Formative Assessment with Cognition in Mind: The Cognitively Based Assessment *of, for* and *as* Learning (CBAL™) Research Initiative at Educational Testing Service

Meirav Arieli-Attali
Educational Testing Service
mattali@ets.org

Traditional standardized educational tests often serve a summative purpose, reporting what students know *after* the learning phase has been completed. In this sense, they serve as an assessment *of* learning, with little or no implication for future instruction. By contrast, assessments may be used *formatively*; that is, they may be used to change the course of instruction as part of the assessment process. In the last 10-15 years, this form of assessment has seen a revival of interest in educational measurement. In this talk, I present different approaches to formative assessment, focusing on one specific formative system that is part of the Cognitively Based Assessment *of*, *for* and *as* Learning (CBAL™) research initiative at Educational Testing Service (Bennett, 2010). The CBAL initiative includes both summative and formative assessment components, connected through a conceptual model that lays out what it means to be competent in a specific content domain. Competency models are developed based on research findings in the cognitive and learning sciences, thus serving as a bridge between educational measurement and theories of cognitive development. Assessment tasks are then created on the basis of these conceptual models. By using technology-enhanced tasks, such as simulations and interactive tasks, we can better diagnose student strengths and weaknesses and recommend further instructional steps ("assessment *for* learning"). Technology also allows us to embed tasks in realistic scenarios in which students can, for example, learn to connect targeted skills to conditions of use ("assessment *as* learning"; Bennett, 2010).

Key words: formative assessment, evidence centered design, technology enhanced assessment

Formative assessment is a term used to denote assessment with the purpose of informing and improving teaching and learning. Although formative assessment systems and practices are fairly new, the roots of the term may be traced back to the 1960s, to distinctions between formative and summative evaluation (Bloom, 1969; Scriven, 1967). Bloom used the distinction to refer to evaluation of the student, while Scriven used it to refer to evaluation of a program. . While summative assessment aims to give an overall evaluation, formative assessment primarily targets *improvement* of the program or the student learning (Black and Wiliam, 2003). In this sense, traditional educational tests often serve a summative purpose, reporting what students know at a particular point in time, usually at the end of a learning phase. As a result, they often have little or no impact on future instruction. Formative assessment, as a reaction to the limitations of summative assessment, aims to address the issue of lack of impact on instruction. The initial purpose of formative assessment is to improve teaching and learning, in the sense of informing and changing the course of instruction based on the information gathered in the *process* of assessment. Some definitions of formative assessment are based on three key instructional processes: (a) establishing where learners are in their learning, (b) where they are going, and (c) how to get there (cf. Leahy & Wiliam, 2011). One widely accepted definition of formative assessment describes it as a classroom-based process in which students and teachers collect evidence of learning in order to understand current learning progress and to make adjustments to learning or to teaching as necessary (Council of Chief State School Officers, 2008; see also, Black, Harrison, Lee, Marshall, & Wiliam, 2003; Black & Wiliam, 1998a, 1998b; Wylie & Lyon, 2012). An important distinction is made here among three perspectives: formative assessment as a "tool" or an "instrument," such as the test itself, as a *process* that can be applied absent any tool, and as an integration of process and methodology that includes instrumentation but includes also teacher support materials, and is part of a more comprehensive approach to instruction (Bennett, 2011). We refer here to the last of those conceptualizations.

In the last 10-15 years, we have seen a surge of formative assessment systems, in many different variations and forms, which do not always agree on the set of practices needed in order to achieve formative goals. In this paper, I present the Cognitively Based Assessment *of*, *for* and *as* Learning (CBAL™) research initiative at Educational Testing Service (ETS) (Bennett & Gitomer, 2009; Bennett, 2010). The CBAL research initiative began at ETS in 2007. As a research initiative, its central goal is to create a model for a future comprehensive system of assessment that documents what students have achieved ("*of* learning"), helps identify how to plan and adjust instruction ("*for* learning"), and is considered by students and teachers to be a worthwhile educational experience in and of itself ("*as* learning") (Bennett, 2010[1]). Thus, CBAL includes both summative and formative assessment components. In this paper, I focus on one unique characteristic that is central to CBAL and on which the formative and the summative assessment systems rest: the incorporation of ideas and findings from the cognitive and learning sciences into the process of assessment development. I will first discuss this aspect and why making the link between cognitive science and psychometrics is needed, then describe the CBAL initiative and its conceptual basis, and lastly illustrate how, by using technology in assessment, we better achieve the goals of an assessment system that is cognitively based.

