. Students enjoy learning statistics, she said, and it often is very useful in strengthening their number sense. A question is if and how a stronger treatment of early statistics can be returned to the standards without leading to "a mile wide and an inch deep." Proper emphasis and more research on the relationship between number reasoning and statistics could address this. She also indicated that with the introduction of the new standards, the question of whether to use an integrated or a siloed approach to high school math is still an issue that schools face at the practical level. The standards support both approaches, so schools need more research to help them consider how best to address this issue.

Dr. Confrey noted that the use of math in careers is a rich target, and she would like to look more closely at how people do work that involves math. Mary Visher said several models provide teacher externships so that they can bring experience from the workplace into their classrooms, such as career academies. Showing the relevance of 10th grade math to a career can be very engaging and motivating. Dr. Visher proposed researching the interaction between successful high school models (e.g., career academies and pathways) and the CCSS, specifically whether such models promote or impede successful implementation of the CCSS.

Daniel Heck questioned how we evaluate learning. It is not helpful to think about the learning trajectory as a means to an end—that is, the ability to answer a specific question. Rather, we should be helping students understand that the learning trajectory is about using prior knowledge to figure out something new, he said. In math, students do not learn ways to do math; they actually do the math, and using technology is not getting around the math but rather a way of doing the math. Persistence is how you learn, said Dr. Heck, and there are particular kinds of math persistence. We should think about the learning experience as a key outcome in itself, he concluded.

Ms. Barrett said that just like we are moving away from a math curriculum that is "a mile wide and an inch deep" to a more focused set of standards, we should prioritize and focus research on what will have the most impact for student and teacher performance on a daily basis. Dr. Middleton appreciated the comments about the applicability of math in daily life and how to judge success. He hoped teachers would evaluate their own success in helping students advance their math skills.

Dr. Confrey suggested the value of looking at the CCSS in ELA that relate quantitative reasoning with English, science, or history and suggested that these provide a means to improve learning.
The group adjourned for lunch at 12:16 p.m. and resumed at 1:00 p.m.

Transforming Classroom Practice in ELA

Moderator: Rebecca Kang McGill-Wilkinson, NCER, IES
Presenter: Catherine Schmidt, Washoe County School District
Ms. Schmidt shared findings from a grassroots effort headed by her and two other teachers. When the CCSS were adopted, teachers were not prepared for the transformation it required. Ms. Schmidt and her colleagues offered a group of teachers time to gather, learn from exemplars, teach what they learned in their classrooms, and then come back to the group and discuss their experiences. The approach took off, and 1,000 teachers

have since been trained. That is, they investigated what it really means to implement the CCSS, shared their experiences with their peers, and learned what they and their students can do.

On the basis of this effort, Ms. Schmidt said that what teachers need most to implement the CCSS is a support system for learning about teaching—specifically, an opportunity to apply the plan-do-study-act model or professional development they already know.

Plan. Planning involves preparation, such as creating video exemplars, and ensuring that teachers have aligned, appropriate curricular materials and the tools to evaluate them. To that end, teachers need to know which commercial products have been vetted, which contribute to building a coherent body of knowledge, and which represent a diverse mix of cultures. Ms. Schmidt noted that changing materials is not sufficient; the CCSS represent a shift in pedagogy.

Do. Doing means paying attention to how teachers teach complex ideas and motivate students to persevere. Teachers need to know what constitutes a rich and rigorous classroom conversation and how to facilitate that. They also need to understand what constitutes a satisfactory presentation from a student in a given grade level.

Study. To answer questions about what works in the classroom, teachers must study, and for that they need accurate assessment data, said Ms. Schmidt. Specifically, teachers need useful, timely data that shed light on why students can or cannot perform a given task. Formative assessments can be helpful. Ms. Schmidt said that teachers trained by her effort are using the evidence from their students and classrooms to determine what to do next. She said school districts should provide teachers with professional development that addresses specific areas in which teachers need to improve. Teachers also need access to the literature on best practices.

Act. Finally, acting involves responding intentionally to what the evidence shows about how students can master content. At this stage, everyone should be able to identify what to do, said Ms. Schmidt, on the basis of the previous steps. In the "act" stage, teachers need to implement interventions. The plan-do-study-act approach is "very Common Core," Ms. Schmidt noted, and it reflects what students are being asked to do. Moreover, it paves the way to replace instructional practices that are not getting results with equitable, research-based learning.

Presenter: Carol McDonald Connor, Arizona State University
Dr. Connor said that now that most states have committed to implementing the CCSS, she believes the most important question to consider is what kinds of instructional strategies will be effective in helping students meet new, higher standards. Much of our current stock of strategies (classroom discussion, integrating reading and writing, encouraging evidence-based thinking) is based on research that defined understanding as the ability to read a passage and answer related questions, yet 30 percent of students are at or below the "basic" level according to the National Assessment of Educational Progress (NAEP). Transformation is needed.

From her research on developmental trajectories around reading comprehension in high-poverty schools, Dr. Connor believes that when focusing on college- and career-readiness, we should keep in mind the importance of a strong foundation in reading. Among students who fail to build such a foundation early, the odds of making up lost ground later in life are very low, she noted. Education is the result of a cumulative process of learning over time. In building the foundation, effective educational strategies will take into account a classroom full of individual students with different needs at different times.

Dr. Connor said some other questions to consider are how to distinguish the fundamental components of reading (e.g., decoding) from the broader skills of reading and writing and how to help teachers understand the synergistic relationship among language, reading, and building meaning. Also, new standards require more

interactive discussion, so students will have to learn to respond to questions differently (i.e., in the context of discussion instead of responding to a teacher's question). Furthermore, Dr. Connor wondered what kinds of support will help students who enter school with no literacy or those with learning disabilities. In addition, to meeting new college- and career-ready standards, students will need to learn the skills of academic language, self-regulation, focus, and perseverance. To support the transformation required, all of these issues should be addressed.

Presenter: Catherine Snow, Harvard Graduate School of Education
Dr. Catherine Snow expressed pessimism about the potential for transformation through the CCSS. She said the approach lacks sufficient grounding in research and evidence. While there are longitudinal studies of reading development and component skills, we still lack understanding about the basic tenets of reading comprehension. Also, we do not yet have agreed-upon measures or assessments for these skills. As Dr. Connor noted, Dr. Snow agreed that much work is based on assessments of the ability to answer multiple-choice questions about text. The next generation of assessments of reading comprehension is being developed in conjunction with implementation of new standards.

Reading comprehension is the interaction that occurs when there is a good match of the reader's skill, the demands of the text, and the demands of the task. The CCSS raise the complexity of both the text and the task without paying much attention to the reader, said Dr. Snow, and in many cases, these are the same readers who have been failing with easier texts and tasks. She acknowledged that in some cases, we may underestimate what students can do, and placing higher demands on them may push them further. However, we do not know whether all students will respond to the higher demands equally. Dr. Snow is particularly concerned that good readers may respond well to more rigorous standards while poor readers fall to the wayside.

Usually, we do not ask students to learn new forms (e.g., more complex text) and new functions (e.g., evidence-based argumentation) simultaneously, as the CCSS are asking students to do, Dr. Snow said. Research should address how much scaffolding is needed and how steep the ramps should be (i.e., how intense the intervention to build up skills should be). While new curricula will not be enough to meet the new standards, we need to determine how far well-developed curricula, which incorporate new supports for teachers, can move us toward the goal.

Dr. Snow said she worries about the CCSS emphasis on close reading. More must be learned about when close reading works, for which students, and in what contexts, as well as how to counter the demotivating aspects of close reading. We still need to know what principles teachers should use to encourage close reading and how long they should spend on those principles. Teacher preparation is crucial, said Dr. Snow. Across grades, ELA teachers are usually good readers, but they are not prepared to lead classes in discussion, argument, critique, and critical thinking, and they may lack mastery of academic language and other skills that they will be required to teach under the CCSS. Furthermore, the CCSS differentiate skills in different topics (e.g., different types of reading skills are needed for ELA when compared with social studies), but we do not know the developmentally appropriate routes for achieving that differentiation. For example, there is no basis in research for the claim that differentiation should begin around grade 6, said Dr. Snow.

