Formative Assessment System for Teachers™ Technical Manual

Last updated: July 10, 2018

520 Nicollet Mall
Suite #910
Minneapolis, MN 55402

612.254.2534
sales@fastbridge.org
www.fastbridge.org

FastBridge Learning

THIS WAY TO FASTER RESULTS.

Formative Assessment System for Teachers™ Technical Manual

Published by Theodore J. Christ and Colleagues, LLC (TJCC)

Distributed by TJCC and FastBridge Learning, LLC (FBL)

520 Nicollet Mall, Suite #910
Minneapolis, MN 55402
Email: sales@fastbridge.org
Website: www.fastbridge.org
Phone: 612-254-2534

Prepared by Theodore J. Christ, PhD as Senior Author and Editor with contributions from (alphabetic order) Yvette Anne Arañas, MA; LeAnne Johnson, PhD; Jessie M. Kember, MA; Stephen Kilgus, PhD; Allyson J. Kiss; Allison M. McCarthy Trentman, PhD; Barbara D. Monaghen PhD; Gena Nelson, MA; Peter Nelson, PhD; Kirsten W. Newell, MA; Ethan R. Van Norman, PhD; Mary Jane White, PhD; and Holly Windram, PhD as Associate Authors and Editors.

Citation:
Theodore J. Christ and Colleagues (2018). Formative Assessment System for Teachers™ Technical Manual, Minneapolis, MN: Author and FastBridge Learning.

Table of Contents

Table of Contents

Table of Contents

Table of Figures

Table of Tables

Table of Contents

Introduction to FAST™ and FastBridge Learning

The Formative Assessment System for Teachers™ (FAST™) Technical Manual provides an overview of FastBridge Learning and a detailed description of the technical evidence supporting FAST™ measures. This document includes the following sections:

- Introduction to FAST™ and FastBridge Learning
- FAST™ CBMreading
- FAST™ earlyReading
- FAST™ aReading
- FAST™ CBMmath (Process & Automaticity)
- FAST™ earlyMath
- FAST™ aMath
- FAST™ SAEBRS
- FAST™ DevMilestones

For each measure presented, information is organized into the following sub-sections:

- Purpose and Use
- Content Description
- Content Development
- Administration
- Scores and Scoring
- Construct Validity
- Evidence for Use as a Screening Tool
- Evidence for Use as a Progress Monitoring Tool

Background and Overview

FAST™ assessments were developed by researchers at universities from around the country, which include the Universities of Minnesota, Georgia, Syracuse, East Carolina, Buffalo, Temple, and Missouri. FAST™ cloud-based technology was developed to support the use of those assessments for learning. Although there is a broad set of potential uses, the system was initially conceptualized to make it easier for teachers.

FAST™ is designed for use within Multi-Tiered Systems of Support (MTSS) and Response to Intervention (RTI) frameworks for early intervention and prevention of deficits and disabilities. It is research- and evidence-based. FAST™ is distinguished and trusted by educators. It is transforming teaching and learning for educators and kids nationwide.

All in One

FAST™ is one, comprehensive, simple cloud-based system with Curriculum-Based Measurement and Computer-Adaptive Tests for universal screening, progress monitoring, MTSS/RTI support, online scoring, and automated reporting. It is easy to implement with online training and resources, automated rostering and SIS integration, nothing to install or maintain, and multi-platform and device support.

Support and Training

Our school support team is accessible and responsive for support via live chat, e-mail, or phone. When combined with our knowledge base—full of quick tips, articles, videos, webinars, and flipped training for staff—in addition to customized online or onsite training, your teachers and administration are supported at every step.

Trusted Results

FAST™ is an evidence-based formative assessment system that was developed by researchers at the University of Minnesota in cooperation with others from around the country. They set out to offer teachers an easier way to access and use the highest quality formative assessments. Researchers and developers are continuously engaged with teachers and other users to refine and develop the best solutions for them.