[1] see also CBAL website http://www.ets.org/research/topics/cbal/initiative

Linking Cognitive Science and Psychometrics

The idea of linking cognitive science and psychometrics is not new. Cronbach, in his 1957 address to the American Psychological Association (APA) called for the unification of differential and experimental psychology, the first being the branch of psychology concerned with the study and measurement of individual differences, and the latter being the roots of what we call today "cognitive science" (Cronbach, 1957). Pellegrino, Baxter and Glaser (1999), in their comprehensive review of findings in cognitive science show how those findings are relevant to and can be incorporated in assessment design. The National Research Council (NRC) in *"Knowing What Students Know"* (2001) also emphasized the need to marry theories of cognition and learning with assessment practices. Their call rests on identifying limitations of current assessments, namely that they (a) are "snapshots" of achievement at particular points in time, thus not capturing progression of students' conceptual understanding over time; (b) often have limited if any useful implications for improving learning and teaching; and moreover (c) do not seem to capture complex knowledge and skills essential for success in the information-based economy of the 21st century, such as organization of knowledge, problem representation, strategy use, self-monitoring skills, and individual contributions to group problem solving. The motivation to forge a link between the theory and principles generated from cognitive science and practices of assessment design appears throughout the recent psychometrics literature (e.g., Embretson & Gorin, 2001; Leighton, & Gierl, 2007a; 2007b).

Recent work has provided a framework for this kind of linkage. Embretson (1999) developed the Cognitive-Design-System (CDS), which focuses on direct incorporation of a priori construct validation into item development, rather than only providing validity evidence retrospectively, as is usually done. This focus means that before items are developed, an understanding of the cognitive processes that are intended to be measured are specified and operationalized into the item stimulus features. Another approach that makes an inherent link between theories of cognition and assessment development, and which serves as the basis of the CBAL initiative, is the Evidence-Centered Design framework (Mislevy, Steinberg, & Almond, 2003). The Evidence-Centered Design (ECD) framework requires specification of a *Student Model*, i.e., a model that states the competencies to be measured and what these competencies include. ECD also requires specifications of an *Evidence Model*, i.e., a model that articulates what in student responses to an item can provide evidence for student knowledge and skills. The Evidence Model is the link between the conceptual Student Model and the assessment scores, thus serving as a portion of the justification or validity argument for making the claims about student competence.

The CBAL research initiative based on the ECD framework

Using the ECD framework, CBAL foundations include a student model, which consists of the *CBAL Competency Models*, and the *CBAL Learning Progressions*. A competency model is domain-specific, states the competencies that a student is expected to acquire in school in this domain (e.g., mathematics, language arts, science), and articulates what each competency entails. Learning progressions (LPs) articulate a *trajectory of learning over time* for a specific competency or topic in a domain (e.g., linear functions, proportional reasoning, modeling in mathematics). Competency models and learning progressions are developed based on a synthesis of findings from cognitive and learning sciences research. Yet, in order to be relevant to the educational system, the models and the progressions must strive to be consistent with curriculum. The development of these models is iterative in nature. For illustration, following is a brief description of the creation and the structure of the CBAL Mathematics Competency Model.

The CBAL Mathematics Competency Model began as a model for the middle school grades, created initially from an extensive literature review conducted by Graf (2009), and followed several revisions (Graf, Harris, Marquez, Fife, & Redman 2009; Haberstroh, Harris, Bauer, Marquez, & Graf, 2010). The model addresses *content* and *process* strands for Grades 6 through 8, consistent with the approach reflected in current mathematics standards (e.g., the NCTM 1999 standards; the U.S. Common Core State Standards, 2011). Process strands refer to mathematical thinking processes that cut across specific topics, while content strands generally follow curriculum topics relevant to the grade level, such as algebra, numbers and operations, measurement and geometry, and so forth. The cross cutting process strands for this model were heavily influenced from a synthesis of cognitive research findings done by Kilpatrick, Swafford, & Findell (2001). These processes define competencies such as the ability in mathematics to model, communicate, argue, etc. (for more details, see Graf, 2009; Graf, Harris, Marquez, Fife, & Redman 2009; Haberstroh, Harris, Bauer, Marquez, & Graf, 2010).

