J·T·L·A

The Journal of Technology, Learning, and Assessment

Volume 10, Number 4 · November 2010

Measuring Cognition of Students with Disabilities Using Technology-Enabled Assessments: Recommendations for a National Research Agenda

Sue Bechard, Jan Sheinker,
Rosemary Abell, Karen Barton, Kelly Burling,
Christopher Camacho, Renée Cameto,
Geneva Haertel, Eric Hansen, Chris Johnstone,
Neal Kingston, Elizabeth (Boo) Murray,
Caroline E. Parker, Doris Redfield, Bill Tucker

www.jtla.org

A publication of the Technology and Assessment Study Collaborative
Caroline A. & Peter S. Lynch School of Education, Boston College

Volume 10, Number 4

Measuring Cognition of Students with Disabilities Using Technology-Enabled Assessments: Recommendations for a National Research Agenda

Sue Bechard, Jan Sheinker, Rosemary Abell, Karen Barton, Kelly Burling, Christopher Camacho, Renée Cameto, Geneva Haertel, Eric Hansen, Chris Johnstone, Neal Kingston, Boo Murray, Caroline E. Parker, Doris Redfield, Bill Tucker

Editor: Michael Russell
russelmh@bc.edu
Technology and Assessment Study Collaborative
Lynch School of Education, Boston College
Chestnut Hill, MA 02467

Copy Editor: Jennifer Higgins
Design: Thomas Hoffmann
Layout: Aimee Levy

JTLA is a free online journal, published by the Technology and Assessment Study Collaborative, Caroline A. & Peter S. Lynch School of Education, Boston College.

Preferred citation:

Bechard, S., Sheinker, J., Abell, R., Barton, K., Burling, K., Camacho, C., Cameto, R., Haertel, G., Hansen, E., Johnstone, C., Kingston, N., Murray, E., Parker, C.E., Redfield, D., and Tucker, B. (2010). Measuring Cognition of Students with Disabilities Using Technology-Enabled Assessments: Recommendations for a Research Agenda. *Journal of Technology, Learning, and Assessment, 10*(4). Retrieved [date] from http://www.jtla.org.

J·T·L·A

Abstract:

This article represents one outcome from the *Invitational Research Symposium on Technology-Enabled and Universally Designed Assessments*, which examined technology-enabled assessments (TEA) and universal design (UD) as they relate to students with disabilities (SWD). It was developed to stimulate research into TEAs designed to better understand the pathways to achievement for the full range of the student population through enhanced measurement capabilities offered by TEA. This paper presents important questions in four critical areas that need to be addressed by research efforts to enhance the measurement of cognition for students with disabilities: (a) better measurement of achievement for students with unique cognitive pathways to learning, (b) how interactive-dynamic assessments can assist investigations into learning progressions, (c) improvement of the validity of assessments for students previously in the margins, and (d) the potential consequences of TEA for students with disabilities. The current efforts for educational reform provide a unique window for action, and test designers are encouraged to take advantage of new opportunities to use TEA in ways that were not possible with paper and pencil tests. Symposium participants describe how technology-enabled assessments have the potential to provide more diagnostic information about students from various assessment sources about progress toward learning targets, generate better information to guide instruction and identify areas of focus for professional development, and create assessments that are more inclusive and measure achievement with improved validity for all students, especially students with disabilities.

J·T·L·A

Measuring Cognition of Students with Disabilities Using Technology-Enabled Assessments: Recommendations for a National Research Agenda

Sue Bechard
Measured Progress
Jan Sheinker
Sheinker Educational Services, Inc.
Rosemary Abell
ROSE Consulting
Karen Barton
McGraw-Hill
Kelly Burling
Pearson
Christopher Camacho
Children's Progress
Renée Cameto
SRI International
Geneva Haertel
SRI International
Eric Hansen
Educational Testing Service
Chris Johnstone
National Center on Educational Outcomes
Neal Kingston
University of Kansas
Elizabeth (Boo) Murray
Center for Applied Special Technology
Caroline E. Parker
Education Development Center, Inc.
Doris Redfield
Edvantia, Inc.
Bill Tucker
Education Sector

Overview

The Invitational Research Symposium on Technology-Enabled and Universally Designed Assessments was held in Arlington, Virginia, on July 23, 2009. Measured Progress and SRI International sponsored this meeting focused on the emerging and dynamic field of technology-enabled assessments (TEA) and the principles of universal design for assessment as they relate to students with disabilities. The symposium brought a group of researchers together from several areas of expertise including educational technology, cognitive psychology, students with disabilities, universal design for learning, and educational assessment. Among the participants were researchers who had completed or were engaged in research involving technology-enabled assessment, universal design for assessment, and/or students with disabilities, focused on two specific areas: cognition and access. The state of educational assessment and technology had recently been described in an article entitled *Beyond the Bubble: Technology and the*

Future of Student Assessment (Tucker, 2009). Tucker drew attention to assessment challenges in the context of a cognitive model for assessment, and envisioned a future for assessment and technology that resolved many of the challenges. Tucker's article provided a foundation for the design of the symposium content, the meeting agenda, and motivated the plan to create a national research agenda that would capture the knowledge, expertise, and vision generated that day.

To launch the symposium deliberations, four cutting-edge research initiatives were conveyed to symposium participants. Chris Camacho with Children's Progress presented an adaptive and scaffolded assessment approach that provided prompts after incorrect responses and selected assessment items based on examinee responses to a previous question. Boo Murray, CAST, described an exemplar of universal design for learning, Strategic Reader, that assesses maze and oral reading fluency via a web-based tool. Jody Clarke, Harvard Graduate School of Education, described immersive virtual performance assessments under development and designed to assess knowledge and skills in science through items embedded within the context of virtual scenarios. Michael Russell of Nimble Tools demonstrated computer administered assessment tasks with embedded tools universally designed to facilitate access to content for students with special needs.

The presentations demonstrating assessment and technology innovations were followed by a large-group dialogue and discussion between the presenters and participants about the research initiatives including unique challenges and particular innovations. This discussion delved into the target areas of cognition and access and surfaced insights and questions arising from consideration of the future of TEAs. The debriefing session resulted in the large group dividing into two subgroups, one tackling issues regarding cognition and the other issues regarding access to assessment content. This article is based on the culmination of the symposium day plus the ongoing interactions among the subgroup members, who met to generate a research agenda regarding technology-enabled educational assessment and measurement of cognition for students with disabilities.

Seventeen participants joined the symposium subgroup that addressed measurement of cognition and students with disabilities. The members of the cognition group included one researcher from a university, six from assessment publishers, seven from research institutes, one researcher from a national technology center, and two educational consultants. The subgroup members communicated via email and telephone conference calls over a nine-month period following the symposium. Fifteen of the participants were contributing authors, writing components of this article. Two members of the symposium planning team facilitated communication, assembled interim drafts, and assimilated revisions from subsequent reviews. Final edits were assembled by the facilitators and reviewed by contributing authors prior to submission for publications.

Introduction

States are rapidly incorporating technology in state achievement assessment programs. Use of technology to administer, score, and deliver assessment results, or technology-enabled assessment (TEA), not only provides the promise of greater efficiency, but more important, of powerful new capabilities to create assessments that better model good instruction and support more valid inferences about student proficiency (Tucker, 2009; Quellmalz & Pellegrino, 2009).

Moreover, the forthcoming reauthorization of the Elementary and Secondary Education Act (ESEA), the emergence of common core standards, and additional federal funding to coordinate assessment efforts through state consortia provide a unique window for action (U.S. Department of Education, 2009). Increasingly, educators and policymakers demand more diagnostic information about students from various assessment sources about progress toward learning targets, data to inform targeted and successful instruction, and data to identify areas of focus for teacher and administrator professional development. The opportunity exists to use TEA in ways that were not possible with paper and pencil tests, to develop assessments that are more inclusive and also to create assessments that measure achievement with improved validity for all students, especially students with disabilities.

As the nation embarks on an unprecedented redesign of its standards and assessments, at a time when the field of TEA is new, it is essential that students with disabilities be considered in research and development efforts at the outset of any endeavor. Otherwise it will become necessary to retrofit the new assessments for this population of students (Thompson, Thurlow, Quenemoen, & Lehr, 2002; Thurlow, 2009), which is an expensive, time-consuming, and less effective proposition. As researchers begin to explore the opportunities that TEA offers to improve measurement of achievement and access of all students, we call for a comprehensive national research agenda that focuses on students with disabilities as part of the larger assessment system. Important issues, questions, and problems need to be articulated to promote research that addresses how to ensure that technology-enabled assessments are appropriately accessible and are employed to increase the validity of the measurement of the knowledge, skills, and abilities (KSAs) of students with disabilities.

A comprehensive assessment system that is fully aligned with challenging content standards is necessary for the acquisition of useful diagnostic information that supports all students in achieving the standards. In this "ideal" comprehensive assessment system, a system of assessments (Table 1) would be designed to validly and reliably measure all of the content standards, not just those that can be most easily and efficiently measured (Herman, 2007). To accomplish this goal, assessments would incorporate multiple measures to reach the highest levels of cognitive complexity prescribed by the standards. Interim assessments would document student progress at key points in the school year to allow for instructional corrections before the final summative assessment is administered (Perie, Marion, Gong, & Wurtzel, 2007). On-going formative assessment that is well aligned to learning progressions lead to achievement of the culminating standards. Learning progressions may be different for students with disabilities than for typical learners. Formative assessments would be continuously administered throughout instruction to make possible immediate instructional adjustments (Kingston & Nash, 2009). These assessments would be dynamic enough to chart all students' learning, even if unique, and to drive instructional efforts efficiently toward students' achievement of the content standards. Each component in a comprehensive system would contribute to constant monitoring of student progress and continuous adjustments in the instructional delivery system (Sheinker & Redfield, 2001). All of the assessments in the system contribute to ongoing diagnostic information.

Table 1: Components of an Assessment System

Although there is no consensus on how to define and interrelate the components of a comprehensive assessment system, for purposes of the discussion that follows, the following definitions will be used:

Formative Assessment	As defined in the RFP (Race to the Top Fund, 2009), "formative assessment means assessment questions, tools, and processes that are embedded in instruction and are used by teachers and students to provide timely feedback for purposes of adjusting instruction to improve learning."
Interim Assessment	As defined in the RFP (Race to the Top Fund, 2009), "interim assessment means an assessment that is given at regular and specific intervals throughout the school year, is designed to evaluate student's knowledge and skills relative to a specific set of academic standards, and produces results that can be aggregated (e.g., by course, grade level, school, or LEA) in order to inform teachers and administrators at the student, classroom, school, and LEA levels."
Summative Assessment	As defined by Perie, Marion, and Gong (2009), "summative assessments are given at the end of instruction to provide information on what was learned. They are generally administered once a semester or year to measure students' performance against district or state content standards. Summative assessments are standardized, usually given statewide (but can be given districtwide) and are often part of an accountability system. While schools may use these data to identify students in need of extra support, they are not designed to provide teachers with timely information about their current students' learning."