Curriculum-Based Measurement

Our Curriculum-Based Measures (CBM) are highly sensitive to growth over brief periods. We offer Common Core-aligned CBM measures with online scoring and automated skills analysis in FAST™ earlyReading, FAST™ earlyMath, FAST™ CBMreading, and FAST™ CBMmath.

Automated Assessments

Our Computer-Adaptive Tests provide a reliable measure of broad achievement and predict high-stakes test outcomes with high accuracy. Automatically adapting to students' skill levels to inform instruction and identify MTSS/RtI grouping, we offer FAST™ aReading and FAST™ aMath.

Prevention and Intervention

Designed for Multi-Tiered Systems of Support (MTSS) and Response to Intervention (RTI), FAST™ makes program implementation easy and efficient with automated scoring, analysis, norming and reporting; customizable screening, benchmarking, instructional recommendations, and progress monitoring.

FastBridge Learning has a strong foundation in both research and theory. FAST™ assessments were created to provide a general estimate of overall achievement in reading and math, as well as provide a tool to identify students at risk for emotional and behavioral problems. For reading and math assessments, item banks have been created containing a variety of items, including those with pictures, words, individual letters and letter sounds, sentences, paragraphs, and combinations of these elements. Overall, FastBridge Learning aims to extend and improve on the quality of currently available assessments.

Administration and Scoring

FAST™ is supported by an extensive set of materials to support teachers and students, including self-directed training modules that allow teachers to become certified to administer each of the assessments. FAST™ assessments can be administered by classroom teachers, special education teachers, school psychologists, and other individuals such as paraprofessionals. Administration times vary depending on which assessment is being administered. Online administrations require a hard copy of the student materials (one copy per student) and access to the FAST™ system (i.e., iPad or computer with Internet connection). Paper-and-pencil assessment administration materials and instructions are available in the Training and Resources section. As with any assessment, only students who can understand the instructions and can make the necessary responses should be administered FAST™ assessments. Assessments should be administered in a quiet area conducive to optimal performance. The brevity of FAST™ assessments aims to minimize examinee fatigue, anxiety, and inattention.

Interpretation of Test Results

The FastBridge Learning software provides various resources to assist administrators with test result interpretations. Methods of notation are also included to provide information regarding those students predicted to be at risk. Exclamation marks (! and !!) indicate the level of risk based on national norms. One exclamation mark refers to some risk, whereas two exclamation marks refer to high risk of reading difficulties or not meeting statewide assessments benchmarks, based on the score. Interpreting FastBridge Learning assessment scores involves a basic understanding of the various scores provided in the FAST™ system and helps to guide instructional and intervention development. FAST™ includes individual, class, and grade level reports for screening, and individual reports for progress monitoring. Additionally, online training modules include sections on administering the assessments, interpreting results, screen casts, and videos.

FAST™ calculates and reports the percentile ranks, or percentiles, of scores relative to same-grade peer performance in the class, school, district, and FAST™ users around the nation. Those percentiles are classified and color-coded in bands: $\leq 19.99^{th}$ (red), 20^{th} to 29.99^{th} (orange), 30^{th} to 84.99^{th} (green) and $\geq 85^{th}$ percentiles (blue). These standards were set to guide resource allocations for early intervention and prevention within multi-tiered systems of support (MTSS).

National norms are used to compare local performance to that of an external group. The standards (color codes) are applied to support decisions about core and system-level supports. Visual analysis of color codes is useful to estimate the typicality of achievement in the local population. They are often used in combination with benchmarks to guide school and district level decisions about instruction, curriculum and system-wide services (e.g., are the school-wide core reading services sufficient to prevent deficit achievement for 80% of students). If FAST™ data indicate that much more than 20% of a school or district's students are below the 20^{th} percentile on national norms, then remediation efforts in that area should be considered as the data suggest that the core instruction is not supporting adequate achievement. If they observe that fewer than 20% of the total school population are below the 20th percentile on national norms, their population is over-performing relative to others. Subsequently, the school should continue using effective services, but identify another domain of focus.