While competency models specify competencies and their relationships, they do not address expected progress in mastering those competencies over time. Learning progressions, in focusing on one competency or one topic, articulate a trajectory of learning and understanding over time, and in this respect, they have the potential to help teachers, students, and policy makers to set expectations about learning and can aid instructional planning. This aspect of LPs is what makes them relevant as a guide for formative assessment (Heritage, 2008; Leahy & Wiliam, 2011). It has been shown that availability of learning progressions can improve teaching (Clements & Sarama, 2004). In CBAL, a learning progression is defined as a description of qualitative change in a student’s level of sophistication for a key concept, process, strategy, practice, or habit of mind. Change in student standing on such a progression may be due to a variety of factors, including maturation and instruction. Each progression is presumed to be modal, i.e., to hold for most, but not all, students; and a progression is provisional, subject to empirical verification and theoretical challenge.

The learning progression approach is not unique to CBAL; it appears in several other educational research studies or programs, and may have different definitions or emphasis (cf. Confrey & Maloney, 2010; Corcoran, Mosher & Rogat, 2009; Leahy & Wiliam, 2011). Although it is acknowledged that learning progressions cannot be independent of curricular sequencing, different approaches may take a different focus and thus vary in their level of consistency with curriculum, and in the extent to which they incorporate research findings. Therefore, a question one might ask is “How do we develop learning progressions that incorporate research findings?” Much of the research literature in the cognitive and learning sciences addresses learning from a *developmental* perspective, often also providing suggestions for sequencing teaching of a topic or a competency. For example, Kalchman, Moss & Case (2001) and Kalchman & Koedinger (2005) proposed a way to teach functions that capitalizes on prior knowledge and developmental abilities, based on Case’s lifelong work in cognitive development. Case argues that a new cognitive concept or structure is developed as a result of a process of integrating separate prior concepts or structures (e.g., Case, 1993). In particular, in order to understand or create the concept of function, Kalchman et al. argue that an integration of previously separate concepts (or "*understandings*") needs to take place, namely an integration of *numeric* understanding and *spatial* understanding. They show that knowledge and understanding of patterns in number sequences on the one hand (numeric understanding), and knowledge and understanding of simple bar graphs on the other hand (spatial understanding) are needed in order to acquire the concept of a function as a relationship between numbers and variables (i.e., $y=mx+b$) and its graphical representation on a coordinate plane. In the approach of Kalchman and her colleagues, students have these separate understandings before any formal teaching of functions takes place. Kalchman et al. argue that teaching should capitalize on this prior knowledge and follow a sequence that both strengthens it and addresses the needs (e.g., focus on the connections) in order to establish a strong conception of functions. Our work in developing and defining a learning progression for functions and linear functions draws upon this cognitive approach.

Research findings often address what are called "misconceptions", which are incorrect or incomplete conceptions that students may hold at intermediate stages in development. In some cases these misconceptions persist unless they are replaced by ideas that are more nearly expert. A well-documented misconception is that proportionality involves a *constant additive difference* (cf. Karplus, Pulos, & Stage, 1983; Noelting, 1980a). This misconception often surfaces when a student is asked to compare ratios or to find an equivalent ratio; for instance, a student may claim that the ratio 5:6 is equivalent to 6:7, because in both cases the difference is one. This misconception may appear as part of developing an understanding of proportionally, but may persist if no direct teaching takes place. As early as Piaget (e.g., Piaget and Inhelder, 1975), researchers have been studying student understanding of proportional reasoning. Findings are consistent over the past 35 years, showing that students' developing concept of proportional reasoning primarily follows three stages: an early *qualitative-intuitive* stage, a *quantitative additive* stage and a *multiplicative structure* stage (e.g., Karplus, Pulos, & Stage, 1983; Lamon, 1993, 2007; Noelting, 1980a; 1980b; Piaget & Inhelder, 1975; Vergnaud, 1983). Our work in developing and defining a learning progression for proportional reasoning draws upon this cognitive approach.