The actualization of a comprehensive system rests on a shift from current assessment practices. "Assessment practices should focus on making students' thinking visible to themselves and others by drawing out their current understanding so that instructional strategies can be selected to support an appropriate course for future learning," (National Research Council, 2001, pp. 90–91). Apart from typical accountability paradigms (and constraints), assessments purposed to provide instructionally relevant information should do more than target evaluations of current understanding, knowledge, or traits. They should provide insight into specific reasoning strategies students use for problem solving and constructing meaning (cognitive pathways), as well as provide timely feedback on the students' acquisition of knowledge and skills (learning progressions). Such assessment "...generally requires more complex tasks that reveal information about thinking patterns, reasoning strategies, and growth in understanding over time" (NRC, 2001, pp. 62–63).

This article has four sections. The first section reviews the potential and current uses of TEA to improve measurement of cognition of students with disabilities. The second section proposes considerations for designing a research agenda to address key issues that are especially pertinent to students with disabilities. The third section poses questions for research related to the essential issues for students with disabilities, and the final section provides suggestions for designing and implementing the research agenda.

Cognitive Pathways and Learning Progressions

For the purposes of this paper, we will explore how cognition may impact student interactions with assessments and propose possible solutions for students with pathways that may differ from standard assumptions about student thinking based on the impact of their disabilities. This paper proposes new thinking, but does so carefully because the notion of cognitive pathways is still relatively new. Currently, there does not yet exist widespread agreement regarding two key terms used throughout this paper: "cognitive pathways" and "learning progressions." However, Pellegrino (2009, pp. 98–99) draws the following distinction.

> From a cognitive standpoint, development and learning are not the same thing. Some types of knowledge are universally acquired in the course of typical development, while other types are learned only with the intervention of deliberate teaching.
>
> In math, the concepts of ordinality and cardinality appear to develop in all non-infants without instruction. In contrast, however, such concepts as mathematical notation, algebra, and Cartesian graphing representations must be taught.

So cognitive pathways apply to the growth or maturation of physio–logical/neurological structures and lead to skills that have not typically been the focus of classroom teaching or assessment. These structures impact most kinds of learning and include working memory, metacognition, and fluid intelligence. Cognitive pathways, or developmental thinking processes, provide the platform for learning progressions, where learners build upon their existing declarative and procedural knowledge.

Typically, a child must understand two-digit subtraction without borrowing before he/she learns two-digit subtraction with borrowing. Both are likely to be easier to learn if the child also understands place values. Together these three concepts form a small piece of a typical learning progression as we now understand it, which might include required and optional pathways. How (and how well) a child integrates these concepts will depend on the underlying cognitive structures and thinking skills available to that child. For example, a student with a smaller working memory capacity might require a different teaching approach and perhaps a different assessment approach to determine whether the student has mastered the concept rather than whether he has remembered the procedure.

Learning progressions—which are the content skill and concept building blocks to learning—may be only slightly less idiosyncratic than cognitive pathways. Heritage (2008) cites six descriptions of learning progressions from recent research literature. All have in common a vertical progression within a domain, but differ in the implied size and breadth of each unit within the progression from Popham's (2007) "carefully sequenced set of building blocks that students must master on route to a more distant curricular aim" to several authors who describe learning progressions as "descriptions of successively more sophisticated ways of thinking about an idea that follow one another as students learn" (Wilson & Berenthal, 2005; Smith, Wiser, Anderson, & Krajcik 2006). In contrast to Popham's definition, this latter description might encompass cognitive pathways as it incorporates thinking or cognitive complexity. Cognitive pathways and learning progressions may be intertwined. Cognitive pathways allow learning to occur, and, as learning occurs, the cognitive pathways are elaborated, thereby increasing capacity for traveling the learning progression.

A critical question that needs to be investigated is, "Do learning progressions differ for students with disabilities?" The presence of a disability may create the need for different cognitive pathways and varying learning progressions to content comprehension. For example, text comprehension is often considered a necessary access skill for success in school and on assessments. However, students who are deaf do not grow up in a phonemic environment, so the cognitive pathways that provide the foundational architecture to support print literacy may differ from those of hearing students (Singleton, Supalla, Litchfield, & Schley, 1998). At the same time, students with learning disabilities (Edyburn, 2000) and students with visual impairments (Johnstone, Thurlow, Altman, Timmons, & Kato, 2009) may learn to read print (or Braille) in similar ways to non-disabled students, but may choose to bypass visual reading and listen to text for comprehension in school. At the very least, students with disabilities for whom print creates a particular challenge may have different font, color, size, and color-overlay preferences that are impossible to reproduce on a paper-based test. In order to design and implement assessments that are accessible, adequate, and valid for use with students of varying abilities, it is important to understand the extent to which the cognitive pathways to learning vary across disabilities. This question is complex and requires attention to numerous considerations.

Potential of TEA

The use of technology within the comprehensive assessment system offers opportunities to investigate the multiple pathways to learning and demonstrating knowledge that are unique for students with disabilities. In addition, experts recommend that assessments provide students with multiple representations of material and multiple means of response (Dolan, Hall, Banerjee, Chun, & Strangman, 2005). Such flexibility is difficult with paper-and-pencil assessments and requires accommodations that are often inefficient, or expensive, or that may threaten the validity of the assessment. For example, having an adult paired with a student to read materials or capture responses may be problematic for a variety of reasons, such as introducing the possibility of changing the intended construct (Sirici & Pitoniak, 2007). Moreover, such approaches foster a dependency on others that is inconsistent with the educational goals for students. Technology-enabled assessments hold forth the promise of increased access (Almond et al., 2010) and supports that are efficient and effective and that maximize the capabilities of students with disabilities to demonstrate what they know and can do. Because students with disabilities are a heterogeneous group, the "alternate pathways to learning and assessment" argument laid out by the Center for Applied Special Technology (CAST) and provided in many online assessments may be a more appropriate way for assessing all students, especially this population.

While still in early stages, there are several projects that exemplify work currently occurring to harness the potential of technology-enabled assessments to better measure achievement for all students. The implications of these advances for a program of research pertinent to students with disabilities are described in this article. The projects described below attempt to reveal cognitive pathways and leverage learning progressions in new ways and at various levels to increase student access to the assessment and usefulness of findings from the assessment for mapping ways to increase learning by informing teaching.

Oregon's Computer-adaptive Assessment (CAT)

CATs have been used for about 30 years, but not for summative assessments for accountability until recently accepted for use in Oregon. Oregon State Department of Education developed a within grade level summative adaptive assessment (http://www.ode.state.or.us/testing/manuals/2007/doc4.2adaptivealgorithm.pdf). CATs save testing time by giving each student only those items likely to yield useful information. Typically, adaptations are determined based on the average difficulty of items across the entire population of test takers. For example, very easy items are not presented to students for whom previous responses indicate the items will surely be responded to correctly. The promise of adaptive testing to measure achievement as currently designed, however, depends on two underlying assumptions: (1) all students are taught in the same scope and sequence and, (2) learning progressions within each content area are largely common across students (Ash, 2008, p. 21). Students must have maximum opportunity to demonstrate knowledge at grade level before assessments are "adapted" and focus shifts to different or less complex constructs. Without these assumptions being met, the ordering of items according to difficulty will be different for different groups of students, particularly students with atypical learning styles or abilities. Results may be misleading when test takers respond to questions in unexpected or idiosyncratic ways (Cahalan-Laitusis, 2009). Given these caveats, however, the potential exists to explore students' learning in a manner different from paper and pencil tests, which typically concentrate items at the proficient/not proficient cut points.

Children's Progress Academic Assessment
(http://www.childrensprogress.com/ products/benefits-of-cpaa.shtml.)

This assessment incorporates adaptive functions with the addition of scaffolding (hints or prompts that increase student access to the construct) in a formative assessment, to provide children with feedback and support. The program covers concepts in early literacy and mathematics. The theoretical foundation of the design is to identify the Zone of Proximal Development, which is defined by Vygotsky as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978). One of the principles behind the assessment is that as much information can be gained from a child's incorrect responses as from correct responses. Adaptivity rules are based on classical psychometrics as well as the conceptual connections between items as described by content experts; learning progressions are considered along with item difficulty indicators. Scaffolding, through hints and prompts in this assessment, is based on distracter analysis that identifies students' common misunderstandings. The goal is to directly tie assessment with instruction and provide instructional recommendations for teachers based on student results.

Harvard University Immersive Virtual Performance Assessment

Another example of the innovations available through technology-enabled assessment is the development of virtual performance assessments by Harvard University (http://virtualassessment.org/). The Harvard project aims to summatively assess achievement in scientific inquiry based on research in Evidence Centered Design proposed by Mislevy, Steinberg, and Almond (2003). The student experiences an immersive virtual performance assessment via computer interface, with 3-D virtual contexts, digital artifacts, and avatar-based identities that capture what students are doing. It has the look and feel of a video game, but is based on an authentic setting allowing a realistic causal model for experimentation. Everything the student does is captured in the database, and this capture allows real-time analyses of the student's path toward solving the problem, comparing student pathways to pathways that would be used by experts in this situation.

CAST Strategic Reader

There are also important efforts underway to develop computer-enabled assessments specifically for classroom applications for students with disabilities. For example, CAST has developed Strategic Reader, a web-based research prototype reading environment (http://cast.org/research/projects/pm.html) that includes curriculum based measurement tool, which can be used for Response to Intervention (RtI) assessments. Two curriculum based measures have been embedded—maze and oral reading fluency. The maze test is immediately scored and results reported in various formats. The oral reading fluency measure is recorded audibly, so teachers can listen to the audio and score later. The computer does all calculation of the scores and produces various displays that allows teachers to view student progress and determine when and which interventions are needed.