FAST™ reports provide tri-annual grade-level benchmarks, which generally correspond with the 15th and 40th percentiles on national norms. Scores below the 15th percentile are classified as "high-risk." Those at-or-above the 15th and below the 40th are "some risk;" and those at or above the 40th are "low risk." This is consistent with established procedures and published recommendations (e.g., RTI Network). It is common practice to use norm-reference standards at the 15th and 40th percentiles; or to use pre-determined standards on state achievement tests.

FAST^{TM} CBMreading

FAST^{TM} CBMreading Purpose and Use

FAST™ Curriculum-Based Measurement for Reading (FAST™ CBMreading) is a version of curriculum-based measurement of oral reading fluency (CBM-R), which was originally developed by Deno and colleagues to index the level and rate of reading achievement (Deno, 1985; Shinn, 1989). FAST™ CBMreading is a simple, efficient, evidence-based assessment used for universal screening in grades 1 through 8, and progress monitoring for grades 1-12. FAST™ CBMreading uses easy, time-efficient assessment procedures to determine a student's general reading ability.

FAST™ CBMreading emerged from a project funded by the Institute for Education Sciences in the US Department of Education entitled Formative Assessment Instrumentation and Procedures for Reading (FAIP-R). Early versions of those passages were used in published research (Ardoin & Christ, 2008; Christ & Ardoin, 2009). The goal in creating the FAST™ CBMreading measures was to systematically develop, evaluate and finalize research-based instrumentation and procedures for accurate, reliable, and valid assessment and evaluation of reading rate.

Students read aloud for one minute from grade- or instructional-level passages. The words read correctly per minute (WRCM) functions as a robust indicator of a reading health and a sensitive indicator of intervention effects. FAST™ CBMreading includes standardized administration and scoring procedures. FAST™ CBMreading provides teachers with timely information about each student's current instructional needs and allows them to plan instruction accordingly, set ambitious but attainable goals for students, and monitor progress toward those goals (Fuchs & Fuchs, 2002).

FAST™ CBMreading is designed for all students in grades 1 through 12. For elementary grades 1 through 5 or 6, measures of fluency with connected text (i.e., CBM-R) are often used as universal screeners for grade-level reading proficiency. Although strong evidence exists in the literature to support the use of CBM-R (Fuchs, Fuchs, & Maxwell, 1988; Kranzler, Brownell, & Miller, 1998; Markell & Deno, 1997; Reschly et al, 2009), support for CBM-R as a universal screener for students who are not yet reading connected text is less robust (Fuchs, Fuchs, & Compton, 2004; National Research Council, 1998). Specifically, CBM-R might not be the best measure for students who cannot yet read at least 10 words in a story. For those students not yet reading connected text with fluency, FAST™ CBMreading results and scores should be interpreted with caution.

FAST^{TM} CBMreading Content Description

Reading involves simultaneous completion of various component processes. To achieve simultaneous coordination across these component processes, instantaneous execution of each component skill is required (Logan, 1997). Reading fluency is achieved so that performance is speeded, effortless, autonomous, and achieved without much conscious awareness (Logan, 1997). Oral reading fluency (ORF) represents the automatic translation of letters into coherent sound representations, unitizing those sound components into recognizable wholes, and automatically accessing lexical representations, processing meaningful connections within and between sentences, relating text meaning to prior information, and making inferences to supply missing information. Logan (1997) described oral reading fluency as the complex orchestration of these skills, establishing it as a reliable measure of reading expertise.

As previously mentioned, FAST™ CBMreading is a version of an ORF measure. FAST™ CBMreading is an effective tool used to measure rate of reading. Indeed, reading disabilities are most frequently associated with deficits in accurate and efficient word identification. Although reading is not merely rapid word identification or the "barking at words" (Samuels, 2007), the use of rate-based measures provide a general measure of reading that can alert teachers to students who have problems and are behind their peers in general reading ability. Overall, FAST™ CBMreading provides a global indicator of reading.