In addition to synthesizing research findings, we also use the *Cognitive Lab* approach in creating or refining a learning progression. The Cognitive Lab method refers to a class of behavioral techniques designed to probe cognitive processes and knowledge representations, in a one-on-one session with a participant. One technique, the *think-aloud*, involves instructing a person to talk aloud his or her thinking while performing a task (Ericsson & Simon, 1993). A second technique, an interview, occurs during or after task completion and is similar to the clinical interview used by Piaget. Cognitive labs can be used in different stages of the development of a learning progression. They can be used at early stages of development in order to better understand the relationship between student understanding of different concepts or at later stages to validate the progression. For example, in our recent work around developing a learning progression for rational numbers (Arieli-Attali & Cayton-Hodges, 2013), we drafted a progression based on analysis of research findings, and then drafted preliminary tasks that we hypothesized would provide evidence to distinguish between adjacent levels. Observations from the cognitive interviews conducted with 14 students from third through fifth grades helped us to refine the definitions of the levels of the hypothesized progression. A review by an expert panel suggested additional revisions and refinement (for more details see Arieli-Attali & Cayton-Hodges, 2013).

The CBAL mathematics branch currently includes learning progressions primarily for middle school, such as Equality and Variable, Proportional Reasoning, and Linear Functions (cf. Arieli-Attali, Wylie, & Bauer, 2012). Current work has added a progression for high school, (i.e., Quadratic Functions; see Graf, Fife, & Marquez, 2013) and for elementary school (i.e., Rational Numbers; see Arieli-Attali & Cayton-Hodges, 2013). While task development ideally starts from the provisional learning progression, the process is iterative enough that progressions may be revised dramatically following task development or new progressions generated after task development. The next section addresses aspects of task development that aim to align with the rationale of an assessment system that is cognitively based.

CBAL scenario-based tasks with simulations: Using technology and cognitive components in assessment

As the presence and importance of technology in society increases, it is essential for assessment programs to use technology in order to maintain their relevance. Technology may also help us achieve the goal of incorporating cognitive components into our assessment, in that it allows researchers and developers to track and presumably assess *thinking processes* rather than to just assess the outcome of these processes. Simulations and interactive tasks may provide a window into student cognition, and in that allow us to better diagnose student strengths and weaknesses in conceptual understanding. Examples of how technology can support the identification of cognitive processes include 1) tracking the steps a student needs to get to an answer, 2) bringing forward erroneous student answers to earlier items to allow another chance to revise the answer, 3) editing behavior or correcting previous answers, or 4) measuring the response time for certain behaviors. In responding to a task, the strategies that students use, as well as whether or not they make use of available tools, may indicate their level of understanding. This type of information can help us to identify misconceptions, difficulties, inconsistencies or gaps in understanding, and to locate student understanding on a hypothesized learning progression. In the following, I illustrate how the specific characteristics of CBAL mathematics tasks allow us to identify the level of student understanding.

CBAL mathematics tasks are extended tasks, comprising a set of 12 to18 items built around the same real-world scenario. Examples include a swimmer who claims to have swum the Atlantic Ocean, an article in a newspaper claiming that Android Apps are growing at a faster rate than iPhone Apps, a person at an airport using an available moving sidewalk, or a water crisis and its implications for using dams for electricity. All of the examples chosen here lend themselves to mathematical modeling using linear functions, and we have other examples for other mathematical models as well. Most mathematics tasks also include a simulation that is presented at the outset of each task, and is available throughout the task. The simulations emulate the properties of the real-world situation, thus making it easier for the student to relate to the problem, but at the same time, making the connection with the mathematical model. For example, in the *Dams and Droughts* task the simulation provides a sink with a faucet and a plug, allowing the student to trace the effect on water level of different parameters, including the rate at which the water flows in or out the sink in a manner similar to how it would flow in or out of a lake with a dam. The *Moving Sidewalk* task includes a simulation that permits the student to observe the rate of motion of the sidewalk and the relationship between time and distance. In a task, called *Proportional Punch*, students can mix different amounts of water and punch to observe how different ratios affect the sweetness of the punch. A *sweetness meter* is provided as part of this simulation (see Figure 1). These examples are just a few that illustrate how a simulation is incorporated into the assessment as a tool to help students make the connection with the mathematical model.