The CCSS emphasize discussion, and many teachers lack the skill to support classroom discussion, said Dr. Snow. Moreover, other factors affect the success of class-based discussions, such as a negative climate that does not inspire trust or provide students a sense of orderliness. We need to think about school structures as well as all the other components, Dr. Snow concluded.

Discussion

Dr. Gamse asked how we can learn from other large-scale programs, such as Head Start, to inform a research agenda on effectiveness. Dr. Connor said the lesson learned from Head Start is to ensure that the assessments tap into the skills we are trying to promote through the CCSS, so that the CCSS is judged on the outcomes

intended. Similarly, Mr. Tucker said, we should reach consensus on a core set of specific skills we believe students should have as a result of the new standards. To prioritize research, he said, we should consider which areas should be tackled in the near-term (i.e., within the next 3–5 years).

Ms. Schmidt pointed out that without aligned curricular materials, we should consider what we can do now to change teacher practice. Dr. Connor suggested looking at some of the existing research about what works (e.g., active listening, taking notes, engaging in discussion) to help students generate new information and connect ideas. She said teachers can ask themselves, "Is what I'm doing helping students be more constructive and generative in what they do?" Dr. Hall said CGCS is partnering with others to support a shift to new practices by informing teachers and encouraging them to share what they are learning.

Dr. Hill suggested looking to earlier efforts to implement new standards in math for lessons about what works. She noted that in many districts, at the request of administrators, teachers are creating their own materials and cobbling together Internet sources; based on her past research, she expects the results will be poor. We are asking teachers to take on very hard tasks without preparation, said Dr. Hill. For research purposes, how do we go beyond the need for professional development and materials to better understand what teachers need, she asked. Finally, Dr. Hill emphasized the need to consider what baseline data we should gather now to facilitate assessment later. The NAEP will have to align with the CCSS so that it can be used to evaluate progression.

Dr. Heck believes we know from past experience what components of learning are important, and we also know that some components are so important that getting on the wrong track can derail a student. He said we should try to understand what curricula can do to support learning and use teacher preparation and development to address what curricula cannot do. Dr. Heck said we should look at the whole and not try to attack all the pieces separately. Dr. Catherine Snow added that we know that professional development works better when it is linked to curricula.

Mr. McCormick stressed that teachers need answers now to respond to the rigorous new standards already in place. He asked where to find information about what works. Ms. Schmidt said teachers can often find answers when they get together and discuss their questions. It is not enough to tell teachers what to do, she emphasized; teachers must have time to strategize. Dr. Supovitz pointed out that by implementing the CCSS, we are defining standards and testing students against those standards, but we have left it up to teachers to figure out how to meet those standards. Nonie Lesaux pointed out that we lack a theory of change to underlie the implementation of new standards.

Ms. Schallau suggested empowering teachers with knowledge about how to conduct classroom discussion and with methods for sharing best practices with their colleagues.

Dr. Visher said that within 5–10 years, we should have a body of research about the CCSS and some models, as well as some rigorous evaluation of interventions to test.

Karen King reminded the group to think about the unintended consequences of previous reform efforts. She pointed out that in the past, "standards" have been conflated with "interventions." The standards (then and now) do not define how to teach. The unintended consequence of the National Council of Teachers of Mathematics' standards was that teachers had to figure out how to define the standards as well as how to implement them. As a result, most teachers did not change their practices, said Dr. King. Instead of a theory of change, we need a baseline built around intervention, and we have to explain what that looks like and provide professional development to support it, she noted. The hypotheses about how states will work together around the changing requirements are what is really different about the new standards, Dr. King concluded.

Ms. Ferguson asked how we can build room at all levels for experimentation that includes some latitude for failure. CEOs always talk about the instructive value of failure, and group problem-solving includes understanding failure, she said.

Dr. Wohlstetter pointed out that the current policy environment is restrictive and may become more crowded over time. If we stick to the state model, she noted, we will never get the collaboration needed for success.

Implications for, and Needs of, Special Populations: Students with Disabilities and ELLs

Moderator: Brett Miller, NICHD
Presenter: Lynn Fuchs, Vanderbilt University
Dr. Fuchs said the signature characteristic of a learning disability is severe, persistent underachievement in a student who has received generally effective instruction and who has intelligence in the normal range. She described the dominant model for addressing students with learning disabilities—inclusion—and the major alternative—intensive intervention. Despite reforms to promote inclusion, achievement among students with learning disabilities has stagnated. Recommendations for implementing the CCSS for learning disabilities continue to focus on inclusion. Rigorous studies of intensive intervention are needed to help students with learning disabilities profit from the CCSS.

Dr. Fuchs analyzed data from a school district that moved from No Child Left Behind (NCLB) toward the CCSS during a timeframe in which students with learning disabilities were randomly assigned to either inclusion or intensive intervention. For the inclusion group, the gap in math achievement compared with their peers grew as the math standards got harder. In the intensive intervention group, the gap narrowed and, in some cases, closed.

Dr. Fuchs said the results of this research suggest strong potential for intensive intervention and signal that relying solely on inclusion under the CCSS may worsen the situation for students with learning disabilities. She called for more research and proposed research using large-scale descriptive studies to (1) illustrate how schools teach students with learning disabilities now, (2) describe student outcomes as a function of inclusion vs. intensive intervention, and (3) estimate the amount of variance in student outcomes attributable to inclusion vs. intensive intervention. Dr. Fuchs also suggested rigorous evaluation studies that index the proximal outcomes, the CCSS assessments, and college- and career-ready outcomes for students with learning disabilities. Such studies should include randomized controlled trials (RCTs) of widely adopted inclusion practices and promising intensive intervention practices, as well as RCTs that compare strong inclusion and strong intensive intervention practices head to head and that contrast promising practices with previously validated programs. Research should also identify the characteristics of students with learning disabilities associated with response to intensive intervention, to expand the framework for intervention in the context of the CCSS.

In addition, Dr. Fuchs pointed out the need for research on the validity of the CCSS for students with learning disabilities—for example, longitudinal research assessing whether the CCSS learning progressions represent the sequence in which students with learning disabilities develop college- and career-ready competence. She also noted the need for research on teaching ways to examine methods for making frustration-level content instructionally meaningful to students with learning disabilities and on studies examining the unintended consequences of practices that inadvertently decrease access to college or high-performing careers.

Presenter: Nonie Lesaux, Harvard Graduate School of Education
Dr. Lesaux said she focuses on ELLs, a group that demonstrates how much of learning is language-based. By the time the average ELL in a linguistically diverse setting goes to middle school, he is one standard deviation below his peers in vocabulary, despite being at an average level of reading, she noted. For ELLs, reading

comprehension does not extend to word knowledge. Dr. Lesaux said that instruction is organized around early word reading but not the meaning-making that students need from the outset to support, sustain, and keep up with the inferences and language demands of the curricula. Importantly, ELLs demonstrate the same or better growth in reading and related skills as their monolingual, English-speaking peers. We have seen small but promising effects with interventions that target word reading, knowledge gaps, and meaning-making, said Dr. Lesaux. Further research could apply what we know from developmental studies, standard practices, and interventions to the CCSS, she noted.

Dr. Lesaux said we have never met the needs of ELLs at scale, and now the CCSS are raising expectations for language skills in the overall population. Under the CCSS, those with good academic language skills will develop even further. In theory, the CCSS promise to provide the kinds of education that ELLs need, but questions remain about the unintended consequences and the result of holding students to expectations (via standardized testing) for which we have not provided scaffolding.

The research questions for ELLs have the same objectives as those for ELA students in general. How can we detect early signals that students need more help? What do language trajectories look like in different students? There is a very limited amount of large-scale research that tells us what to expect with respect to accelerated learning or what growth should look like in that context. Dr. Lesaux said we should think about the role of text in the curriculum as a platform and help teachers carry out meaning-making. There is a lot of rhetoric around the very specific elements of ELL instruction, but there is very little empirical data, said Dr. Lesaux.