FAST™ CBMreading Passage Development

The FAST™ CBMreading passages have been systematically developed and field tested over several years. The goal in creating the FAST™ CBMreading measures was to systematically develop, evaluate, and finalize research-based instrumentation and procedures for reliable assessment and evaluation of reading rate. The developers sought to reduce the amount of residual error found in other oral reading fluency (ORF) assessments. Both Christ and Ardoin (2009) and Ardoin and Christ (2009) found that the error variance in the FAST™ CBMreading was less than in other published ORF forms.

The FAST™ CBMreading passages were initially developed and field tested with at least 500 students per level. All passages were designed with detailed specifications and in consultation with educators and content experts. The researchers analyzed data from three rounds of field testing and edited passages to optimize the semantic, syntactic, and cultural elements.

Specifications and Passage Construction

General passage development followed the process and standards presented by Schmeiser and Welch (2006; see Figure 1) in the fourth edition of Educational Measurement (Brennan, 2006).

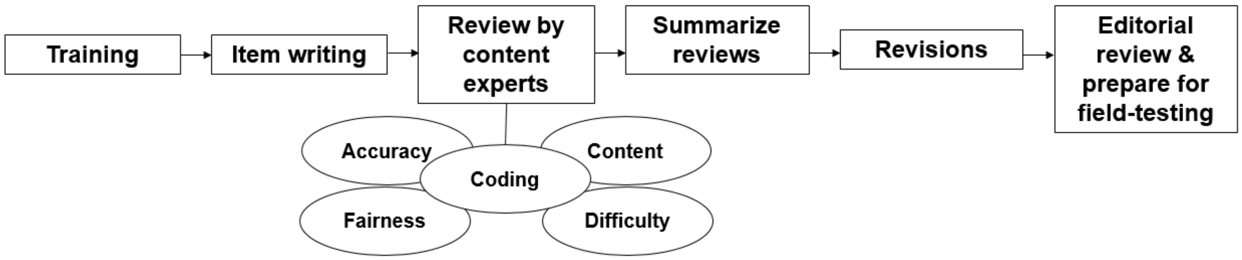

Figure 1. Formal Passage Development Process

In addition to the process and standards of developing passages presented by Schmeiser and Welch (2006) text difficulty had to be considered. To address the limitations of prior passage construction methods, relevant research in reading comprehension was taken into consideration (Graesser, Singer, & Trabasso, 1994; Kintsch & van Dijk, 1978; Kintsch, 1998; van den Broek, 1994; Zwaan & Rapp, 2008). Specifically, text type, paragraph and sentence structure, word and language usage, and cohesion were selected as criteria for development of all FAST™ CBMreading passages.

Text Type

The initial goal for developing FAST™ CBMreading materials was to develop passages that were consistent for students in grades 1-6 and provided few confounds with a student's background knowledge. Therefore, one type of text (narrative) was selected for use. In general, narratives tell a story from fictional or fact-based events (Stein & Glenn, 1975). The episodes and situations in a narrative provide the underlying structure of a story and can be reflected in episodes and situations experienced in real life. Indeed, the purpose of these stories is typically to reflect everyday experiences in most people's lives. As a result, the type of underlying cognitive processes required to understand a narrative text is more natural, easy to remember, and reflective of oral discourse as well as personal experience (Graesser, Olde, & Klettke, 2002; Rubin, 1995). To make the passages appropriate, the writers were instructed to develop a goal-action/attempt-outcome structure for each passage story. They were given instructions that their passage stories should consist of one or more characters who have a goal, act or attempt to meet that goal, and experience an outcome where their initial goal is met or not met.