Current Status of TEA

Evidence of state reform efforts to incorporate technology into standards and assessments has been demonstrated in a recent National Center for Educational Statistics (2009) report. By 2008–09, all 50 states had standards for students that included technology, 29 states had established a virtual school, 26 states offered online large-scale assessments, 13 states tested students on technology standards, and 3 states required that teachers and administrators be trained in the use of technology. While more than half of the states now employ large-scale online assessments, the use of technology for assessments is proceeding more slowly than for instruction and has not resulted in changes to traditional approaches to testing. Most of the online tests used by state programs are replications of these traditional paper and pencil versions consisting primarily of multiple choice questions. Only a few include essays, and even fewer use computer-adaptive testing, which has been used for years for certification and admissions testing.

Reasons for using technology are multiple: (1) technology is at the core of virtually every aspect of our daily lives and work, and we must leverage it to provide engaging and powerful assessments that measure student achievement in more complete, authentic, and meaningful ways; (2) technology-based assessment systems will be pivotal in improving student learning and generating data that can be used to continuously improve the education system at all levels (Office of Educational Technology, 2010, p. 3); (3) many students come to school as "digital natives" who have spent their entire lives using the toys and tools of the digital age, and because of this exposure, they are starting to think and process information differently (Prensky, 2001); and (4) technology provides the potential to transform the daily educational experience much as it is transforming other parts of society. Technologies gaining use in classrooms include not only desktop and laptop computers, but digital whiteboards, cell phones and smart phones, handheld or mini computers, clickers, Alphasmart keyboards, LCD projectors, USB drives, portable media players, and digital cameras (Simba Information, 2009), although schools continue to face struggles with the barriers of cost and access to hardware, software, and infrastructure.

Considerations for Designing a Research Agenda

In defining the key issues and suggesting research designs, the authors considered several priorities for building an evidence base for technology-enabled assessments, specifically tailored for students with disabilities:

- How can we build a theoretical and empirical base for a new generation of assessments that allow us to measure and foster skills that cannot currently be measured well for all students?
- What would a comprehensive assessment system that builds on the opportunities that TEAs provide look like? And how might the assessment tools within the system using TEAs provide more valid, complete, and actionable information for teaching and learning than our current assessments, which typically indicate what students with disabilities do not know or can not do?
- How do we design the research to gather the evidence needed to help policymakers answer questions about efficacy, efficiency, validity, and value-added characteristics of technology-enabled assessments for students with disabilities? And how do we do this and implement it within a timeframe that is useful to practitioners and serves the pressures faced by policymakers?

Key Issues and Assumptions for Framing a Research Agenda

Better Understanding of Cognition

Technology may provide an opportunity to increase our depth of understanding about cognition and the variance that likely exists in how students with a diversity of disabilities progress toward learning outcomes. In building on what is currently known about how students achieve mastery in reading, mathematics, and science, we can anticipate that investigations of how uniform or varied the ways in which students progress from novice to expert may be increasingly possible using technology-enabled assessments. For example, can students with disabilities that affect language or reading fluency approach mathematics problems through graphically based problem-solving strategies to reach the same understanding as typical learners? Or can investigations clarify whether, as Mislevy contends, "Rather than seeing such variables as independently existing characteristics of people, we can view them as summaries of patterns in situated behaviors" (Mislevy, 2009, p.1) that make for similarities in what students do on assessments. What can we learn about the strategies that students use to understand text and draw inferences across passages, for example, when working memory or executive functioning skills may be an issue? Technology-enabled assessment may provide an increased opportunity to capture the complexity of competent performance necessary to design assessments from which we can make more valid inferences about the achievement of all students (Pellegrino, 2004; Tucker, 2009).

In order to fulfill the potential of TEA to provide greater accuracy in determining what students, especially students with disabilities, know and have mastered, the authors focused on the following issues in their discussions:

- Information derived from TEA is meaningful to the instructional decisions teachers need to make.
- Teachers use the information gained from these assessments to guide instruction.
- These assessments contribute to improvements in achievement, especially the higher levels of cognitive complexity that students may achieve.
- Students, especially students with disabilities, have access to all assessments in the system and be fully engaged by the items and tasks.

A dilemma for measuring achievement of students with disabilities is the uncertainty about potential barriers that the assessments' structure and format might present to an accurate determination of what the student knows and can do. Technology offers flexible tools and formats to increase access. For example, font size and color can be individually adjusted for optimum visual access. The possibilities have not been fully investigated and may not yet be known as the full potential of technology unfolds. (See the article in this issue, Technology-Enabled and Universally Designed Assessment: Considering Access in Measuring the Achievement of Students with Disabilities—A Foundation for Research [Almond et al., 2010], for a more in-depth discussion of this topic and a future research agenda focused on access for students with disabilities.)

More Rigorous Standards

States have struggled to measure more rigorous standards using current assessment methodologies. States and national organizations debate the importance of assessing skills beyond basic academics that meet international benchmarks (NGA, CCSSO, & Achieve, 2008; Jerald, 2009). The Council of Chief State School Officers (CCSSO) recognized that discrepancies existed in the degree of challenge evident in state assessments resulting from differences among state standards. CCSSO and the National Governors Association Center for Best Practices (NGA Center) partnered with Achieve, Inc., ACT, and the College Board to initiate a process to develop a common core of state standards. This common core is intended to prepare students to compete nationally and internationally. It is still unclear what impact these standards will have on state assessments, and it is likely to be some time before the outcome is apparent. At present, states may not be currently assessing the most challenging standards that already exist in their policy documents. For example, reading standards in many states address skills like critiquing and interacting with various forms of text (Johnstone, Thurlow, Thompson, & Clapper, 2008), but may not assess these areas on statewide tests (Johnstone et al., 2007; Johnstone & Thurlow, in press). Efficiency and costs may limit opportunities to accomplish this goal. TEAs may open doors for the assessment of already existing high standards or the adoption of more rigorous standards, perhaps even common core standards, where the limitations of traditional assessments previously excluded them (Tucker, 2009). The innovative approaches to assessment that TEAs make possible go beyond the limited representations of learning shown on traditional assessments and provide a window into cognition that traditional assessments cannot. This innovation is especially critical for students whose disabilities have prohibited them from demonstrating what they know and can do on traditional paper and pencil tests.

Improved Validity

The utility and expanded use of TEA hinges on the establishment of evidence of validity. Establishing the validity of these assessments is necessary for their use in summative assessments for accountability and establishing a relationship between formative assessments and summative outcomes. Technology's flexibility, range, and ability to collect and combine multiple sources of data offer the promise of bridging the often-perceived distance between learning and assessment. Evidence of validity will be necessary to defend TEAs against arguments for comparability to the status quo.

An assumption here is that comparability is not necessarily evidence of validity. Innovations in assessment that increase accuracy of results and open windows into the varying cognitive pathways that individual students exhibit should not be held hostage to "comparability" with traditional forms of assessment. Technology-enabled assessments provide the opportunity to gather more information about the cognitive pathways and learning progressions individual students follow in reaching full achievement of the standards. Rather than achieving comparability between two significantly different forms of assessment, the emphasis for the technical quality of any assessment should be increased validity for the intended purpose (AERA, APA & NCME, 1999).

The No Child Left Behind Act Peer Review Guidance calls for evidence of comparability (NCLB, 2004). Comparability between TEAs and paper and pencil assessments and the results they produce may cease to be an issue if:

- The TEA provides increased access for students with disabilities.
- The increased depth and range measured by the assessment becomes the primary source of any incomparability.
- Evidence of increased validity of technology-enabled assessment becomes the source of incomparability because the TEAs provide more accurate evidence of what students, especially students with disabilities, know and can do compared to paper and pencil tests.
- The source of any incomparability of results across years in the transition from paper form to technology platform is determined to be the result of improvement of the inferences that can be derived from TEA.

Whether or not the upcoming reauthorization of ESEA redefines assessment or accountability parameters, best practice in assessment requires pursuit of improvement in the validity of assessment results, the inferences drawn from it, and the usefulness of that information for improving student learning.

Evidence of Value

Finally, only when adequate infrastructure is present in schools can TEAs be fully implemented. Motivating policymakers to provide the needed funding for adequate infrastructure requires convincing evidence of the efficacy, efficiency, validity, and value-added possibilities that technology-enabled assessment offers. Research on TEAs thus far has demonstrated their potential to assess students in ways that provide a deeper understanding of student learning and that provide information that cannot otherwise be acquired in other formats (Clarke-Midura, Dede, & Mayrath, 2010). More evidence is needed to demonstrate and document these attributes, including research that substantiates the capacity of TEAs to improve teaching and learning of students with and without disabilities by providing valid, accessible, and practical assessments of cognitively challenging content and skills.

One example of how TEAs can make possible an assessment that can provide a deeper understanding of student learning, particularly of more complex content applications, is the immersive virtual assessments currently being developed and researched at the Harvard Graduate School of Education. This research project promises to make possible the assessment of science inquiry skills, an area of interest to content teachers and assessment designers that has proven difficult if not impossible to measure on conventional assessments (Clarke-Midura et al., 2010). This type of assessment may enhance the ability of students with disabilities to demonstrate more complex learning when more visual information that relies less on text and more on interactivity, such as "drag and drop" and audio instruction, is available so that communication of the task is not totally or mostly dependent on or confounded by language. Harvard's Graduate School of Education's immersive virtual performance assessments provide strong visual images made uniquely possible via TEA to measure students' science inquiry skills. Figure 1 displays an image from River City, an immersive virtual assessment of middle school science that takes students through a visually rich virtual simulation of environments. River City presents students with a problem and asks them to develop a hypothesis and procedure, virtually test their hypothesis, describe their findings, and make recommendations.

Figure 1: Harvard Graduate School of Education's immersive virtual performance assessments provide strong visual images made uniquely possible via TEA to measure students' science inquiry

Still, there are a number of important research questions to think about as we discuss this topic focusing on students with disabilities. The research questions focus investigations around the following key questions: (1) How do cognitive pathways and learning progressions vary across disabilities, and how can technology-administered tests take advantage of this knowledge? (2) How does the use of technology impact the ability to measure the constructs of interest, especially for students with disabilities? (3) What is the impact of using technology to administer tests, including performance assessments, on item creation, item type, and analysis of assessment results? (4) Which data matter for instruction and for score interpretation? and (5) How do we model the data to be useful for various purposes?