Presenter: Stephen Elliott, Arizona State University
Dr. Elliott put forth a conceptualization on the basis of 50 years of research that the opportunity to learn is a function of time, content, and quality. Together, these three variables do a pretty good job of predicting student achievement. He summarized some of the research on the opportunity to learn for students with disabilities. Most teachers are unable to cover all of the required academic content standards before students take state assessments, and students with disabilities in inclusive classroom settings receive even less of the required content. Time spent on instruction and the amount of content covered are the biggest predictors of achievement, said Dr. Elliott.

Dr. Elliott said that better measurement of the opportunity to learn and development of instruments teachers can use to track opportunities to learn could form the basis of professional development. Existing methods to gather information about the opportunity to learn are insufficient. Dr. Elliott proposed that research address the following topics:

- **Policies and systems.** Students with disabilities do not get equal instruction time or content coverage, especially those in inclusive classrooms (the pervasive model). More research is needed to determine whether gaps in opportunity to learn are systemic and why they occur.
- **Embedded professional development and interventions for teachers.** Teachers need more support to do their work. Interim assessments of students' achievement along with measures of the opportunity to learn could provide teachers information on ways to improve their instruction and the content on which to focus.
- **Instruction and effects on performance.** Looking at the components of the opportunity to learn frequently (as often as daily) could improve understanding about the progression of learning. Ideally, researchers could follow the progression of the same students over many years.
- **Test score validity and accountability.** Many students are not taught the content on which we test them, yet we use the test results to make inferences about achievement and quality of instruction.

Discussion

Ms. Cahill pointed out that diversity within all of the special populations discussed poses further challenges to teachers. Dr. Lesaux noted that the largest and fastest-growing group of students is U.S.-born children of Latino immigrants. These children may be in preschool but are not necessarily monolingual English speakers. It is difficult to get good descriptive data on these children, some of whom may be long-term ELLs or may get "stuck" at an intermediate stage. Heterogeneity abounds, and in the United States, it is further confounded in groups with high-linguistic diversity and low incomes, said Dr. Lesaux.

Dr. Fuchs said differentiation of instruction for students with varying needs and skills within a regular classroom is an attractive idea, which has never been realized in an effective manner or in a way that a typical teacher can manage. She believes there are models teachers can use to differentiate more effectively and more fruitful ways to provide the opportunity to learn for more students (e.g., building classroom routines that systematically provide opportunities for differentiation). Dr. Fuchs also pointed to the need to identify ways that assessment data can be communicated to teachers with instruction recommendations. Teachers receive a lot of assessment feedback on student performance, but this feedback needs to be organized in ways that communicate clearly about how to connect that feedback to effective instructional practices that support differentiation (such as peer tutoring, technology-supported strategies, or small-group instruction) and about which activities are suited for which students, she noted. Ultimately, we need to align research with instructional recommendations, Dr. Fuchs observed.

Dr. Elliott said he was surprised by the lack of differentiation that teachers employ, suggesting teachers may need more training on differentiating in the classroom, particularly how to do it while they are implementing new standards and practices. He added that the amount of noninstructional time that students with disabilities spend in the classroom keeps coming up as a factor in their progress.

Dr. Middleton asked how to conceive quality of instruction as it relates to each of the special populations discussed. Dr. Elliott responded that he looks first at the cognitive emphasis and expectations of the teacher, an important predictor of performance that is reflected in how teachers ask questions. He also looks at evidence-based practices (e.g., using visual representations, think-aloud exercises).

Dr. Fuchs said that broad instructional dimensions (such as the opportunity to respond) are important, but researchers often fail to address the content-specific nature of instruction that distinguishes higher and lower quality. For example, to meet new learning objectives, classroom teachers in her sample focused almost exclusively on teaching tricks for finding solutions to fraction problems. Those tricks can pose barriers to students achieving understanding about important principles regarding fractions.

Dr. Lesaux said that ELLs need opportunities to produce oral and written language that go beyond question-and-answer, which demonstrates the need for a different environment to learn language. Dr. Fuchs noted that students who respond to questions in class already have the skills to do so, while those with learning disabilities may not even try and are not encouraged to persevere.

Mr. Thomas said the discussion shines a light on the tendency for special education teachers to "stay in the shallow end of the pool." He said special education teachers in his school are generally the most empathic but also the most insecure about their content knowledge. He asked how teachers can be encouraged to embrace academic rigor and combine it with their strong skills dealing with students with special needs. Dr. Elliott said that most teachers say they need more time—for example, more time to cover content with students with special needs before they take the state assessment tests. Dr. Fuchs added that most schools have only one or two special education teachers, and those teachers cannot be expected to have content knowledge that covers the entire curriculum.

Dr. Gamse pointed out that we have an opportunity to compare the effects of the CCSS across states that have very different definitions and resources available for students with learning disabilities and ELLs.

Dr. Confrey asked whether there are lessons to be learned from previous school reform efforts about particular types of scaffolding needed for students with special needs. Dr. Fuchs said there are few studies describing what is needed to make curricula like the CCSS work for students with disabilities, and it is important to frame a research strategy specific to the CCSS and students with disabilities.

Dr. Hill noted that math knowledge among special education teachers is somewhat below the national average, which poses a serious concern. She agreed with Dr. Fuchs and noted that she is worried that special education teachers may be marked down on observation metrics when their students are not "producing" CCSS-aligned thinking and reasoning; therefore, evaluation of special education teachers should be specifically addressed, Dr. Hill said. Dr. Fuchs agreed but noted that instructional methods exist that improve response among students with learning disabilities. She also noted that students with learning disabilities should not be penalized for not reaching expected goals. If progress toward goals is not recognized, then there's no incentive for schools to work hard on behalf of students with disabilities. Moreover, schools will lose good teachers. The same applies to ELLs, said Dr. Lesaux.

Dr. Supovitz said that in schools, ELL teachers play a central networking role in providing instructional advice and information about reading and ELA but are completely on the periphery in math. He asked whether ELLs and students with special needs have different outcomes in math and ELA under the CCSS. Dr. Lesaux said that ELLs dispel the prevailing myth that math is language-free, and Dr. Fuchs added that language comprehension is a strong predictor of both math and ELA performance. However, math is an area in which students with a history of low performance can make great progress with instruction, said Dr. Fuchs, underscoring that such students can and do learn.

District and School Leader Support for Implementation of College- and Career-Ready Standards

Moderator: James Benson, NCER, IES
Presenter: Marielena DeSanctis, Broward County Public School System
Dr. DeSanctis described her experience first as a high school principal adopting the CCSS and then as a curriculum instruction leader for the system. As a principal, she recognized that the CCSS were not aligned with the state assessment, and she focused on gathering performance and trend data into a dashboard for rapid assessment of progress. As a curriculum leader, Dr. DeSanctis worked with a group of teachers to develop a website, www.definingthecore.com, to provide professional development on demand. The website gives information and invites public discussion about the CCSS to minimize the chance of a backlash against the CCSS from people who are poorly informed. In addition, the website gives parents a better idea of what students should be doing in class, so teachers are getting the message about changing their practices from both sides, Dr. DeSanctis said.

Professional development opportunities and webinars all focus on how to change teacher practices in specific areas of ELA and math, said Dr. DeSanctis. She combined observation data with teacher surveys about the CCSS implementation and found that administrators and teachers have different perceptions about progress. To get a handle on progress, the school system categorized schools as silver, gold, or platinum on the basis of their success so far with the CCSS and created some teams to work with those at the lowest levels of performance.

Dr. DeSanctis said progress in her school district is complicated by the fact that different offices are responsible for curriculum instruction and teacher evaluation. The system also struggles with a lack of good baseline data about current teacher practices. Also, even as the school district is pushing for more professional development, the state is overhauling its professional development systems. Moreover, in

addition to state assessment and other tests, the state now requires end-of-course examinations for every course. As a result of the shifting landscape of school reform in Florida, the state has four cohorts of students in high school with different sets of graduation requirements. Computer systems to track progress and communicate with parents about changing policies need to be updated.