Paragraph and Sentence Structure

Specific guidelines were also provided for paragraph and sentence structure. This was necessary to ensure a parallel text structure across the passages. Each writer was instructed to use three or four paragraphs within each passage and, when possible, include a main idea sentence at the beginning of each paragraph that would introduce and help organize content for the reader. Writers were also instructed not to use complex punctuation such as colons and semicolons to reflect text that is familiar to primary grade levels, as well as to encourage a more direct style of writing.

Word and Language Usage

Writers were asked to address a variety of word and language restrictions in their passage writing so that an accurate measure of reading ability would not be directly or indirectly influenced by these factors. For example, use of biased language might cause some writers to pause or have an emotional response to the passage that may alter reading outcomes. Therefore, bias-free language was required when describing people of varying age, race, appearance, education, and lifestyle. Writers were also instructed to avoid the use of technical words. Again, use of these words might cause some readers to pause or become frustrated due to lack of appropriate background knowledge and this would alter reading outcomes. When technical terms needed to be included to maintain the integrity of the story, the writer was asked to reconsider whether the story needed that information and rewrite as necessary. However, if the writer felt that the information was necessary, the writer elaborated with a sentence or part of a sentence explaining or defining the word or phrase. Use of familiar or more common words was encouraged so that word choice would not cause readers problems when moving through a passage. Writers examined relevant primary grade word lists and curriculum materials to help them identify appropriate language, structure, and word use for developing the passages for each level. The use of dialogue was not allowed because dialogue requires readers to use different types of cognitive processes to understand text meaning. Specifically, findings from research in discourse comprehension suggest that dialogue use requires the reader to develop a representation of the story that may activate information in a more complicated way compared to a story without dialogue (e.g., see Kintsch & van Dijk, 1978). To avoid differences in reading outcomes due to use of dialogue, it was avoided. Finally, writers were asked to avoid the use of clichéd words or phrases, catch phrases, or colloquialisms such as slang, jargon, and idioms. These were important types of words and language to avoid because we wanted the texts to be understood by many people and these kinds of words and language can hinder general understanding.

Cohesion

Cohesion is defined as the result of using cohesive devices in a text (e.g., features, words, phrases, or sentences) that help the reader interconnect parts of a text more easily and create a coherent representation (Graesser et al., 2004). A cohesive text is one that is easier to read and allows the reader to more readily develop an accurate, coherent representation of text meaning. Writers were to incorporate cohesive cues in the texts to eliminate potential confounds due to poor or ineffective writing; however, the use of cohesion in texts had to be varied to address the needs of different ability levels of readers within each passage set. For example, readers who have stronger reading skills than others might find extensive use of cohesive cues to be disruptive or make the text seem too simplistic whereas readers who struggle to read might find these same cues helpful (McNamara & Kintsch, 1996). Therefore, the goal was to provide an average level of cohesive cues across all texts without explicitly manipulating levels of cohesive elements

(Graesser, McNamara, & Louwerse, 2003). Writers varied their use and application of the following text elements during the writing process:

- ambiguous pronouns with nouns,
- elaborations that help connect unfamiliar and familiar concepts across the text,
- connectives to create relations between sentences or ideas,
- words to increase conceptual overlap between or across sentences,
- temporal order of events,
- explicit & implicit goals,
- causal relations across the text, and
- words that were familiar, concrete, limited to one meaning, easy to imagine, meaningful, and age-appropriate.

Multi-Grade Level Passage Sets

Although the initial goal was to develop unique sets of passage forms for each grade level, the research findings indicated that multi-grade passage sets were more appropriate and sensitive to student needs (Compton, Appleton, & Hosp, 2004; Hiebert & Fisher, 2007). An important consideration in developing the multi-level passages was the fact that reading ability can vary substantially within and across grade levels (Vellutino, 2003). For this reason, three form levels corresponding to multiple grades were developed. The resulting forms were divided into Levels A, B and C according to the passage difficulty as follows:

- Level A = Grade 1
- Level B = Grades 2-3
- Level C = Grades 4-6

There were 39 Level A passages, 60 Level B, and 60 Level C passages. The passages were further organized for use as screening and progress monitoring forms.