Designing Measures of Learning Progressions

Simply stated, learning and development are not the same thing. Cognitive pathways are physiological structures. Learning progressions are building blocks that students master on route to a learning outcome. Evidence-centered design (ECD) offers an approach to assessment design that incorporates attention to both cognitive pathways and learning progressions. This assessment design framework is used to increase construct validity by executing a rigorous procedure consisting of five layers: domain analysis, domain modeling, conceptual assessment framework, and compilation, and the fifth layer, which is a four-phase delivery architecture (Almond, Steinberg, & Mislevy, 2002; Mislevy et al., 2003; Mislevy & Haertel, 2006). The result is an assessment argument that is able to link theories of cognition to student performances as evidence. In using ECD to design an assessment of a learning progression, the assessment designers conduct a domain analysis to identify the cognitive constructs that are of interest in the domain. During the domain modeling phase, the knowledge, skills, and abilities (KSAs) to be assessed are further specified to reflect a particular learning progression of interest. The ECD framework requires designers to create design patterns and task templates. These documents link the knowledge and skills being assessed to the kinds of evidence that will reveal that students have attained the knowledge and skills associated with the learning progression and may hold implications about their cognitive pathways.

Design patterns created during the domain modeling phase and task templates can be used for the design of multiple items/tasks. As part of the framework, designers identify the features of items/tasks that can be varied to detect different levels of performance of the learning progression or pathway. For example, a variable feature may be the type of performance a student is expected to provide (an explanation vs. recalling a fact), the type of data display used (histogram vs. representation such as scatterplot), or the number of data points in a graph. TEA offers the opportunity to customize items/tasks at delivery, when specific items/tasks are chosen for given students. While it is important to keep the end in mind, for example, achievement targets expressed in achievement-level descriptors, designers have found that TEAs can provide a wealth of information about the varying cognitive pathways and learning progressions students follow toward these achievement targets. Working backward from the ultimate goal as described by the achievement-level descriptors, designers can map significant points along the content-specific learning progression and escalations in cognitive complexity leading to the goal.

Other assessment designers have discussed at length the need to attend to the cognitive load of assessments by creating collections of items that reach the highest levels of cognitive complexity, but also map the cognitive pathway (Burkhardt, 2009; Schoenfeld, 2009). TEA provides an opportunity to increase understanding about what students with disabilities know and can do and how they know it. By deconstructing the cognitive features of content KSAs with attention to the cognitive complexity of what is being required, designers of TEA can increase the accuracy with which the nature of mastery for individual students is measured relative to identified learning outcomes. This information is useful for skilled instruction as well as accurate measurement of degree of mastery of content concepts.

Assessments that take cognition into account "… rely on detailed models of the goals and processes involved in mental performance … (NRC, 2001, p. 63)." The information about cognitive pathways needed for such design efforts is only now being acquired. Assessment designs that allow researchers to investigate the cognitive pathways and learning progressions of individuals and groups of students have been largely unavailable. Models like the Children's Progress Academic Assessment and the ECD framework provide some clues to how TEA might open the door to the design of assessments that reveal these unique aspects of cognition, learning, and performance.

Measuring the Construct of Interest Using TEA

Scientific knowledge about computer-based tools, sound assessment design, student access to demonstrating performance, and cognition should be the foundation for building technology-enabled assessments. Building upon such a foundation is likely to enhance the validity or trust in and utility of the assessment results. Haladyna and Downing (2004) note that "the most fundamental step in validation is defining the construct" (p. 25). Evidence of the application of sound assessment design per accepted design principles, such as those described by Downing and Haladyna (2006), and evidence-centered design (ECD) (Mislevy et al., 2003; Mislevy & Haertel, 2006) should be considered in both the design of and research on TEAs, particularly in regards to investigating validity evidence to demonstrate that the assessment is measuring the construct of interest (Clarke-Midura et al., 2010).

Bejar et al. (2003) noted in their discussion of Embretson (1983) that "construct representation [is] a key aspect of test validity concerned with understanding the cognitive mechanisms related to the item solution and item features that call on these mechanisms" (p. 4). This concept is true for all types of assessments. In the context of TEAs and their environments, clear articulation of the measurement construct is key to understanding the role of technology in the measure, the potential benefits of technology, and its potential pitfalls.

The interfaces, tools, modes of presentation of the content, and modes of response are item and assessment features with the potential to impact student performance. If performance is impacted because the feature is relevant to the construct being measured (construct-relevant), then it is functioning appropriately; otherwise, the feature is construct-irrelevant and is evidence against validity of the assessment. For example, if an assessment requires a student to read and understand an unfamiliar vocabulary word before he/she can demonstrate understanding of the attributes of a main character in a narrative passage, the word itself may be irrelevant to and an obstacle to the student's ability to explain what he/she learned about the character from the entire passage.

Questions researchers hope to answer are whether cognitive load is increased, and if so, by how much, and is this additional cognitive load a feature of the construct being measured or external to the construct (Paas, Renkl, & Sweller, 2006). For example, in the work of Clarke-Midura et al. (2010), does the use of an avatar introduce a construct-irrelevant feature that impacts the validity of results or the inferences about construct knowledge that can be made? The impact of features like this that are possible in TEA is not as well researched or understood as the impact of features of traditional paper-based assessments. It is, therefore, important that future research investigates the impacts that technology environments have on student performance to understand what features relate to or that might interfere with the validity of the assessment, especially those which may be affected by a student's disability, and, just as important, what can be done to minimize construct irrelevant impacts.

However, for all of these implementations of technology in assessment, the benefits of each "enhancement" must be weighed against the potential negative impacts. The test item creation process requires continual investigation of how students think about and respond to specific test item features (Leighton & Gokiert, 2005). Although extensive research has been performed to examine the statistical properties of test items, less effort has been directed at understanding the potential for item features to introduce or reduce construct-irrelevant variance (Ferrara et al., 2003; Haladyna & Downing, 2004; Leighton & Gokiert, 2005; Gorin, 2007).

Therefore, cognitive models for how students will interact with the assessment items, given their unique cognitive pathways, and the technological features of TEAs must be posited and investigated. These cognitive models should not only be statistical in nature. They should include both the problem-solving strategies that students use within specific academic content domains and those they use for navigating and using the technology. Thorough investigation of these models across student populations will help to identify irrelevant and confounding elements within the assessment, provide greater validity through better understanding of the knowledge and processes students use to respond to assessment tasks, provide clarity about the needs for data capture during live assessment, identify potential mechanisms for using and understanding complex data, and identify how students can best demonstrate their learning.

Understanding access and cognition in light of the measured domain-specific constructs within a TEA should provide improvements in (1) the assessment's ability to measure what students actually know; (2) the assessment's ability to measure how students use the knowledge, understanding and skills in new and different contexts; (3) students' abilities to navigate and respond within the assessment environment—a critical element for students with physical and cognitive disabilities; and (4) instructionally useful information about student thinking, areas needing improvement and extension, and strategies to close learning gaps, correct misunderstandings or enhance learning progress by mapping next steps to increased complexity.

Further, the current psychometrics and test designs consistently yield relatively low levels of precision or high levels of measurement error for students at the "extremes" of performance (e.g., lowest and highest levels). The lack of precision at the lowest performance levels particularly affects students with disabilities, and yields low validity, trust, or utility of such scores. When students either top out or bottom out, the assessment is likely not efficiently measuring students' actual knowledge or providing useful information for extending that knowledge, even for top performers. Computer-based assessment tools could adapt to, and extend, such performance opportunities at the extremes, thereby adding variability in scores, and increasing reliability and validity. Such computer-based assessment tools might include the use of adaptive algorithms for item selection (CAT), incrementally providing hints or supports (scaffolding), and continuous feedback, simulated environments, and other approaches not yet imagined.

Measuring Incomplete Understandings Using TEA

Assessments, if they are to be instructionally supportive, should be built in ways that reveal the cognitive strategies students utilize in accessing and responding to test items, including both correct and incorrect responses, so that both remediation and enrichment instruction can be appropriately and successfully applied. As Pellegrino and his collaborators note, there is "value [in] describing students' incomplete understandings" (NRC, 2001, p. 84). As noted before, TEAs make possible the opportunity to gather a wealth of information about these cognitive strategies. When technology is employed, opportunities increase to make assessments more interactive and dynamic than they might be otherwise, so design of the assessment becomes critical. In this section, we discuss interactive, dynamic, and scaffolded assessment. The following terms are used to discuss these topics in this section.

The terms "interactive" and "dynamic" are used together and separately in the literature. The definition of interactive-dynamic assessment comes from the field of intelligence testing and the dissatisfaction with conventional methods to provide information about individual's learning that can be translated directly into practice by educators and members of other helping professions. An important characteristic of interactive-dynamic assessment is identification of specific obstacles that may be hindering cognitive performance and specification of conditions under which intellectual performance can be facilitated (Tzuriel & Haywood, 1992, pp. 8–9). For the purpose of this article, we discuss several types of interactive assessments and assessment strategies made possible through technology. Interactive assessments can provide the opportunity for feedback to students and teachers during the assessment process.

Table 2: Interactive Assessment

Interactive assessment is a type of TEA that can impact the flow of instruction and is responsive to student performance (Byers, 2001).

Dynamic Assessment	Dynamic assessment is a type of interactive assessment based on Vygotsky's Zone of Proximal Development that integrates instruction and assessment (Poehner, 2008). Dynamic assessment provides corrective feedback in response to student failure to measure both the product and process of learning (Caffrey et al., 2008).
Scaffolding	Scaffolding is a dynamic assessment strategy used to enhance items derived from supports provided during learning that are gradually removed when learning becomes solidified and/or the learner becomes more independent. Scaffolding in assessment includes structural assistance introduced to organize information or guide responses embedded in the presentation of the item or task (Perie, 2009, p. 377). Scaffolding is expected to affect the knowledge, skills, and abilities (KSAs) required to respond to a task, but is not intended to change the construct being measured. More information may be obtained regarding student cognition than in static assessments when only partial understanding of the construct is evident.
Computerized Adaptive Testing (CAT)	Computerized adaptive testing (CAT) is a type of interactive assessment, which redesigns educational measuring instruments for delivery by interactive computers. Its objective is to select, for each examinee, the set of test questions from a precalibrated item bank that simultaneously most effectively and efficiently measures that person on the trait (Thompson & Weiss, 2009).