One unintended consequence of new teacher evaluation requirements is that teachers are no longer willing to swap classes temporarily in an effort to take advantage of another teacher's expertise or skill in a given topic, a tactic, Dr. DeSanctis said, she occasionally employed as a principal. She is working with local university faculty to build teachers' content knowledge. The system is struggling with instructional resources, and teachers still want to purchase curriculum kits just for the activities and materials.

Presenter: James Spillane, Northwestern University
Dr. Spillane said one important lesson from the implementation of college- and career-ready standards so far is that despite increased state activities around instruction, especially in ELA and math, and increased influence from national, federal, and state groups, school districts remain the source of fundamental infrastructure for schools. Over the past 25 years, school districts have been the entities that create policy to guide classroom teaching in ELA and math, and, especially in light of NCLB, other subjects have fallen by the wayside. Policies and programs are critical, said Dr. Spillane—they form the basis of instructional guidance that teachers and administrators use to decide what to teach. At the same time, the level of guidance differs radically across school districts.

In part, the variability stems from differences in human capital—both the intellectual capacity at the school district level and the number of bodies available to do the work. School districts interpret standards and devise their own ways of putting them into practice. People tend to construct fairly conservative understandings of how to put standards into practice, said Dr. Spillane; when new information arises, we distill it to the simplest form. For example, if math is problem-solving, we focus on using story problems to teach math. Many school districts do not have sufficient support for ongoing interpretation of the meaning of new standards, Dr. Spillane observed. In implementing college- and career-ready standards, it is critical to recognize that the infrastructure that school districts create will be the primary source that teachers and leaders use to respond.

Along those lines, Dr. Spillane suggested evaluating those school districts that support knowledge development about teaching ELA and math at different levels. The findings could be archived and made accessible to others. In considering how to measure the guidance that school districts provide, Dr. Spillane suggested looking closely at materials for overall coherence. In addition, we should consider whether the guidance provides a surface understanding of the material or a deeper understanding of what students are learning. We should consider how to create infrastructures at the school district level that support ongoing learning for administrators and teachers, said Dr. Spillane. We should also determine the extent to which an infrastructure is anchored in learning, as the formal structure of school districts is disconnected from classrooms. Finally, evidence suggests that high-stakes accountability at the state level ensures that local school districts pay attention to state requirements, although the impact on learning is mixed.

Dr. Spillane pointed out that reading and ELA have better infrastructures in place than math does. He stressed the need for school districts to organize support for learning in a way that takes subject matter seriously.

At the school level, principals play an important role, but given the vast array of administrative staff schools have, Dr. Spillane recommended thinking more systematically about the use of human resources. Instead of focusing all of our attention on principals, it may be helpful to consider the roles of others in managing and leading teachers. Dr. Spillane also suggested considering the relationships between schools within school districts.

Presenter: Paul LeMahieu, Carnegie Foundation for the Advancement of Teaching
Dr. LeMahieu said it remains challenging to put complex ideas into practice effectively, reliably, and at scale. Change requires a combination of ideas, will, and execution, he noted. While we have now a solid base of ideas about how to improve education, we need more knowledge about how to execute those ideas in ways that go beyond calling for more professional development. He outlined his thoughts about what research should entail.

The first step is reconceiving the challenge of implementation, said Dr. LeMahieu. From a scientific point of view, we prize fidelity of implementation. However, we have a growing set of ideas that have been proven to have good effects but little knowledge about implementing them to get those effects, because context matters. One approach that takes into account implementation in context is the application of improvement science to promote integrative adaptation (or, what some have referred to generically as design-based implementation). Under this approach, ideas are tested in context and refined to ensure effectiveness. The concept stems from manufacturing initially, but more recently in the social services sector, including health care where workers test out potential solutions in practice and adapt them as appropriate to promote continuous improvement.

Dr. LeMahieu said his organization promotes improvement research through networked improvement communities, which promises opportunities to accelerate learning; address real and complex problems in practice; and engage a broad array of expertise. The networks are professional communities of academic scholars and experts working together on high-leverage problems. They are similar to communities of practice but also are (1) focused on a common measureable aim, (2) guided by a shared definition of the problem and the system that produces the problem, (3) disciplined by the methodology of improvement science, and (4) coordinated by the application of tenets of network and knowledge management.

Dr. LeMahieu called for developing a science of putting complex ideas into practice. He suggested the following areas for research:

- Identify and learn about the types of problems that are amenable to solutions based in improvement science.
- Explore knowledge management (the culling, organization, and presentation of knowledge) so that knowledge can be spread more efficiently and put to work throughout the education community and beyond.
- Develop and adapt more tools to support improvement science. Foster the development of analytics and other mechanisms that can support genuine insight.
- Develop, test, and refine principles and guide the development of practical measures to support improvement science. (Whereas measures of accountability prize construct and conceptual validity, measures of improvement value predictive validity.) Ideally, measurements reveal leading and lagging indicators.
- Build up the field. Determine how to better prepare educational practitioners to provide them with the identity and tools to see and participate in improvement as an element of professional responsibility and give them the tools to produce improvement related to knowledge.
- Frame the work of educators in terms of processes—that is, as a rational enterprise of complex interdependent processes aimed at producing anticipated outcomes.

Finally, Dr. LeMahieu said, we need to explore when standardization is appropriate and when variation in practice is good. As a rule of thumb in education, he said, a variation introduced for the benefits of adults or the system should probably be controlled, while a variation that benefits students should be promoted.

Discussion

Dr. Connor said that research-practitioner partnerships have the potential to be powerful. Dr. DeSanctis commented that her system is partnering with several universities and community colleges. She described one case in which the research partner took on a classroom role and learned a great deal about the many challenges teachers face beyond lesson planning and teaching.

Dr. Middleton noted that a research agenda on understanding and supporting leadership is critical. We know very little about how leadership works, and it is a promising area of research, he said. Also, we need to be able to identify practices that align with the intent of the innovation and try to figure out how individuals and social systems can come to be so out of alignment. Dr. Middleton pointed out that some schools exhibit all the trappings of standards, reform, and college- and career-readiness, but the guts of those efforts either do not exist or are misaligned. He also noted that measurement models should account for behaviors that do not fit neatly on a continuum for rank ordering.

Ms. Cahill said that in the past, problems have been defined and interpreted too simplistically. She asked how we could encourage people to specify problems at a more sophisticated level.

A participant agreed that research should focus on understanding leadership as a practice. The difficulty of defining problems stems from the fact that organizational routines and processes are not well developed. Dr. LeMahieu said the problem of specificity begins with seeing work as a series of processes, something educators do not usually do. Instead of an organizational structure, educators would benefit from having a process map, he said. Ideally, interventions are problem-focused and user-centered, he added.

Dr. LeMahieu proposed adopting the term "integrity" over "fidelity" of implementation. "Fidelity" suggests a mandate to do exactly what is prescribed, while "integrity" suggests understanding deeply the ideas that animate the reform effort and remaining true to them while adapting processes for local implementation. "Integrity" suggests the need for users to understand the ideas behind an intervention and stay true to them while testing their implementation to ensure that it is effective. We need to explore how interventions are designed, Dr. LeMaheiu insists, because designing for integrity in implementation is a different process leading to a different designed object than designing for fidelity (with which we have greater familiarity).

Dr. Spillane added that schools tend to focus on one type of data when they define their problems. They also tend to come up with interesting hypotheses when data are shared.

Mr. McCormick suggested thinking about what needs to be different and gathering data on what seems to work.

Dr. Hill supported the need to evaluate coherence (of guidance, materials) as part of the research agenda. Given the competing demands on schools and school districts, Dr. Hill noted that research cannot focus solely on implementation of the CCSS but rather must consider how the CCSS fit into the systems that teachers and school districts face.