Item Writers

Writers included university graduate students studying English, Linguistics, and other Text or Discourse-related disciplines. Twenty writers were selected across the three different national sites. All writers attended a 2-3-hour training meeting and wrote between 10-20 passages using an iterative process in which they wrote, revised, and provided feedback to other writers. Editing took place throughout this process and a final revision/editing process resulted in the completed set of passages.

FAST™ CBMreading Administration

FAST™ CBMreading includes standardized administration and scoring procedures. Students read each passage aloud for 1 minute while the examiner records any errors in the online system. When used for screening, each student completes three passages and the median score is used as the estimate of the student's current reading skill. Including transitions between passages, screening administrations typically require about 5 minutes. When used for progress monitoring, students complete one passage and require just over one minute to complete. All administration procedures are described in the online system along with the examiner directions and scoring rules. The examiner reads the directions from the online system and records the student's errors using a computer or tablet device while the student completes the assessment. This feature alleviates the need for the examiner to tally and enter student scores after the administration.

FAST™ CBMreading Scores and Scoring

Score Types

There are three parts to each FAST™ CBMreading score: (a) total words read, (b) errors, and (c) words ready correctly (WRC). WRC is the primary metric used in reporting student performance on FAST™ CBMreading. Total words read refers to the total number of words read by the student, including correct and incorrect responses. Number of errors is the total number of errors the student made during the one-minute administration time. Words read correctly per minute (WRCPM) is the number of words read correctly per minute. This is computed by taking the total number of words read and subtracting the number of errors the student made.

In addition to the above score details, each student's accuracy when reading is important. Competent readers need to read with at least 95% accuracy to understand the text. For this reason, the FAST™ CBMreading reports include indicators of student accuracy whenever it falls below 95%.

Benchmark Scores

Benchmarks were established for FAST™ CBMreading to help teachers accurately identify students who are at risk or not at risk for reading difficulties. Benchmarks for FAST™ CBMreading were set by examining data from students who completed both the FAST™ CBMreading assessment and another "high stakes" assessment such as a state test. These data were analyzed to determine the predictive validity of FAST™ CBMreading scores in relation to student performance on the other assessment. The assessment of oral reading rate with CBM-R is well established in the literature for use to benchmark student progress (Wayman et al., 2007). Results indicate that FAST™ CBMreading is highly predictive of student's scores on other reading assessments. Specifically, FAST™ CBMreading scores that fall at or below the 15th percentile predict that a student is at high risk of not passing the second assessment. FAST™ CBMreading scores from the 16th through 39th percentiles indicate some risk of not passing the second assessment. Benchmark scores are available for FAST™ CBMreading for grades 1 through 8 and time of the year (i.e., fall, winter, spring).

Normative Scores

Normative scores for FAST™ CBMreading reflect typical performance by percentile rankings. These data characterize typical performance for each season by grade level and are intended to establish a baseline distribution for FAST™ CBMreading. FAST™ reports includes multiple norms so that teachers can compare student performance with different groups of students. Specifically, the reports include class or group, school, district, and national norm distributions. The “local” norms for the class or grade, school, and district provide educators with information about how each student performs in relation to the other students participating in the same instruction. The National norms provide information about how individual student performance compares with all the students of the same grade in the U.S. national FAST™ data system.

Score Interpretations

FAST™ CBMreading scores are interpreted in relation to expected scores for students in each grade level. In addition, FAST™ provides information about expected scores by grade level for the three benchmark screening time points: (a) fall, (b), winter, and (c)spring. Grade-level score expectations are organized into two types: benchmarks and norms.