Dynamic Assessment

Dynamic assessments have been shown to have significant benefits for assessment and instruction (Grigorenko, 2009). A critical feature of a dynamic assessment approach is to link assessment with instruction effectively by providing actionable information to support instructional adjustments (Black & Wiliam, 1998). Although there are many different approaches that fall under the umbrella of dynamic assessment, there are a few key aspects that are consistent. Generally, dynamic assessments quantify a child's learning potential, whereas static assessments typically gauge a child's state of preexisting knowledge. The primary difference between the two approaches lies in the fact that dynamic assessments adopt a Vygotskian approach to assessment. That is, dynamic assessments typically provide various types of scaffolding after incorrect responses to dissociate what a child can do independently versus what a child can do when provided with scaffolding. Through the scaffolding procedures, dynamic assessments can shed light on particular misunderstandings that may be responsible for a child's incorrect response. On the other hand, because static assessment does not provide scaffolding after incorrect responses, it is able to reveal only two states of understanding, unaided success and unaided failure (Fuchs et al., 2008). For the purposes of accountability over the past several years, static assessments seemed useful. In these assessments, students either knew or did not know grade-level material. Such assessments were useful for broadly understanding student success on grade-level standards, but have been less useful for instructional decision making.

In terms of instructional utility, dynamic assessment has been named as a promising alternative to static assessments for students with special needs, because this type of assessment may be able to determine the adequacy of students' responses to interventions (Fuchs & Fuchs, 2008; Fuchs, Fuchs, McMaster, & Al-Otaiba, 2003). Specifically, research has shown that dynamic assessments may predict an individual's potential to learn better than static measures (Grigorenko & Sternberg, 1998; Sternberg & Grigorenko, 2002). Dynamic assessment is able to account for variations in performance that may be due to factors such as prior instruction or a misunderstanding of the task directions, which cannot be accounted for by static testing (Samuels, Schermer, & Reinking, 1992). Therefore, it might be possible to use dynamic assessment to better predict whether the student is likely to respond to instructional interventions or whether a more intensive intervention should be prescribed earlier on in the intervention process.

Scaffolding

A primary question for dynamic assessment is: How does the child respond to the feedback and scaffolding procedures? It is this reaction to feedback and scaffolding that dynamic assessments attempt to quantify. Scaffolded assessments logically include items to allow for gradually increasing the difficulty or cognitive complexity of the items/steps related to the concept to be measured and hopefully to lead to a truer measure of the students' level of performance or achievement. Scaffolded items may be developed to logically build meaning when complex concepts are embedded and contain first steps/questions that address underlying/background knowledge and skills, while subsequent steps might more closely approach the content standards intended for the grade level being assessed. For students who seem to be performing above or below the desired achievement level, scaffolded items may allow for assessing content above and below the complexity/difficulty level of the grade level content standard or within agreed upon "bands" of difficulty within a grade level.

Scaffolded items may also deconstruct the complexity of the construct being measured to allow students to show partial understanding of complex KSAs. A project currently in progress for a General Supervision Enhancement Grant for Montana is investigating the effects of scaffolding strategies in an alternate assessment based on modified achievement standards. Four scaffolding strategies have been developed for both reading and mathematics assessments, which are designed to utilize instructional prompts teachers typically employ to redirect students from common misconceptions to the targeted KSA (Bechard, 2010).

Through the use of scaffolded items, students can presumably better direct their own learning because they know what they know and what they do not yet understand, and teachers can likely better tailor instruction to meet student-learning needs. By providing feedback and scaffolding, dynamic assessment makes the evaluation a bidirectional process intended to mimic the child's daily learning environment.

Identifying Instructionally Useful Error Patterns Using TEA

Recent research, test development, and score reporting have paid little attention to the identification of errors or error patterns for examinees, either as a whole or for subgroups. Even for traditional (i.e., summative, paper-pencil, static) assessments, investigations of errors can readily be conducted to provide instructional information, as well as assessment design and item development improvements. As assessments become more technologically enabled, innovations in item types and test administrations, such as scaffolded items or adaptive assessments, could greatly improve the type and amount of instructionally relevant information available from the assessments. TEAs have the potential to provide information about what students do or do not know, but also what cognitive specific strategies students do or do not apply.

There will certainly be challenges in the identification of errors and, especially, patterns of errors, particularly if we are dependent upon existing assessment designs, psychometrics, and data management practices. The use of TEAs may prove especially beneficial in overcoming such challenges. Research should seek to evaluate the benefit of TEAs in the identification of errors and patterns of errors, thus increasing the dynamic nature and utility of such assessments.

Research should seek out and explicitly investigate TEAs that are instructionally useful. These TEAS should be designed to provide instructional implications that can be used to guide the learning activities of students at varying levels of knowledge and skills. The research should be grounded in sound learning theories and teachers' instructional experiences that identify students' errors as they engage in learning. The erroneous strategies can be identified and documented to provide information about where a student is performing along a learning progression, providing a valid foundation upon which sound instruction can be based.

Which Data Matter in TEA?

TEAs have the potential to acquire, store, process, and communicate a wide range of data that can be used to make sound inferences about what students know and can do, including complex forms of learning and cognition. How data are to be used has an important impact on how the data should be presented and the level of detail for data collection. For a summative assessment, the main inferences typically drawn might be those related to the status of the student's knowledge, skill, and ability at the end of a learning experience (e.g., a semester or year of study). In summative assessments, data about the learning progressions that students follow to learning outcomes may be more difficult to collect given the broad range of goals that are being assessed.

On the other hand, while taking a formative assessment, the main inferences typically drawn are those related to what the student is currently learning (or not) in the classroom. Formative assessments may provide more opportunities to collect data on learning progressions given the fewer learning outcomes that are typically assessed in any single formative assessment. TEAs may provide multiple tasks that teachers could administer formatively and frequently that embody effective instruction. These might include a series of probes that could be used to investigate further student understanding. Instructional response to findings from formative assessments may also follow in real time, more immediately than from summative assessments, especially those that are paper-based. If the assessments results are intended to inform instruction, it may be a matter of great importance what routes have been followed and what patterns of errors students have made, because improvements can be made to future instruction based on this information.

Since there may be greater variability in the pathways followed by students with disabilities to attain particular learning outcomes, the benefits conferred by TEAs may be greater for students with disabilities. Furthermore, it may be possible using TEAs to collect data that document students' patterns of reasoning and make more valid inferences about hard-to-measure concepts (e.g., complex problem solving skills in mathematics and inquiry processes in science). In these hard-to-assess areas, it is important to document the patterns of reasoning and errors that students make as they engage in multistep, iterative solution strategies. TEAs are able to capture this kind of information for the individual student. To use such data confidently, it is necessary to validate the inferences. Inferences that are important for students with disabilities may vary from or be larger than the set of inferences that are important for students without disabilities, because of, for example, the additional inferences that may need to be made about the accessibility of the assessment tasks.

For summative assessments, TEAs provide increased opportunities to gather data about student performance more closely aligned with the intended depth of knowledge of the assessment and beyond. TEAs offer the opportunity to gather data from student performance using the range of tools for designing innovative items and tasks utilizing simulations, enhanced graphic design, and flexibility in presenting the tasks in different ways to accommodate diverse learners. Some aspects of achievement that have proven resistant to all efforts to assess them, such as extended writing, mathematical reasoning, and scientific inquiry, may yield to the unique range of tools that TEAs make available. These advantages promise to add value to the known opportunities technology provides as described by Darling-Hammond and Ducommun (2010):

> Technology also organizes data about student learning, enhancing system accountability for instruction and reporting by providing more efficient, accurate, and timely information to teachers, parents, administrators, and policymakers. In the current U.S. context, technology can help to integrate information at all levels of the system as part of a longitudinal state data system, contributing to a rich profile of accomplishment for every student.

Not only will there be more data, TEA promises richer data that includes information about more complex items and tasks (Darling-Hammond & Ducommun, 2010). The opportunity to measure students' learning progressions and better understand their cognitive pathways is a significant addition to the data available from past assessments. The challenge will lie in our ability to analyze the data effectively, once gathered.

Research Questions Addressing TEA and Students with Disabilities

As discussed in this article the opportunities provided by TEA to enrich knowledge of student cognition, learning, and achievement and the impact of disability on learning are many. The challenge lies in a thorough investigation of the many dimensions of this still-emerging form of assessment. This discussion has attempted to briefly review the opportunities and challenges that TEA provides to help formulate research questions that can guide the design, deployment, and validation of TEA in the service of students with disabilities.

Research is needed to pool scientific knowledge across the fields of computer-based tools, assessment design, and cognition, as well as instructional strategies, for students with and without disabilities, particularly at the higher and lower ability levels, and to investigate the benefits and challenges of TEAs relative to traditional methods and modes of assessment. This article proposes that researchers and experts across disciplinary fields in these areas coordinate efforts to address these issues, guided by the following research questions regarding the use of TEAs, particularly for students with disabilities.

Measurement of Cognition

1. What is known about cognitive pathways and learning progressions that help in the assessment of students' content mastery?
   - How are cognitive pathways different or the same for students with disabilities?
   - How are learning progressions different or the same for students with disabilities?
   - What is the relationship between cognitive pathways and learning progressions for students with disabilities?
   - Are cognitive pathways and learning progressions different for students with different disabilities?
   - What are the learning progressions for students with different types of disabilities in various academic disciplines?
2. Can students' positions on the continuum of development and learning toward targeted outcomes be determined accurately?
   - What are identifiable benchmarks or milestones that can be assessed?

3. How can instruction based on evidence-based learning progressions be effectively used to better understand the cognitive pathways appropriate to a student's abilities and learning needs?

4. Can patterns of errors be identified to provide evidence for charting effective learning progressions? Can errors and patterns of errors be revealed by the following assessment types:

    - Those with a variety of designs and uses, such as paper-pencil, online, adaptive, dynamic, formative, summative?
    - Those at various levels of aggregation, such as total sample, subgroups, individual students?

5. If patterns of errors can be identified, what data can be validly extracted from those errors and error patterns to benefit the following areas?

    - Item and test design and development.
    - Performance assessment development.
    - Score reporting.
    - Cognitive modeling.
    - Production of instructionally relevant and useful data, instructional models, remediation, and enrichment.
    - Involvement of students in their own learning.
    - Design of professional development initiatives.

Interactive-Dynamic Assessment

1. How can TEA be used to build models of progressive learning and performance of students with disabilities?

    - How can the TEA map learning over time?
    - Does adapting and/or scaffolding of assessment items provide the instructional information teachers and students need to chart learning progressions? How useful is this information to teachers in guiding instruction?
    - In what ways could technology improve the diagnostic assessment process in a comprehensive assessment system? An intervention process?

2. How do students with disabilities respond to varying assessment conditions and environments?

- How do students respond to the feedback and scaffolding procedures inherent in dynamic assessment?
- What student strategies contribute to correct and incorrect responses within a TEA?
- What is the impact on performance and/or cognition of construct-irrelevant variables or barriers (i.e., access, navigability, accommodations)?