Whether a student is using the simulation for answering particular items may be an indication of that student's level of understanding or mastery. In all learning progressions, earlier levels refer to more concrete understandings where the student still needs manipulative or visual representations, and higher levels refer to ones where student can abstract or generalize without this type of aid. As an illustration let us draw from the *Proportional Punch* task: we find that Level 2 students often use the simulation, but answer the item incorrectly; Level 3 students often use the simulation and answer correctly; and Level 4 students do not use the simulation and answer correctly. This example illustrates how different interactive behaviors can be linked to strategies and to levels in a progression, behaviors that could not easily be captured in traditional paper-and-pencil tests. Moreover, embedding a task in realistic scenarios may help students make the connection between targeted skills (e.g., procedures or models in mathematics) and conditions of use in real-world problems. This is one aspect of the CBAL term "assessment *as* learning." The preliminary feedback we have received suggests that extended scenarios and simulations increase student engagement and motivation, and in turn may increase the validity of the assessment and perhaps learning itself.

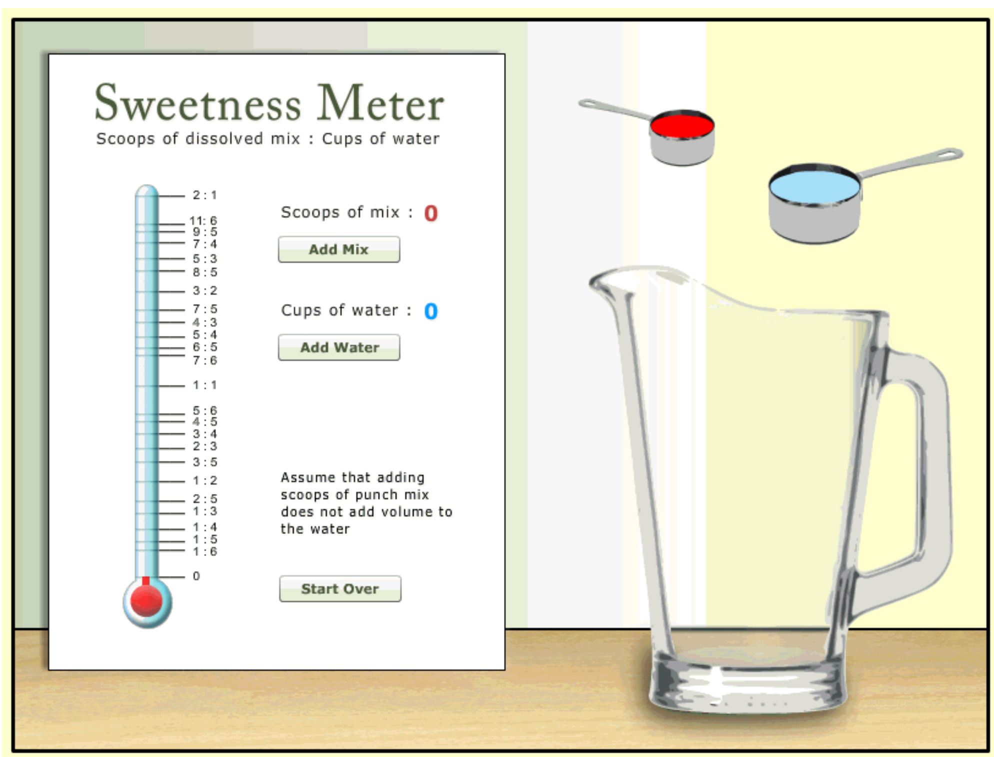

Figure 1. Sweetness-meter simulation as part of the Proportional Punch CBAL task.