FAST™ CBMreading Construct Validity

Content-Related Validity Evidence

The design specifications for FAST™ CBMreading relate directly to their evidence of content validity. Each passage set was designed with the intent to address specific criteria aimed to maximize both utility and sensitivity. Specific guidelines were provided for paragraph and sentence structure. This was necessary to ensure a parallel text structure across the passages. Each writer was instructed to use three or four paragraphs within each passage and, when possible, include a main idea sentence at the beginning of each paragraph that would introduce and help organize content for the reader. Writers were also instructed to not use complex punctuation such as colons and semi-colons to reflect text that is familiar to primary grade levels as well as to encourage a more direct style of writing.

The Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science and Technical Subjects (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) is a synthesis of information gathered from state departments of education, assessment developers, parents, students, educators, and other pertinent sources to prepare K-12 students for college and careers. This process is headed by the Council of Chief State School Officers and the National Governors Association. The Standards related to oral reading fluency are found within Foundational Skills in Reading. These standards are primarily relevant to K-5 children and include the working knowledge of the following subcategories:

- Print Concepts: the ability to demonstrate the organization and basic feature of print.
- Phonological Awareness: demonstrate understanding of spoken words, syllables, and sounds or phonemes.

- Phonics and Word Recognition: the skill of applying grade-level phonics and word analysis skills in decoding words.
- Fluency: Reading on-level texts with sufficient purpose, accuracy, and fluency to support comprehension.

FAST™ CBMreading is an assessment that incorporates all or the above skills and provides an indicator of a student's oral reading fluency.

Criterion-Related Validity Evidence

Predictive and concurrent criterion validity for grades 1-6 are available using several different other assessments (i.e., Test of Silent Reading Efficiency and Comprehension (TOSREC), Measures of Academic Progress (MAP), AIMSweb and DIBELS Next), providing evidence of criterion-related validity. Where applicable, the delay between FAST™ CBMreading administration and criterion administration is stated. Concurrent validity coefficients for FAST™ CBMreading grade-level passages are provided in

Table 1. Predictive validity coefficients for FAST™ CBMreading grade-level passages are provided in

Table 2. All coefficients were derived from students across three states: Georgia, Minnesota, and New York.

Table 1. Concurrent Validity Coefficients for FAST™ CBMreading

Criterion	Grade	N	Correlation
AIMSweb	1	399	.95
	2	425	.97
	3	402	.95
	4	445	.96
	5	447	.96
	6	229	.95
DIBELS Next	1	399	.95
	2	463	.92
	3	483	.96
	4	485	.95
	5	503	.95
	6	225	.95
MAP	1	--	--
	2	237	.81
	3	231	.78
	4	233	.73
	5	219	.66
	6	212	.69

TOSREC	1	218	.86
	2	246	.81
	3	233	.81
	4	228	.79
	5	244	.81
	6	222	.82

Note. N = sample size.

Table 2. Predictive Validity Coefficients for FAST™ CBMreading

Criterion	Grade	N	Time Lapse in Weeks		Correlation
			Mean	SD	
AIMSweb	1	385	18.68	3.04	.91
	2	413	18.56	2.44	.93
	3	391	18.98	2.50	.91
	4	427	19.00	2.32	.94
	5	431	19.00	2.51	.93
	6	220	17.45	0.86	.94
DIBELS Next	1	425	35.57	2.02	.82
	2	80	35.93	1.61	.74
	3	76	35.79	1.47	.91
	4	74	35.67	1.41	.90
	5	85	35.67	1.41	.93
	6	--	--	--	--
MAP	1	--	--	--	--
	2	240	35.23	1.42	.76
	3	233	35.47	1.27	.73
	4	235	35.23	0.88	.69
	5	220	35.29	1.13	.65
	6	212	35.06	0.96	.71
TOSREC	1	44	35.57	2.02	.47
	2	35	35.94	1.61	.56
	3	33	35.79	1.48	.69
	4	35	35.67	1.41	.52
	6	18	35.06	0.96	.87

Note. N = sample size; SD = standard deviation.