3. What effects do interactive-dynamic assessments have on the way in which teachers deliver curriculum and instruction to students with disabilities?

Validity Evidence

1. What data are needed to articulate the validity arguments associated with the intended inferences of technology-enabled formative and summative assessments, particularly for students with disabilities?
2. How do we expand and improve upon current psychometric methods that will be needed in a TEA environment (e.g., computer adaptive models, multidimensional IRT models, performance assessment measurement, statistical models for simulated environments, additional precision needed at the high and low ends of the scale)?
    - Can appropriate psychometric models be developed to facilitate diagnostic probing within the context of cognitive pathways and learning progressions for students with disabilities?
3. In what ways do strategies used to design items for static, paper-pencil assessments apply to TEA (e.g., accessibility guidelines, universal design)?
4. Can appropriate psychometric models be developed for TEAs linked to a learning progression and cognitive pathway for students with disabilities?
5. What are the constructs, both relevant and/or irrelevant, that are measured within the TEA environment?
6. What inferences can be made with acceptable validity on the basis of scores or results of formative and summative technology-enabled assessments for students with disabilities?
    - What frameworks and structures are available for framing these arguments (e.g., claims, evidence)?

- How do these frameworks represent key concepts—such as assessment purpose, the targeted proficiency to be measured or fostered, the task situation, characteristics of the students, and accessibility barriers—in a way that supports the design of assessments that achieve desired quality characteristics such as accessibility, validity, learning effectiveness, and learning efficiency (Hansen, Zapata-Rivera, & Feng, 2009)?
- How well do these frameworks address the requirements of data from complex constructed responses and data about learning? What enhancements to such frameworks are needed?

7. How can empirical research (e.g., analyses of test content, students' cognitive processes, criterion-related evidence) best support the claims (and counterclaims) for the intended uses of results in the validity argument that has been articulated?
8. How can Response to Intervention (RTI) techniques (Fuchs & Fuchs, 1998, 2006) be used effectively to enhance the validity of assessment results and subsequent instructional interventions?

Consequences of TEA for Students with Disabilities

1. What benefits can be identified for using TEA strategies with children with special needs? Are these benefits different for students in the general education curriculum? Benefits may include the following:
    - Measuring student performance on the defined construct of interest.
    - Measuring aspects of student cognition (i.e. cognitive load, problem solving).
    - Linking assessment with improved cognition, instruction, and learning.
        - Providing information about effective and less effective student cognitive strategies.
        - Providing information to support and improve instruction and student learning.
2. Are there negative consequences for students with disabilities using such types of testing?
    - While more precise understandings of points on learning progressions will help to guide instruction, is there potential

for individualized approaches to undermine targeting grade-level standards as outlined by IDEA (2004)?

Designing and Implementing the Research Agenda

Given the research questions described, designing and implementing the research agenda presents a number of challenges. Research questions need to be prioritized. Design of the research will require attention to both the unique opportunities provided by TEA and to the technical quality requirements of any assessment.

Challenges

Known challenges should be minimized while unknown challenges encountered during the research should be exposed. Measuring students' cognitive strategies in any assessment is a challenge. Some of the challenges to assessing various types of students with TEAs include student access, school infrastructure and technical capabilities, curriculum, teacher preparation, professional development, policy, and others. To respond to these challenges, researchers need to:

- Determine the degree to which TEA enhances measurement of students' cognitive strategies.
- Clarify the impact of cognitive pathways on learning progressions and their role in the design of TEAs for students with disabilities.
- Identify and apply what is already known about the use of technology in education. When technology is involved, the challenge increases due to uneven use of the tools of technology in classroom instruction.
- Identify generalizable findings. The wide range of disability types/categories, both between disability categories and within disability categories, will make it difficult to generalize findings especially with regard to low-incidence disabilities. The lack of consistency across schools and districts in the categorizing of students exacerbates this challenge.
- Conduct rigorous research based on random assignment and control groups. This may prove difficult given the many in-place programs and regulatory requirements.
- Identify incentives. While there is a clear need for systematic, coordinated, collaborative, and innovative approaches to addressing this set of research questions, incentives for engaging in such approaches are lacking.

- Find or develop the political will to invest the necessary time and funding to support systematic, coordinated, collaborative, and innovative approaches to yield effective implementations that can be brought to scale.
- Develop and implement training at all levels (e.g., implementers, developers, practitioners, and those who train them).
- Revise practices, policies, and regulations to ensure privacy and information protection while enabling a model of assessment that includes ongoing student learning data gathering and sharing for continuous improvement (Office of Educational Technology, 2010).

Timeframe

The research should include short- and long-term goals and involve both immediate and longitudinal analyses. The research should include a combination of observational qualitative research with the use of an existing TEA, followed by data analyses over time.

Immediate Tasks

Immediate tasks include (1) comprehensive reviews of extant, fugitive, and emerging research to determine what is already known about cognitive pathways relative to differing disabilities; (2) the development of a taxonomy of research tasks whereby different researchers could undertake different tasks; (3) creation of a catalyst network that facilitates and manages the synthesis of results across research initiatives; (4) enlistment of funding sources and development of political will to the research agenda; and (5) rapid prototyping and pilot testing of promising practices.

Longer-term Tasks

Longer-term tasks include (1) conducting rigorous research on promising practices that have been pilot tested; (2) conducting impact and consequential validity studies; (3) bringing promising practices to scale and conducting research on the scale-up process/effectiveness; and (4) ensuring adequate training for those who will implement various aspects of the approaches based on research findings.

Design

The design of the TEA should follow standard assessment guidelines and meet the technical quality requirements set forth in the *Standards for Educational and Psychological Tests* (AERA, APA, & NCME, 1999) and best practice (Downing & Haladyna, 2006). Statistical and qualitative data analyses should be rigorous and appropriate. For example, specific measurement models or statistical analyses should be used only when the assumptions underlying the models and statistical applications can be adequately met. Researchers need to incorporate a research design as close to experimental as possible, where students are randomly grouped by treatment or control conditions. The design does not necessarily require, for example, a comparable paper-based assessment; however some ability to differentiate true benefits of the TEA above and beyond current approaches (whatever those may be) is important.

Special attention should be paid to the interaction of student access to the assessment and the construct being measured, especially as it impacts the performance of students with disabilities. The investigations of access should not be limited to accommodations and other access tools. It should include how TEA is able to provide access to barrier-free construct-relevant performance, which requires greater attention to the target construct and how it is measured. For example, examinee response flexibility and the capability and flexibility of the TEA to capture data about the construct of interest is critical. As such, TEA tools and the research methods should attend to influences on, and the interactions between access and construct variables (Almond et al., 2010).

The TEA research design should also incorporate specific methods for collecting and rigorously analyzing and testing models of student cognition, to include learning progressions and changes in cognition across time and task. In either case, the design should capture not only what students know, but also what they do not know and the strategies they use.

Programmatic evaluation of the instructional and learning environment is also an important feature of the design to consider. Understanding of the environment is important to the utility of the TEA. For example, a TEA may measure the construct perfectly yet be completely unusable in the classroom, provide little instructionally relevant information, or be unsupported by teachers, technical limits, or competing policy.

The following steps, in addition to standard test development practice, are recommended for addressing the research questions taking into account the timeframe of immediate and loner-term tasks suggested above. Some of the steps may be undertaken simultaneously.

1. Identify or conduct reviews of extant, fugitive, and emerging research on topics such as the following:
    - Cognitive pathways, with particular attention to students with disabilities of various kinds.
    - Learning progressions, with particular attention to students with disabilities of various kinds.

- Ways in which technology has been used effectively to increase access to learning and valid assessment for students with various disabilities.

2. Design and implement research-based pilot tests and performance assessments of learning progression models that are grounded in cognitive pathway theory and/or models. Different pilot projects would address different types of disabilities.
   - Use the pilot test findings to refine approaches and validate the progressions for a variety of students.
   - Try the refined approaches with a different sample of students.
   - Refine and test bringing the approach to scale.
3. Design and pilot test assessments that are informed by the learning progression research conducted in step 2. Different pilot test projects would address different testing environments, such as the following:
   - Formative testing (during units of instruction to allow students and teachers to modify learning and instruction).
   - Benchmark/interim testing (between units to monitor within year progress toward annual goals).
   - Accountability testing (summative, typically end-of-year to determine level of proficiency for policymakers).
4. Establish the technical quality of the assessments, including the extent to which they pinpoint a student's progress relative to a learning progression and benchmark.
5. Determine how technology can add value to the instructional and assessment processes, with the following methods:
   - Identify or design prototypes.
   - Pilot test prototypes.
   - Refine, validate, and try out with another sample.
   - Refine and bring to scale.

Documentation

Finally, as with research on any assessment, plans for gathering evidence of both the technical quality and consequential validity of TEAs should be set forth at the beginning of the research design. The level of technical quality required for assessments of students with disabilities, when the assessments will be used to make placement or other high-stakes decisions, should be at least as high as for students who do not have disabilities. In addition to adequate levels of reliability and validity for the inferences that will be made using the assessment data, consequential validity evidence should be documented and acted upon. It is highly desirable that student performance in response to varying interventions be documented to monitor progress relative to learning targets and to adjust instruction. Such documentation will benefit from user-friendly technology-enhanced processes that allow practitioners to quickly document and assess teaching-learning efforts.

Conclusions

With the upcoming reauthorization of the Elementary and Secondary Education Act (ESEA), an opportunity exists to use TEA to develop assessments that are not only more inclusive for more students, but also significantly improve the validity of assessment results for all students, especially students with disabilities. There is an urgent need to articulate a research agenda for TEA that capitalizes on the increased access such assessments provide for students with disabilities. The shift to TEA is already underway, driven by states' increasing use of technology to assess all students, as educators seek more diagnostic information from their assessment system and look for better ways to evaluate skills not easily measured on paper-and-pencil assessments.

TEA holds the promise of particular benefit for students with disabilities for understanding their cognitive pathways and learning progressions, and the ways in which these may be different than those for typical students. There are many potential benefits of using technology for assessment for all students that have been discussed for more than a decade (Bennett, 1995, 2002):

- Allowing for cost savings and faster reporting when compared to traditional modes of assessment (multiple choice and constructed response).
- Potentially minimizing sources of construct-irrelevant variance when accommodation or support is provided.
- Allowing the assessment to be more authentically situated.