Dr. Hart pointed out that the discussion throughout the day did not define implementation. Research can tell us what to do but not always how to do it, he said. Also, the question often overlooked is how well an intervention was implemented—that is, how do we know whether it was implemented with fidelity. Ms. Schmidt agreed that integrity of implementation is a big concern. She pointed out the need to consider the integrity of the process as well, by considering what we tell teachers and what we ask from them. Accountability and leadership have to go together, she said. Furthermore, once we define the problem and build systems to address it, we have to have some faith in those systems, Ms. Schmidt noted.

Dr. Spillane said we could gather insight from efforts to build systems to support learning about the CCSS at the school level, by school districts, and by charter schools. Investments could be made in promising practices, he said. There is a lot to be learned from comparing and contrasting across schools and infrastructure designed for different populations of students.

Dr. DeSanctis said practitioners would appreciate help from researchers to better understand the context of their research so that practitioners can adapt the interventions appropriately.

Dr. LeMahieu said the many threads of discussion throughout the day support his suggestion that we need research that focuses on the execution of ideas. He noted that IES has recently put out a request for proposals for improvement research, and he hoped someone would run some analytics data on respondents because he suspects that there will be considerable interest in the program, even given its very modest scope, because the field appreciates its value and is deeply interested in it.

Dr. Brock thanked the group for the rich conversation of the day and adjourned the meeting for the day at 5:10 p.m.

Day Two: August 20, 2013

Opening

Thomas Brock, NCER, IES
Dr. Brock welcomed the group at 9:04 a.m. He thanked the group for the good discussion of the previous day and hoped today's meeting would provide an opportunity to winnow down the areas of interest to identify some major research objectives and priorities.

Measuring Classroom Practice to Improve Instruction

Moderator: Karen King, NSF
Presenter: Jonathan Supovitz, University of Pennsylvania
Dr. Supovitz encouraged the TWG to think about a framework for learning over time, noting that standards represent learning objectives over the broadest arc of time and, at the same time, serve as an organizing principle for education along with other policy instruments. Class instruction, on the other hand, addresses learning on a daily basis. Dr. Supovitz noted that by measuring classroom practice with annual surveys, infrequent activities are often overrepresented and frequent activities are underrepresented. Therefore, some alternative measures focus on gathering data immediately, through frequent logging or sampling of class practices at the end of the day, for example.

Dr. Supovitz emphasized the advantages that technology can provide in measurement, such as videotaping classroom instruction. He described an online tool (Teacher Analysis of Student Knowledge [TASK]) that helps teachers understand how to assess student work in relation to the expectations of the CCSS. Daily online logs are relatively quick and painless for teachers to use and yield a sophisticated stream of information. A New York City project provides teachers with an application to document the strategies that students employ, which informs researchers about how teachers organize their work.

Dr. Supovitz noted that the researchers may gather and assess data, then use the results to influence practice eventually. Alternatively, measures can be used more immediately as a sort of professional development tool, particularly in the context of iterative research (e.g., design-based implementation research). Dr. Supovitz described another tool, the ELA Common Core Observation Guide, which can be used as a measure but also creates a common language of instruction and understanding of what implementation means and how it looks at different levels. Such a tool builds coherence because the observers and teachers are aware of the language used.

In summary, Dr. Supovitz said that as long as there is consensus about which domains are important to measure, it is acceptable for various parties use different measures. Technology provides the opportunity to expand how and what we measure.

Presenter: Heather Hill, Harvard Graduate School of Education
Dr. Hill pointed out that the results of surveys of teaching practice are difficult to interpret, because many teachers believe they are implementing the CCSS, while classroom observation shows they are not. She noted that recent research findings indicate that information from surveys of teachers' knowledge about their students reliably predicts student outcomes.

Measurement of the CCSS in math is complicated by the fact that the standards ask for student thinking rather than teacher behaviors; students think and reason mathematically in their heads, so there is less to observe than standards that specify teacher actions. Moreover, "It is hard to measure what is not there," said Dr. Hill. She described her extensive work to develop observations to assess the implementation of CCSS-like classroom mathematics practices, but she was disheartened by the results of the methods of effective teaching (MET) study, which showed that in most cases, instruction is didactic and students say little during class.

Observation instruments geared toward measuring infrequent practices, like those implied by the CCSS, will have a difficult time demarcating more average teachers, and thus may have low reliability.

Dr. Hill added that it is challenging to determine how much students are learning during a math class, particularly if the class breaks into small groups (because videotape does not capture what is going on in all the groups). Teachers can benefit a great deal from videotape, however, by identifying opportunities for improving instruction.

Also, it is difficult to know how many lessons should be observed to provide a good measure of the construct of interest. In most states, the number of observations is arbitrary. Dr. Hill also noted the importance of high-quality raters, because researchers need to know that their assessment criteria are being applied to subjects in a uniform way. Ensuring high-quality raters includes activities such as prescreening applicants, designing rigorous certification tests, providing ongoing professional development around the instrument, and monitoring raters over time to prevent drift.

Finally, Dr. Hill said she is becoming more skeptical about the usefulness of current measurements, as the most consistent predictor of learning appears to be classroom behavior (i.e., students are in their seats, "not climbing the walls"). In a few years, she said, we may know whether the mathematics tests we have now are good at capturing anything other than behavior in the classroom.

Discussion

Dr. Elliott posited that better behavior may translate into more time spent on the task at hand. It may also represent the skill of a teacher in keeping students engaged.

Mr. McCormick asked whether it would be helpful to focus on certain parts of the ELA Common Core Observation Guide and work with teachers on putting in place observations and improving instruction around those parts. Dr. Supovitz said that any time teachers have the opportunity to slow down and be more deliberate, that is helpful. He believes that some observation and measurement instruments can be used as an alternative to traditional professional development approaches. Dr. Hill said that measuring a lot of indicators can be frustrating, and she applauded the notion of narrowing the focus to fewer items. Ideally, teachers are attending to the measurement instrument that will be used to evaluate progress, she said; if it is a good instrument, it prompts teachers to do a better job.

Dr. Catherine Snow posed the question of what might be the worst possible consequences of using measurement to improve instruction. Examples given include teachers insisting on including all of the content of a lesson, even when some students are years behind their peers in grasping the content; teaching time focused entirely on the text; or discussions that bring in evidence that is not in the text. Dr. Supovitz said one researcher provides teachers with tablets to note throughout their lessons what they are covering and at what grade level.

Dr. Confrey asked how the shortcomings of observational research are being addressed. Dr. Hill said there are some data from small samples indicating correlations between student performance and tests that require more reasoning. Also, other studies are asking students to do more cognitive work and are aligning instruments to better assess whether, for example, students perform better when teachers engage more students in discussion, said Dr. Hill, and we will know in a few years whether these hypotheses are borne out by research.

Dr. Confrey asked how researchers see the development of a theory of action or framework of learning over time. Dr. Supovitz said researchers are either working at the level of individual classrooms via annual assessments or of standards over the long term, and we need something in between to better understand the arc of learning. Dr. Hill agreed that research is limited; she said that she thinks an offshoot of the MET study

has taped classrooms up to 50 times a year, and such tapes could be a starting point for looking at the architecture of learning.

Dr. Middleton pointed out that much research does not specify dimensions of practice, such as frequency, intensity, depth, duration or density of instruction per time unit, synchrony (relationships among lessons), alignment/focus, reach, and feedback. Qualities of learning are difficult to measure. The goal of research is to determine what configuration of practices influence students to think deeply enough so they reach a more productive level or more advanced understanding of the topic, said Dr. Middleton. In the classroom, it is hard to predict what students know or do not know, and very different practices seem to result in learning and understanding, he said. Learning is strongly influenced by social factors, Dr. Middleton pointed out; at the high school level, the number-one factor in math is a sense of belonging in the school, he said. He emphasized the importance of considering operational definitions and determining how precise they must be to understand what we are seeing and then to generalize the information across the field.

Dr. Hill agreed that it is very difficult to assess quality, so she uses instruments that allow for some variation (e.g., distinctions between major and minor errors). Dr. Middleton said he knew of no publications about how to establish and maintain a group of researchers (including raters) that uses a common language and approach and Dr. Hill should consider sharing her experience.