Summary and Conclusion

This paper gives a brief overview of the CBAL assessment system that includes both summative and formative components, and focuses on the incorporation into assessment of theory and research findings from the cognitive and learning sciences. The paper illustrates via the CBAL mathematics strand how we incorporate cognitive aspects in the development of the competency models and learning progressions, and continue to preserve these components in the development of assessments that are technology enhanced. In this short paper we only touch upon these issues. An important outcome of incorporating research findings into assessment is that it also bridges research and pedagogy, bringing major findings and ideas into classroom practices (NRC, 2001). The fact that we are using research finding to create the formative assessment materials that teachers are then using in the classroom means that teachers are applying those research findings in their instruction.

The paper also shows how we capitalize on technology to serve the goal of cognitively based assessment, capturing student thinking processes. Yet, technology also poses challenges to assessment that this paper does not address. It is not always easy to make sense of the overwhelming amount of data that these assessment types can produce (Bennett, Jenkins, Persky, & Weiss, 2003). Not all the behaviors captured by the computer are relevant to the construct at hand, and it is even more important and more difficult in these cases to identify the presence of construct-irrelevant variance. Moreover, the fact that students can be tracked as they proceed through an assessment (e.g., mouse clicks) may refute the classical notion of assessment. Student behavior may change as a result of their awareness that their behaviors are recorded and may be used for scoring. These issues and others are the focus of our current research work in CBAL and at ETS.

References

Arieli-Attali, M., & Cayton-Hodges, G. (2013). *Expanding the CBAL Mathematics Assessment to elementary grades: The development of a competency model and a rational number learning progression*. Manuscript submitted for publication.

Arieli-Attali, M., Wylie, E.C., & Bauer, M. I. (2012, April). *The use of three learning progressions in supporting formative assessment in middle school mathematics*. Paper presented at the meeting of the American Educational Research Association, Vancouver, Canada.

Bennett, R. E. (2010). Cognitively Based Assessment *of*, *for*, and *as* Learning: A preliminary theory of action for summative and formative assessment. *Measurement: Interdisciplinary Research and Perspectives, 8*, 70-91.

Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, *18(1),* 5-25.

Bennett, R. E., & Gitomer, D. H. (2009). Transforming K–12 assessment. In C. Wyatt-Smith & J. Cumming (Eds.), *Assessment issues of the 21st Century* (pp. 43-61). New York, NY: Springer .

Bennett, R. E., Jenkins, F., Persky, H., & Weiss, A. (2003). Assessing complex problem solving performances. *Assessment in Education: Principles, Policy, and Practice, 10*, 347–359.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham, UK: Open University Press.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice 5(1)*, 7–74.

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappa 80(2)*, 139–48.

Black, P., & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal*, *29(5)*, 623–37.

Bloom, B.S. (1969). Some theoretical issues relating to educational evaluation. In R.W. Tyler (Ed.). *Educational evaluation: New roles, new means. The 63rd yearbook of the National Society for the Study of Education, part 2 (Vol. 69)*, (pp. 26–50). Chicago, IL: University of Chicago Press.

Case, R. (1993). Theories of learning and theories of development. *Educational Psychologist, 28*, 219-233.

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, *6(2)*, 81-89.

Common Core Standards Writing Team (2011). *Progressions for the Common Core State Standards in Mathematics (draft)*. Retrieved from http://commoncoretools.files.wordpress.com/2011/08/ccss_progression_nf_35_2011_08_12.pdf

Confrey, J., & Maloney, M. (2010). The construction, refinement, and early validation of the equipartitioning learning trajectory. Proceedings of the 9th Annual Conference of the Learning Sciences (Vol 1, pp. 968-975). Chicago, IL.

Corcoran, T., Mosher, F. A., & Rogat, A.. (2009). *Learning progressions in science: An evidence-based approach to reform* (Research Report 63). New York: Consortium for Policy Research in Education, Teachers College, Columbia University.