- Allowing for additional data about student interaction with a problem space to be collected.
- Measuring student use of technology-based tools and resources for problem solving in a technology-rich environment that may be analogous to how the activity would be completed in the classroom and/or real world.

In summary, there are prototypes that demonstrate the possibilities TEA offers to assess a full range of content and skills with interactive-dynamic formats that are flexible and can be personalized. Computerized adaptive testing (CAT) redesigns educational measurement instruments for interactive delivery. Its objective is to select, for each examinee, the set of test questions from a precalibrated item bank that most efficiently measures that person's achievement (Thompson & Weiss, 2009). Dynamic assessment provides corrective feedback in response to student failure to measure both the product and process of learning (Caffrey et al., 2008). Scaffolding is a dynamic assessment strategy used to enhance items by providing support during learning that are gradually removed when learning becomes solidified and/or the learner becomes more independent. Scaffolding in assessment includes structural assistance introduced to organize information or guide responses embedded in the presentation of the item or task (Perie et al., 2009). Another example of the innovations available through technology-enabled assessment is virtual performance assessment, where the student experiences an immersive virtual performance assessment, via computer interface, with 3-D virtual contexts, digital artifacts, and avatar-based identities that capture what students are doing (Clarke-Midura et al., 2010; Mislevy et al., 2003). Requirements for assessments in Response to Intervention (RtI) have motivated the development of innovative TEA specifically for students with disabilities. Among these are web-based research prototype reading environments that include a curriculum-based measurement tool—maze and oral reading fluency. These dynamic and interactive forms of assessment hold promise for improving the range and depth of assessments for all students, especially students with disabilities.

With the need to document the technical quality of these assessments, research is needed to establish their validity, especially for students with disabilities. Critical research topics include the following:

1. Improving measurement of cognition by investigating how the cognitive pathways and learning progressions of students with disabilities may converge with or differ from those of typical students and how error patterns may reveal the challenges students with disabilities face in following learning progressions to achievement of standards. These investigations hold promise for using this deepened understanding of the learning of students with disabilities to inform teaching strategies.
2. Use of interactive-dynamic assessments, including gaming technology, simulations, collaboration environments, and virtual worlds, with students with disabilities to help them navigate the technology environment and access the assessments to demonstrate their achievement of complex skills and performances embedded in standards.
3. Identifying the evidence needed for establishing validity for TEA, especially for students with disabilities. This evidence may require the development of new psychometric models, identifying the data needed to support inferences of TEA, minimizing construct irrelevant variance and increasing measurement precision. The challenge lies in making full use of the extensive data that TEA makes it possible to collect about students' performance.

Challenges for conducting research on TEA include technology and resource limitations, access of researchers to adequate populations to produce meaningful findings, and the political will to pursue the next evolution in assessment. Rather than continuing down the path of adapting traditional assessment strategies, policymakers, assessment providers, and other stakeholders will need to marshal the resources and commitment necessary to fully realize the promise of TEA to make innovation in assessment possible.

The authors offer recommendations for designing a national research agenda. Immediate opportunities are provided by expanding the types of investigations already underway in interactive-dynamic assessment, providing increased access and range through scaffolding and expanded access tools, and expanding assessment range and depth through virtual performance assessments. TEA makes available a wealth of information about cognitive pathways and learning progression. Researchers need to take action to fully utilize the opportunities to investigate these promising windows into student learning.

These innovations need validation for application to large-scale summative assessments and expanded applications in interim and formative assessment. Immediate tasks include the following:

1. Comprehensive reviews of extant, fugitive, and emerging research to determine what is already known about cognitive pathways relative to differing disabilities.
2. The development of a taxonomy of research tasks whereby different researchers could undertake different tasks.
3. Creation of a catalyst network that facilitates and manages the synthesis of results across research initiatives.
4. Enlistment of funding sources and development of political will to implement the research agenda.
5. Rapid prototyping and pilot testing of promising practices.

Longer-term tasks are as follows:

1. Conducting rigorous research on promising practices that have been pilot tested.
2. Conducting impact and consequential validity studies.
3. Bringing promising practices to scale and conducting research on the scale-up process/effectiveness.
4. Ensuring adequate training for those who will implement various aspects of the approaches based on research findings.

Careful attention to documentation of the technical quality, especially the validity, of TEA should be set forth at the beginning of the research design. Establishing high technical quality for assessments of students with disabilities, when the assessments will be used to make placement or other high-stakes decisions, should be at least as high a priority as that for students who do not have disabilities.

References

Almond, P., Winter, P., Cameto, R., Russell, M., Sato, E., Clarke, J., Torres, C., Haertel, G., Dolan, B., Beddow, P., & Lazarus, S. (2010). *Technology-enabled and universally designed assessment: Considering access in measuring the achievement of students with disabilities— A foundation for research*. Dover, NH: Measured Progress, and Menlo Park, CA: SRI International.

Almond, R.G., Steinberg, L.S., & Mislevy, R.J. (2002). Enhancing the design and delivery of assessment systems: A four-process architecture. *Journal of Technology, Learning, and Assessment, 1*(5).

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Ash, K. (2008). Computer-adaptive testing addresses individual student needs, but cost and logistical challenges persist. *Education Week. 28*(13), 19–21. Retrieved August 24, 2009, from http://www.edweek.org/ew/articles/2008.

Bechard, S., & Snow, J. (2010). *Identifying students in need of modified achievement standards and developing valid assessments*. Poster session presented at the annual meeting of the American Educational Research Association, Denver, CO.

Bejar, I.I., Lawless, R.R., Morley, M.E., Wagner, M.E., Bennett, R.E., & Revuelta, J. (2003). A feasibility study of on-the-fly item generation in adaptive testing. *Journal of Technology, Learning and Assessment, 2*(3).

Bennett, R.E. (1995). *Computer-based testing for examinees with disabilities: On the road to generalized accommodation*. ETS Research Memorandum (RM-95-1). Princeton, NJ: Educational Testing Service.

Bennett, R.E. (2002). Inexorable and inevitable: The continuing story of technology and assessment. *Journal of Technology, Learning, and Assessment, 1*(1).

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment, *Phi Delta Kappan. 5*, 139–148.

Burkhardt, H. (2009). On strategic design. *Educational Designer, 1*(3). Retrieved January 10, 2010, from http://www.educational/desginer.org/ed/volume1/issue3/article9.

Byers, C. (2001). Interactive assessment: An approach to enhance teaching and learning. *Journal of Interactive Learning Research, 12*(4), 359–374. Chesapeake, VA: AACE.

Caffrey, E., Fuchs, D., & Fuchs, L.S. (2008). The predictive validity of dynamic assessment: A review. *The Journal of Special Education, 41*, 254–270.

Cahalan-Laitusis, C. (2009, November). Public testimony at the U.S. Department of Education Race to the Top Assessment Program: Public and Expert Input Meeting. Atlanta: GA. 152-158. Retrieved March 3, 2010, at http://www2.ed.gov/programs/racetothetop-assessment/atlanta-transcript-2.pdf.

Clarke-Midura, J., Dede, C., & Mayrath, M. (2010, May). *Designing immersive virtual environments for assessing inquiry*. Paper presented at the meeting of the American Education Research Association, Denver, CO.

Darling-Hammond, L., & Ducommun, C.E. (2010). *Performance counts: Assessment systems that support high-quality learning*. Washington, DC: Council of Chief State School Officers.

Dolan, R.P., Hall, T.E., Banerjee, M., Chun, E., & Strangman, N. (2005). Applying principles of universal design to test delivery: The effect of computer-based read-aloud on test performance of high school students with learning disabilities. *Journal of Technology, Learning, and Assessment, 3*(7).

Downing, S M., & Haladyna, T.M. (Eds.). (2006). *Handbook of test development*. Mahwah, NJ: Lawrence Erlbaum Associates.

Edyburn, D.L. (2000). Assistive technology and students with mild disabilities. Focus on *Exceptional Children, 32*(9), 1–24.

Embretson, S.E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93*, 179–197.

Ferrara, S., Duncan, T.G., Perie, M., Freed, R., McGivern, J., & Chilukuri, R. (2003, April). *Item construct validity: Early results from a study of the relationship between intended and actual cognitive demands in a middle school science assessment*. Paper presented at the American Educational Research Association, Chicago, IL.

Fuchs, D., & Fuchs, L.S. (1998). Treatment validity: A unifying concept for reconceptualizing the treatment of learning disabilities. *Learning Disabilities Research and Practice, 13*, 204–219.

Fuchs, D., & Fuchs, L. S. (2006). Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*, 93–99.

Fuchs, D., Fuchs, L.S., McMaster, K.N., & Al-Otaiba, S. (2003). Identifying children at risk for reading failure: Curriculum-based measurement and the dual-discrepancy approach. In H.L. Swanson & K.R. Harris (Eds.), *Handbook of learning disabilities* (pp.431–449). New York: Guilford Press.

Fuchs, L.S., Compton, D.L., Fuchs, D., Hollenbeck, K.N., Craddock, C., & Hamlett, C.L. (2008). Dynamic assessment of algebraic learning in predicting third graders' development of mathematical problem solving. *Journal of Educational Psychology, 100*(4), 829–850.

Fuchs, L.S., & Fuchs, D. (2008). *What is scientifically based research on progress monitoring?* Washington, DC: National Center for Student Progress Monitoring.

Gorin, J.S. (2007). Test design with cognition in mind. *Educationa Measurement: Issues and Practice, 25*(4), 21–35.

Grigorenko, E.L., Ed. (2009). Dynamic assessment and response to interventions. *Journal of Learning Disabilities, 42*, 111–132.

Grigorenko, E.L., & Sternberg, R.J. (1998). Dynamic testing. *Psychological Bulletin, 124*, 75–111.

Haladyna, T.M., & Downing, S.M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice 23*(1), 17–27.

Hansen, E.G., Zapata-Rivera, D., & Feng, M. (2009, April). *Beyond accessibility: Evidence Centered Design for improving the efficiency of learning-centered assessments*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, California. April 16, 2009.

Heritage, M. (2008, February). *Learning progressions: Supporting instruction and formative assessment*. Council of Chief State School Officers: Washington, DC.

Herman, J.L. (2007, March). *Issues in the use of multiple measures for NCLB*. National Center for Research on Evaluation, Standards and Student Testing (CRESST): Los Angeles, CA.

Individuals with Disabilities Education Act of 2004, Pub L. No 108-446 Stat. 2647 (2004).

Jerald, C. D. (2009, July). *Defining a 21st century education*. Washington, DC: The Center for Public Education.