Dr. Connor said she and her colleagues conduct classroom observations at the level of the individual student in an effort to understand patterns of individual differences, the impact of these patterns, and what might be considered effective instruction for different students. Their research reveals that some students need more of a certain type of individual instruction than do students with a different constellation of literacy skills. The research further indicates that outcomes are not affected by the total amount of literacy instruction but rather how well the amount and type of instruction matches the needs of individual students. She encouraged researchers not to dismiss the importance of understanding what individual students are learning. Dr. Connor said she finds more variability among students within a class than between classes, and researchers can lose that insight into variability when they only look at whole classrooms. As we ramp up for the CCSS, we should look not just at what teachers need to do differently but also what we expect students to do differently, Dr. Connor suggested.

Dr. DeSanctis pointed out that we are asking administrators to act as researchers (e.g., to observe classrooms and develop professional development to address concerns), but administrators are not researchers. Instruments like the TASK website described by Dr. Supovitz are exciting and will be helpful. She noted that a lot of teacher resources for implementing the CCSS focus on ELA, but math is more about understanding what a student is thinking, and we do not know what that looks like. Dr. DeSanctis hoped that research would bring some clarity to what good math instruction looks like so that school districts can use those findings for professional development.

Dr. Heck said some good surveys can provide useful measures, but they cannot measure quality. In terms of how often to observe or measure classroom practice, Dr. Heck said that in his experience, observation of multiple classes allows researchers to see the beginning and end of a lesson, but teachers never seem to get around to summarizing the content of the lesson for the class, no matter how often observation occurs.

Dr. Heck continued that it is important to have an architecture of learning over the long term—an architecture that is focused on student behaviors and instruction that supports the desired behavior (motivation, calling on prior knowledge, making evidence-based claims, etc.). Ideally, practitioners are using a framework that helps students move through the process of learning over time.

Dr. Heck noted that the CCSS should not claim to be the source of ideas about deep, rich engagement in math practice, for example. The CCSS must be implemented with quality as well as fidelity to be successful, he said, and more goes into quality than what is described in the standards. If standards are implemented with quality, they will work—but that is true of most things, said Dr. Heck.

Regarding lesson architecture, Dr. Heck said, timing is everything. The right question is only right if it is asked at the right time. Researchers can work with teachers to get a better idea of timing and what is going on in students' heads.

Dr. Wohlstetter pointed out the underlying theme of the discussion seems to highlight the relationship between teaching and learning. Teachers vary in their perception of the relationship between the CCSS goals and teacher evaluations. She asked whether schools could use a single instrument that addresses both the implementation of the CCSS and evaluation of teachers. Dr. Hill said the two most popular instruments do combine the two, and at least one of them has the potential to promote better instruction. She added that a lot of the variation in teacher value-added scores has yet to be explained, and the lack of alignment of teacher evaluation and implementation requirements adds to teachers' burdens. Dr. Supovitz said he is struck by how many teachers are practice-oriented but do not see the connection between practice and performance. Typically, teachers try to make inferences about their practices on the basis of student performance data, which is difficult to do. Efforts are underway to illustrate explicit connections between practices and student comprehension.

Dr. LeMahieu questioned whether one assessment mechanism can address CCSS implementation, accountability, and genuine improvement. The attributes of an assessment tool (e.g., timeliness, definitional attributes of technical rigor, the nature of evidence produced and feedback given) should be well suited to the purpose of the assessment and that suggests substantial differences in assessment for accountability, research, or improvement. Dr. LeMahieu called for the exposition of a science around different kinds of assessment with particular emphasis on assessment for improvement. Dr. Supovitz pointed out that there is an uncomfortable tension between how people respond to an environment of accountability and their willingness to accept feedback intended to improve practice. It is important to consider how the assessment is framed within a larger set of policies around incentives, rewards, and punishments, he noted.

Dr. Fuchs cautioned against drawing causal inferences on the basis of observational (correlational) data. She noted, for example, that Dr. Hill's data contradict experimental evidence showing that certain instructional features improve student learning (e.g., teachers encouraging students to compare/contrast solution strategies). Dr. Fuchs explained that Dr. Hill's classroom observations reveal little variation in the instructional methods teachers employ to teach math. Without variation in teachers' instructional methods, it's not *possible* for instructional variables to explain variance in student outcomes. Therefore, to draw causal inferences about the type of instructional features that support student learning, researchers and policymakers should rely on randomized control trials over observational data.

Ms. Schmidt noted that if the goal is to improve instruction, teachers must change their practices. Therefore, researchers should focus on working with teachers to determine what might work in the classroom, rather than coming in with a static approach. The relationship between researchers and practitioners should be more fluid than ever as the two partners investigate the outcomes of the CCSS. Dr. Supovitz added that such partnership is in itself a professional development activity.

A participant pointed out that various concepts about teaching practice affect how we measure practice, and the social interaction component cannot be adequately measured by aggregating behavior. It may be worthwhile to distinguish and measure student-teacher interaction, student-student interaction, teacher-class interaction, etc. In addition, we should be aware that some teachers are already using measurement instruments to guide their practices, regardless of whether the instruments were designed to support practice.

Mr. Thomas stressed the need to establish a foundation of trust to support effective evaluation. That foundation can be enhanced by providing teachers with coaching or mentoring that helps them build their capacity and expand their knowledge. Dr. Hill noted that some schools have implemented peer evaluation processes.

Ms. Barrett said it seems obvious that there is a disconnect between teachers, states, and researchers about what quality is and what we are looking for. We need to find a way to close that gap and speak the same language, she said. If we do, the research will be more useful for practitioners and researchers, and it will have more impact on instruction.

Strategy Session: Research on College- and Career-Ready Standards

Moderator: John Easton, IES
Presenters:
Beth Gamse, Abt Associates
Barbara Means, SRI International

Dr. Easton posed questions and topics to the presenters throughout the session (identified here in boldface and italics).

What are key themes from past day?

Dr. Gamse pointed out that every conversation of the TWG has revealed significant complexity across different systems. At its core, the CCSS and assessments are designed to be an overarching system encompassing many facets of the educational enterprise and were intended to create a system where there was none before. Notably, the intent of the system is to ensure that all students have equal access to learning experiences, something that not been part of our loosely tethered education enterprise to date, said Dr. Gamse. Another theme has been the need to identify and address all the actors involved—students, teachers, principals, school districts, etc.—and how their roles fit together.

The content of the CCSS has been a focus of conversation. Participants also raised the need to understand the processes that tie the content together: for example, professional development for teachers, continuous improvement strategies, etc. Outcomes have been a major topic, and some hope to categorize the many outcomes into manageable groups. It is hoped that research can identify some indicators that signal progress toward broad, long-term outcomes such as equity, learning, and performance, Dr. Gamse noted.

Dr. Means said she was struck by the participants' interest in ensuring that the CCSS are implemented on a deep level, not a surface level. There seems to be consensus that practitioners should apply what they know about good instruction, and "integrity of implementation" may be more important than "fidelity of implementation." Also, participants seem to be encouraging thinking more deeply about how the CCSS are configured and what constellations of components are important.

Dr. Means expressed surprise that the participants focused more on practice than student outcomes, which she feels is positive, but she noted that assessment should address both processes and outcomes and take into consideration the accountability systems that are in place. Student outcomes remain important to determine whether teacher practices are changing in ways that produce the kind of learning desired.

There is serious work to be done to articulate how the CCSS differ from previous standards and what they mean for instruction, said Dr. Means. She was struck by how poorly we understand many of the standards, as illustrated by the lack of definitions of key terms such as academic language and the push to learn new forms and new functions at the same time. She suggested working through multiple theories of action in response to the CCSS, because there is real tension between what we say the standards are and what they imply. We need

to better understand the tenets underlying the CCSS, such as the premise that giving students more challenging materials will lead to higher performance. We also need to articulate the standards so that we can identify the research questions and develop instruments to collect baseline data. In addition, more work is needed to develop the instruments to measure the right things.