Council of Chief State School Officers (2008). Formative assessment: Examples of practice. A work product initiated and led by Caroline Wylie, Educational Testing Service, for the Formative Assessment for Students and Teachers Collaborative. Council of Chief State School Officers, Washington, DC.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12(11)*, 671-684.

Embretson, S. E. (1999). Cognitive psychology applied to testing. In F. T. Durso (Ed.), Handbook of applied cognition (pp. 629-660). Chichester, England: Wiley .

Embretson, S., & Gorin, J. (2001). Improving construct validity with cognitive psychology principles. *Journal of Educational Measurement*, *38*, 343-368.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Graf, E. A. (2009). *Defining mathematics competency in the service of cognitively based assessment for grades 6 through 8* (Research Report 09-42). Princeton, NJ: Educational Testing Service.

Graf, E. A., Fife, J. H., & Marquez, E. (2013). *The development of a quadratic functions learning progression and associated task models*. Unpublished manuscript. Princeton, NJ: Educational testing Service.

Graf, E. A., Harris, K., Marquez, E., Fife, J., & Redman, M. (2009). *Cognitively Based Assessment of, for, and as Learning (CBAL) in mathematics: A design and first steps toward implementation* (Research Memorandum 09-07). Princeton, NJ: Educational Testing Service.

Haberstroh, J., Harris, K., Bauer, M., Marquez, E. & Graf, E.A. (2010). *CBAL mathematics final report 2009*. Princeton, NJ: Educational Testing Service.
Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment*. Washington, DC: Formative Assessment for Teachers and Students, State Collaborative on Assessment and Student Standards of the Council of Chief State School Officers, Washington, DC.
Kalchman, M., & Koedinger, K. R. (2005). Teaching and learning functions. In National Research Council (Ed.), *How students learn* (pp. 351-393). Washington, DC: National Academies Press.
Kalchman, M., Moss, J., & Case, R. (2001). Psychological models for the development of mathematical understanding: Rational numbers and functions. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: 25 years of progress,* (pp. 1-38). Mahwah, NJ: Erlbaum.
Karplus, R., Pulos, S., & Stage, E. K. (1983). Proportional reasoning of early adolescents. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 45-86). New York, NY: Academic Press.
Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academies Press.
Lamon, S. J. (1993). Ratio and proportion: Connecting content and children's thinking. *Journal for Research in Mathematics Education, 24(1),* 41-61.
Lamon, S. J. (2007). Rational numbers and proportional reasoning: Toward a theoretical framework. In F. K. Kessler (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629-667). Charlotte, NC: Information Age Publishing.
Leahy, S. & Wiliam, D. (2011, April). *Devising learning progressions*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.
Leighton, J. P., & Gierl, M. J. (2007a). *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge, UK: Cambridge University Press.
Leighton, J. P., & Gierl, M. J. (2007b). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practice*, *26*(2), 3-16.
Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). Focus article: On the structure of educational assessments. *Measurement: Interdisciplinary research and perspectives, 1(1),* 3-62.
National Research Council (2001). *Knowing what students know: The science and design of educational measurement.* Washington, DC: National Academies Press,
Noelting, G. (1980a). The development of proportional reasoning and the ratio concept: Part I - Differentiation of stages. *Educational Studies in Mathematics, 11*, 217-253.
Noelting, G. (1980b). The development of proportional reasoning and the ratio concept: Part 11 - Problem structure at successive stages; problem solving strategies and the mechanism of adaptive restructuring. *Educational Studies in Mathematics, 11*, 331- 363.
Pellegrino, J. W., Baxter, G. P., & Glaser, R. (1999). Addressing the" two disciplines" problem: Linking theories of cognition and learning with assessment and instructional practice. *Review of research in education, 24*, 307-353.
Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. New York, NY: Norton.

Scriven, M. (1967). The methodology of evaluation. In *Perspectives of curriculum evaluation*, ed. R.W. Tyler, R.M. Gagne, and M. Scriven, 39–83. Chicago, IL: Rand McNally.

Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 127-174). New York, NY: Academic Press.

Wylie, C. & Lyon, C. (2012). Formative assessment – supporting students' learning. *Educational Testing Service R&D Connection*, *19*, 1-12. Princeton, NJ.