Johnstone, C.J., Moen, R.E., Thurlow, M.L., Matchett, D., Hausmann, K.E., & Scullin, S. (2007). What do state reading test specifications specify? Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Johnstone, C. & Thurlow, M. (in press). Statewide testing of reading and possible implications for students with disabilities. *Journal of Special Education*.

Johnstone, C., Thurlow, M., Altman, J., Timmons, J., & Kato, K. (2009). Assistive technology approaches for large-scale assessment: Perceptions of teachers of students with visual impairments. *Exceptionality, 17*(2), 66–75.

Johnstone, C.J., Thurlow, M.L., Thompson, S.J., & Clapper, A.T. (2008). The potential for multi-modal approaches to reading for students with disabilities as found in state reading standards. *Journal of Disability Policy Studies, 18*(4), 219–229.

Kingston, N., & Nash, B. (2009, April). *The efficacy of formative assessment: A meta-analysis*. Paper presented at the annual meeting of the American Educational Research Association. San Diego, CA.

Leighton, J.P., & Gokiert, R.J. (2005, April). *The cognitive effects of test item features: Informing item generation by identifying construct irrelevant variance*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Montreal, Quebec, Canada.

Mislevy, R.J., & Haertel, G.D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6–20.

Mislevy, R.J., Steinberg, L.S., & Almond, R.G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*, 3–67.

Mislevy, R.J. (2009, February). *Validity from the perspective of model-based reasoning* (CRESST Report 752). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing.

National Center for Education Statistics. (2009). *Table 2.7: Use of technology and capacity to use technology, by state: 2008–09*. State Education Reforms (SER). Assessment and Standards. Retrieved May 19, 2010, from: http://nces.ed.gov/programs/statereform/aas.asp.

National Governor's Association (NGA), Council of Chief State School Officers (CCSSO), and Achieve (2008). *Benchmarking for success: Ensuring U.S. students receive a world-class education*. Washington, DC: Authors.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. The Committee on the Foundations of Assessment. Pellegrino, J., Chudowsky, N., & Glaser, R., (Eds.). Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

No Child Left Behind Act of 2001, Pub L. No. 107-110, 115 Stat. 1425 (2002).

Office of Educational Technology. (2010, March). *Executive summary: Draft national educational technology plan 2010*. Washington, DC: U.S. Department of Education. Retrieved May 19, 2010, from http://www.ed.gov/technology/netp-2010.

Paas, F., Renkl, A., &Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychology, 38*(1), 1–4.

Pellegrino, J. (2004). Connecting cognition and assessment. In CASET Associates (Ed.), Math/Science Partnerships Workshop: Assessment of student learning (pp. 2–59). Washington, DC: The National Academies National Research Council National Science Resource Center.

Pellegrino, J. (2009). *The challenges of conceptualizing what low achievers know and how to assess their competence. In Perie, M (Ed.), Considerations for the Alternate Assessment based on Modified Achievement Standards (AA-MAS): Understanding the eligible population and applying that knowledge to their instruction and assessment– A white paper commissioned by the New York Comprehensive Center in collaboration with the New York State Education Department Project* (pp. 90–151). Dover, NH: Center for Assessment.

Perie, M., Marion, S., Gong, B., & Wurtzel, J. (November, 2007). *Role of interim assessments in a comprehensive assessment system: A policy brief*. Dover, NH: National Center for the Improvement of Educational Assessment.

Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues and Practice, 28*, 5–13.

Perie, M. (2009). Glossary. In Perie, M. (Ed.), *Considerations for the Alternate Assessment based on Modified Achievement Standards (AA-MAS): Understanding the eligible population and applying that knowledge to their instruction and assessment–A white paper commissioned by the New York Comprehensive Center in collaboration with the New York State Education Department Project* (pp. 90–151). Dover, NH: Center for Assessment.

Poehner, M.E. (2008). Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development. *Educational Linguistics, 9*, 3–20.

Popham, J.W. (April 2007). The lowdown on learning progressions. *Educational Leadership, 64*(7), 83–84.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon, 9*(5). NCB University Press, October 2001. Retrieved May 5, 2009, from: http://pre2005.flexiblelearning.net.au/ projects/resources/Digital_Natives_Digital_Immigrants.pdf.

Quellmalz, E.S., & Pellegrino, J. (2009). Technology and testing. *Science, 323*(5910), 75–79. Retrieved October 22, 2009, from http://www.sciencemag.org/cgi/content/full/323/5910/75.

Race to the Top Fund: Notice of proposed priorities, requirements, definitions, and selection criteria, 74(144) Fed. Reg. 37804 (July 29, 2009), pp. 37803–37837.

Samuels, S.J., Schermer, N., & Reinking, D. (1992). Reading fluency: Techniques for making decoding automatic. In S. J. Samuels & A E. Farstrup (Eds.), *What research says about reading instruction* (2nd ed., pp. 124–144). Newark, DE: International Reading Association.

Schoenfeld, A.H. (2009). Bridging the cultures of educational research and design. *Educational Designer, 1*(2). Retrieved January 10, 2010, from http://www.educational/desginer.org/ed/volume1/issue3/article5.

Sheinker, J., & Redfield, D. (2001). *Handbook for professional development in assessment literacy*. Washington, DC: Council of Chief State School Officers.

Simba Information. (2009, March). *K–12 technology tools & trends 2009* (ID: CURP2041370). Available at http://www.simbainformation.com/pub/2041370.html.

Singleton, J.L., Supalla, S., Litchfield, S., & Schley, S. (1998). From sign to word: Considering modality constraints in ASL/English bilingual education. *Topics in Language Disorders, 18*(4), 16–29.

Sireci, S.G., & Pitoniak, M.J. (2007). Assessment accommodations: What have we learned from research? In Cahalan Laitusis, C., & Cook, L.L. (Eds.). *Large-scale assessments and accommodations: What works?* Arlington, VA: Council for Exceptional Children.

Smith, C., Wiser, M., Anderson, C., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and atomic-molecular theory. *Measurement, 14*(1&2), 1–98.

Sternberg, R., & Grigorenko, Y. (2002). *Dynamic testing*. New York: Cambridge University Press.

Thompson, N., & Wiess, D. (2009). Computerised and adaptive testing in educational assessment. In F. Sheuermann & J. Björnsson (Eds.), *The transition to computer-based assessment. New approaches to skills assessment and implications for large-scale testing* (pp. 127–133). Luxemburg: Office for Official Publications of the European Communities.

Thompson, S.J., Thurlow, M.L., Quenemoen, R.F., & Lehr, C.A. (2002). *Access to computer-based testing for students with disabilities* (Synthesis Report 45). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved February 6, 2004, from http://education.umn.edu/NCEO/OnlinePubs/Synthesis45.html.

Thurlow, M.L. (2009, November 18). *Assessments for students with disabilities*. Expert presenter at the U.S. Department of Education Race to the Top Assessment Program: Public and Expert Input Meeting. Atlanta: GA.

Truziel, D., & Haywood, H.C. (1992). The development of interactive-dynamic approaches to assessment of learning potential. In H. C. Haywood & D. Truziel (Eds.), *Interactive assessments*. New York: Springer-Verlag.

Tucker, B. (2009, February). *Beyond the bubble: Technology and the future of student assessment* (Education Sector Reports). Washington, DC: Education Sector.

U.S. Department of Education (2009, December). *Executive summary: Race to the top assessment program, Notice of public and expert input meetings and request for input*. Retrieved December 30, 2009, from http://www2.ed.gov/programs/racetothetop-assessment/executive-summary.pdf.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wilson, M.R., & Berenthal, M.W. (2005). *Systems for state science assessment*. Washington, DC: National Academies Press.

Author Biographies

Sue Bechard, Ph.D., is a consultant for Inclusive Educational Assessment. Her research focuses on the development of and improvements to assessments for students with disabilities. She recently directed three federal research grants addressing assessments based on alternate and modified achievement standards and serves as an advisor to other assessment research projects. During her ten year tenure at Measured Progress, she investigated issues related to large-scale assessment and students in special populations, working with state departments of education to design and implement alternate assessment programs, research studies, and professional development opportunities. She can be reached at sbechard@me.com.

Jan Sheinker, Ed.D., is a consultant working with States on standards, descriptors, curriculum, assessment, instruction, school improvement, and systems alignment for general and special populations. In more than 38 years in public education, Sheinker has taught and supervised general and special education classrooms and assessment teams, authored instructional and professional development materials, and facilitated the development of state and district standards, and general and special education assessments. Sheinker is an affiliate of and consultant to several national organizations and a technical assistance provider for the U. S. Department of Education. Her email address is jansheinker@gmail.com.

Rosemary Abell is a Principal Consultant at ROSE Consulting and can be reached at rosemaryfitton@aol.com.

Karen Barton is a Principal Research Scientist at CTB/McGraw-Hill and can be reached at karen_barton@ctb.com.

Kelly Burling is a Senior Research Scientist at Pearson Assessments and can be reached at kelly.burling@pearson.com.

Christopher Camacho is former Director of Research at Children's Progress, and is now at LanguageMate as Director of Research. He can be reached at chris.camacho@me.com.

Renée Cameto is a Senior Research Social Scientist at SRI International and can be reached at renee.cameto@sri.com.

Geneva Haertel is Director of Assessment Research and Design, Center for Technology and Learning, SRI International and can be reached at geneva.haertel@sri.com.

Eric Hansen is a Research Scientist at Educational Testing Service (ETS) and can be reached at ehansen@ets.org.

Chris Johnstone is a Research Associate at National Center on Educational Outcomes (NCEO) and can be reached at john4810@umn.edu.

Neal Kingston is Director of the Center for Educational Testing and Evaluation at University of Kansas and can be reached at nkingsto@ku.edu.

Elizabeth (Boo) Murray is a Senior Research Scientist/Instructional Designer at Center for Applied Special Technology (CAST) and can be reached at bmurray@cast.org.

Caroline E. Parker is a Senior Research Scientist at Education Development Center, Inc. (EDC) and can be reached at cparker@edc.org.

Doris Redfield is President and CEO of Edvantia, Inc. and can be reached at doris.redfield@edvantia.org.

Bill Tucker is Managing Director at Education Sector and can be reached at btucker@educationsector.org.

J·T·L·A

The Journal of Technology, Learning, and Assessment

Editorial Board

www.jtla.org

Technology and Assessment Study Collaborative
Caroline A. & Peter S. Lynch School of Education, Boston College