Dr. Means said we need to be informed by pioneering districts about their challenges, practices, and burning research questions. These districts can help identify the most pressing research questions. A common theme of discussion has been the need for coherence across content standards, teaching approaches, professional development and supports, and assessment. At the school and district levels, we need coherence across initiatives (such as teacher evaluation and CCSS implementation). Finally, said Dr. Means, there was great interest in the use of new technology to support research.

Discussion

Dr. Catherine Snow said a huge research agenda can be built around developing interim outcome measures that ultimately feed into overall assessment. We need to identify the mediators of good performance and develop measures specific to those mediators, then evaluate those mediators periodically to assess our hypotheses. We need to understand if the mediators are processes (such as writing) or the skills that students develop as a result of processes.

Dr. Connor asked for input on whether the assessment tools in use will provide a fair, honest assessment of what the standards aim to achieve. Dr. Hart asked whether the assessment results will lead us to understand why students do or do not succeed. It is not clear whether the CCSS are more difficult than current standards or whether the assessments are more difficult, he noted. Dr. Hart said we need more information about the critical components and what the initial results of assessment mean.

Dr. Confrey worries that the new assessments will report results at the level of claims, not targets. For example, they may tell us more about what schools are teaching than about what students are learning. Therefore, we need to develop interim measures to assess the impact. In the future, formative and diagnostic assessments will be an area of innovation, Dr. Confrey predicted, and we will rely less on end-of-year tests to inform instructional improvement.

Mr. McCormick said that understanding the scoring and interpretation of the assessments requires partnerships between researchers and school districts. He hoped efforts would be made to identify examples of student work that demonstrate a high level of achievement in relation to the standards and to share them with teachers before implementing the assessments.

In crafting the research agenda, Ms. Ferguson called for attention to how research interacts with policy and how policymakers will interpret research efforts. The assessments are going to result in a wave of concern about how kids are tested, she said. It is important to consider now how school districts and states will communicate about the assessment results and the research involved. Articulating a theory of change will be particularly important, Ms. Ferguson noted.

Dr. Gamse agreed that communication with the public and policymakers is key. In doing so, it is important to step back and describe the incredible amount of work, engagement, encouragement, and vision required to get states and stakeholders to commit to implement the CCSS.

Dr. Gamse said that one conversation that has not occurred yet is identifying the actors that will make up the community of practice that is committed to doing meaningful work and that will be needed to ensure the success of any research agenda. We do not yet know what resources will be available to support such a community. Dr. Gamse suggested keeping in mind that individual research agendas will not look at all the ideas and priorities on the table, but groups of researchers can address those ideas and priorities collectively.

If we craft a coherent, comprehensive research agenda to implement over the next few years, what criteria should we apply and what important factors should it address so that the next generation sees that we had the interests of all stakeholders in mind?

Dr. Means said the goals of the CCSS are to reduce inequalities of opportunities to learn and free up resources to improve the quality of teaching and learning. What she hopes research will answer is whether implementing one set of standards nationally (rather than 50 state-specific sets) succeeded in improving equality of opportunity. In addition, before research can answer the question "Did it work?" it will have to identify what was implemented and how. She hoped we would leave a legacy of research on implementing change.

Dr. Means said she was heartened to learn how many people are applying lessons learned from other programs, and more research is needed to support such efforts. To help schools and school districts improve outcomes and close gaps in the opportunity to learn, Dr. Means supported Dr. LeMahieu's improvement science approach. School districts are not set up to share knowledge, she said, so we should create partnerships between researchers, practitioners, and those designing interventions. To build capacity and understand the impact of interventions, said Dr. Means, we should focus on the iterative (or plan-do-study-act) approach. She said iterative research and design refinement should be a major goal of education research.

Dr. Gamse suggested thinking about how the research agenda engages all the actors and looks at the processes in an integrative fashion. It may be necessary to have a master plan for the research agenda and to determine who will monitor that plan as the implementation context changes. Coordination across different settings should be a feature of an overarching agenda. The agenda should take into account the utility of research for policymakers, school boards, etc. and ensure that research findings can be applied in various contexts.

Dr. Gamse suggested considering what incentives might be offered to encourage research participation. Other domains, such as public welfare, have offered, for example, waivers to states that participate in large studies. Dr. Gamse also suggested considering how the lessons of previous research are valued and prioritized. Also, she asked, what tradeoffs will be required to conserve resources, and when will the research take place?

Discussion

Dr. Hart noted that we often discuss moving from research to practice, but miss a critical step that involves creating processes and models that allow practitioners to take good research to scale. Dr. Connor described the benefits of evaluating individual students, yet that may not be feasible in some cases, she said.

Dr. Visher said research should build on what we already know. She expects the system will spawn some promising strategies for professional development, for example, and it may be helpful to use those as a starting point for supporting evidence-based practices. Such an approach informs practice without adding to the burden of schools and school districts.

Dr. King said she sees multiple research agendas at different levels for different purposes. She would like research to address what a common approach—which the CCSS is trying to promote—affords an educational system like ours (that does not work like a system) and what it takes away. One research agenda could focus on the impact of the common approach (e.g., does a common approach to teacher education around academic topics lead to a unified teaching approach around the country?). Another agenda could focus on what happens in the classroom and how standards should be organized to support learning in specific topics. Other agendas could address issues at the policy level and the school district level. Different people would engage differently with each agenda.

Dr. DeSanctis observed that even though it is difficult to measure what is not there (as Dr. Hill described), we need more insight about where we are. We also need to understand whether our assessments are aligned with the CCSS and what school districts can use to assess whether student performance is improving. Dr. DeSanctis also called for research on how state policies moderate outcomes, how notions of academic freedom vs. a common core of knowledge play out, and what unintended consequences have arisen. She noted that the CCSS are supposed to increase collaboration; research should determine whether they have and, if so, whether such collaboration has a positive effect.

Dr. Spillane reminded participants that politics will play a role in the CCSS implementation. Political differences at the state level will play out at the school district level.

Dr. Middleton said research should include an effort to understand what it means to be college- and career-ready. There is some research around college readiness but little attention to career readiness. Higher education relies on class standings, SATs, and school ranking compared to other high schools, he noted; the SATs predict little other than freshman college performance. What students need to succeed in engineering school, for example, is not addressed by any assessments.

Dr. Catherine Snow expressed concern that the CCSS are an overreaction to the last set of bad practices, and that 10 years from now there will be a similar overreaction to the CCSS. One indicator of success will be the absence of such a backlash. She would like to see us eliminate the phrase, "research to practice" and attend to practitioner input into research. Dr. Snow hoped we could develop research mechanisms that foster collaboration with practitioners to define the problems and answer questions that cannot be resolved at the school district level.

Dr. LeMahieu said one criterion for research should be that it focus on enhancing the capacity of the system over the long term. If we agree, for example, that we have a vision of research and development that can guide researchers in working with practitioners, he said, we should explore how that capacity can be spread throughout the whole system.

Dr. Hill noted that research dollars can shift with fads. To explore the implementation of changes in the classroom, funding mechanisms should be revised to privilege more descriptive research, such as case studies, which can be powerful. Dr. Means agreed that quality research and improvement research should be grounded in what actors are doing. There is a place for RCTs and efficacy studies, especially once we have promising interventions fully developed and refined, she added. Successful interventions for the CCSS will not crop up overnight, Dr. Means pointed out. Typically, pilot-testing and refining interventions take a decade, but technology could speed up the process. However, if we try to address the whole research agenda using rigorous experimental design, we will fail, she said.

Dr. Easton said he would like to see a model for research that knows when we need a causal estimate and when other evidence is more appropriate. Good descriptive research has been undervalued for a while, he acknowledged. We need a system that incorporates feedback loops, recognizes the need for causal evidence, and knows how to ask the right question at the right time. Dr. Earle said a research agenda can be constructed to address questions in phases, so that in the early phases of implementation, you ask questions that can be answered on the basis of early interventions. She recommended a phased research agenda.

Ms. Ferguson said the common rhetoric around the CCSS is that they are used as a validation for ensuring college- and career-readiness. The research agenda should consider the expectations that will arise around the CCSS and whether research is in a position to control or address those expectations, she said. She would not like to see wrongheaded expectations undermine everything the CCSS offer.

Dr. Confrey said we cannot miss the opportunity to understand the baseline. We should develop the narrative now about what we think the baseline is and what data we can collect to address the goals (e.g., addressing equity, enhancing international competitiveness, and better understanding resource use).

Dr. Wohlstetter said the research agenda should have short-, medium-, and long-term goals. The key audience for the short-term research will be those in charge of implementation. Some tasks are common to all of the schools, school districts, and states—such as communicating about what the CCSS are. Also, focusing on descriptive research and common practices would be helpful for implementers. Such an approach could facilitate collaboration across states as they see opportunities to share resources.

What are our various roles in accomplishing our agenda? How do we develop a community of stakeholders to engage in the research agenda, especially around communication, feedback, and early descriptive work?

Dr. Gamse noted that several communities are participating in research in different ways, and the TWG represents one community of practice. Grantees working on similar problems; local, regional, or state-based education stakeholders; and practitioners within a discipline could all form such communities. We may need more than one community given the complexity of the CCSS research agenda, said Dr. Gamse. We should also think about how to communicate among and across the various stakeholders. Other models should be considered, such as developing regional centers similar to the NSF's engineering research centers (although not duplicating the IES Regional Education Laboratories), she noted. The timeframe for research affects the development of communities of practice, because it takes time to build rapport. Dr. Gamse added that a community could engage around specific outcomes of interest, such as how teachers learn or individual learning differences and equity of participation. It may be helpful to have a mixed model of different communities who will commit their time and intellectual energy to advance a common agenda.

Dr. Means said she believes we have a community of researchers interested in common areas, but we could improve in two areas. For example, the Carnegie model of a networked community is excellent, and we should have more mechanisms to gather people together to jointly negotiate around a key problem. In addition, efforts should be made to provide policymakers with research information that is synthesized, well communicated, and focused on key issues of concern. She reiterated her point that "research is helping, not just observing."

Discussion

Dr. Confrey suggested that in the area of digital transformation, the private sector is investing considerable resources in innovation. She suggested that an interesting model would be to foster partnerships for research on these innovations that would leverage the private investments with public research having the advantage of links into peer review and independent evaluations.

Dr. Catherine Snow said the publishers of curricula should be engaged early on. Dr. Gamse said that some studies in reading involved publishers in the enterprise. Dr. Earle said that NSF creating curricula had a positive effect on other publishers.

Dr. Means said that states working with researchers on policy issues could form a great community of practice.

Dr. Heck said that if a research agenda is defined and funded, it is important to communicate the agenda broadly so that researchers can determine where their own work fits and whether others are doing related research, which fosters networking. Such networking happens informally now, but it would be helpful to create a space that clearly outlines what is underway so that school districts, for example, could find research partners. Also, said Dr. Heck, some funding mechanisms create communities, such as centers or institutes,

and some projects give rise to mini-communities. Even a little face-to-face interaction helps inform researchers' work and improves dissemination, he said. Dr. Spillane added that some state and national business groups are interested in getting involved.

Mr. Tucker said he expects to see more aggregation platforms like Edmodo and other tools to bring people together to work on specific issues or collect data. On the workforce side, he emphasized that the CCSS are just a means to reach college- and career-readiness, so institutes of higher education and community colleges should be involved.

Dr. Hill said that practitioner/researcher collaborations are a great idea, but we do not have the capacity to put them in place at scale. She cautioned that such efforts are only one slice of the pie. Dr. Visher said collaboration is possible on a smaller scale—for example, practitioners, researchers, and workforce investment board members could engage with one school district so that practice and research inform each other.

Dr. Gamse said that one goal of the CCSS is to reduce the amount of data collection and the number of measures required over the long term. For the research agenda, it will be helpful to exert some restraint to avoid duplicating efforts needlessly.

Dr. Benson appreciated the participants' focus on short-term research approaches. There are many questions to answer about conducting short-term implementation research, particularly in the framework of federal funding opportunities. Partnerships are emerging as a way to do short-term research and fill the knowledge gaps between standards and practice. The more frequently partners communicate, the more quickly questions will be answered, Dr. Benson said.

Concluding Remarks

Thomas Brock, NCER, IES

Dr. Brock reiterated a point made by Dr. King in the previous session: that there is no single research agenda, but in fact there are several research agendas to pursue. He also expressed the need for future convenings to encourage dialogue and information-sharing between researchers, practitioners, and funders. As findings start to become available, one of the biggest challenges will be synthesizing and communicating major lessons to policymakers and supporting the efforts of states and school districts to make improvements. Dr. Brock categorized the input of the TWG into five agendas.

1. Partnership Agenda

The TWG indicated a need for strong partnerships between researchers and practitioners to focus on issues that states and school districts care about and that will help them make improvements in a timely way. One approach is to fund design-based implementation or continuous improvement research. IES is trying to build up such partnerships with some new funding opportunities announced earlier this year. Dr. Brock called for feedback about these new grants, and said that there may be interest in expanding the program if it is meeting the needs of the field and appears to be successful.

2. Descriptive Research Agenda

The TWG also identified a need for research that will help states and school districts learn from some of the "early adopters" of college- and career-readiness standards. Case studies and other high-quality descriptive research can illustrate how states and school districts grapple with major challenges and suggest strategies that others may want to adopt or avoid. Two issues worthy of special focus are strategies that school leaders and teachers use to translate the goals of the CCSS into practice. Another important topic is how the CCSS and related initiatives are affected by contextual factors such as teacher evaluation initiatives (which may or may not be aligned with the CCSS) and state and local fiscal conditions. Ideally, case studies will use mixed-methods designs, blending qualitative interviews and observations with teacher surveys and administrative

records data to document what changes in schools and classrooms, explore how these changes come about, and generate hypotheses on factors that lead to better outcomes.

3. Efficacy Agenda

The TWG agreed that it would be impossible to conduct a random assignment evaluation of the overall effects of the CCSS or related initiatives, but some participants felt that there may be a role for experimentation on specific strategies or choices that states and school districts are facing as they implement college- and career-ready standards. For example, RCTs could be used to assess the effectiveness of professional development programs for teachers, or to compare the effectiveness of different curricula in core subjects like English and math. Such studies could help states and school districts sort through the myriad options available to them and help them adopt policies and practices that are most effective.

4. Longitudinal Research Agenda

The TWG discussion revealed a need for longitudinal research to determine how students fare as they move from one grade level to the next and eventually into college and the workforce. Such research would provide important feedback to policymakers and school personnel on whether students are sufficiently prepared and progressing as intended. Looking across states and school districts, there may also be opportunities to exploit natural variations to identify policies, practices, and conditions that lead to better outcomes. While such research would not be causal, it could generate hypotheses that could later be tested using more rigorous designs. At present, relatively few states have good longitudinal data systems that track students from K-12 into postsecondary education and the workforce. IES has made some grants to states for this purpose, but much more needs to be done to increase capacity.

5. Measurement Agenda

Finally, the TWG identified a need for better data collection tools to capture changes in schools and classrooms. Given that implementation of college- and career-ready standards is already underway in some places, there is an especially urgent need for baseline data collection. Videotaping classrooms may be one way to start gathering information that can be used to assess changes in teacher practices over time. While it often takes many years to develop good measures, Dr. Brock hoped that we would not let the "perfect" be the enemy of "good." As discussed during the TWG meeting, some good measures are already available and may only require minimal modification to capture some of the more important attitudes and practices that are expected to change with the introduction of new standards.

In conclusion, Dr. Brock said the summary report of this meeting will be available to the public, and he hoped the ideas in it would be considered not just by federal funders but by foundations and others. He encouraged the TWG to continue providing their input to IES staff. Dr. Brock thanked the contractors who assisted with the meeting: meeting reporter Dana Trevas, administrative liaison Rayniece Anderson, and audio engineer Bruce Tran. Finally, he thanked all the participants for their time and for the rich conversation.

The meeting adjourned at 1:00 p.m.