Reliability-Related Validity Evidence

Evidence of FAST™ CBMreading internal consistency across passages is provided in Table 3. Data were gathered across three states: Georgia, Minnesota, and New York.

Table 3. Internal Consistency Coefficients for FAST™ CBMreading Passages

				Coefficient	
Passage Level	N	Number of Passages	Number of Weeks	Range	Median
Level A	231	60	< 2	.91 - .92	.92
Level B	488	60	< 2	.89 - .91	.90
Level C	513	60	< 2	.88 - .93	.91

Note. N = sample size.

Data collected across several studies from three states (George, Minnesota, New York) are summarized in Table 4. The information represents evidence for alternate-form reliability of FAST™ CBMreading, and overall reliability of the performance level score.

Table 4. Alternate-Form Reliability Coefficients and SEM for FAST™ CBMreading

				Coefficient		
Passage Level	N	Number of Passages	Number of Weeks	Range	Median	SEM
Level A	231	39	< 2	.62 - .86	.74	5.40
Level B	488	60	< 2	.65 - .82	.75	8.54
Level C	513	60	< 2	.78 - .88	.83	10.41
Level A	231	39	< 2	.89 - .94	.92	3.03
Level B	488	60	< 2	.87 - .92	.90	4.97
Level C	513	60	< 2	.92 - .95	.94	7.06

Note: N = sample size; SEM = standard error of measurement.

Inter-rater reliability evidence from data in Georgia, Minnesota, and New York is presented in Table 5.

Table 5. Inter-Rater Reliability Coefficients for FAST™ CBMreading

		Coefficient	
Passage Level	N	Range	Median
Level A	146	0.83 - 1.00	0.97
Level B	1391	0.93 – 0.97	0.97
Level C	1345	0.83 - 1.00	0.98

Note: N = sample size.

Evidence for Use of FAST™ CBMreading as a Screening Tool

Classification accuracy refers to how well an assessment identifies students who do or do not possess the skills measured by the test. FAST™ CBMreading classification accuracy information is provided for first through sixth grades, using the Measures of Academic Progress (MAP) and Test of Silent Reading Efficiency and Comprehension (TOSREC) as the criterion measures. Measures of classification accuracy were used to determine decision thresholds using criteria related to sensitivity, specificity, and area under the curve (AUC). Specifically, specificity and sensitivity were computed at different cut scores in relation to maximum AUC values. Decisions for final benchmark percentiles were generated based on maximizing each criterion at each cut score (i.e., when the cut score maximized specificity ≥ .70, and sensitivity was also ≥ .70; see Silberglitt & Hintze, 2005). In the scenario for which a value of .70 could not be achieved for either specificity or sensitivity, precedence was given to maximizing specificity.

FAST™ CBMreading classification accuracy was initially determined based on a sample of 1,153 students in the state of Minnesota, spanning across three regions. Data were collected during the 2012-13 school year. The sample consisted of approximately 45% males and 55% females. Approximately 20% of the students involved were eligible for free and reduced lunch. Most students were White (52%). The remainder of the sample consisted of approximately 30% Hispanic, 12% Black, 4% Asian or Pacific Islander, and 1% American Indian or Alaska Native. Approximately 15% of students were receiving special education services. All participants were proficient in English. See Table 6 for classification accuracy results.

Table 6. Classification Accuracy by Grade Level for FAST™ CBMreading Passages

Grade Level	N	Cut Score	AUC	Sensitivity	Specificity	Classification	Time Lag	Criterion
				20th Percentile				
1	171	16.5	0.81	0.75	0.63	0.71	2 to 4	TOSREC
2	206	42.5	0.93	0.88	0.87	0.88	2 to 4	
3	188	75.5	0.89	0.84	0.83	0.84	2 to 4	
4	181	108.5	0.87	0.78	0.82	0.79	2 to 4	
5	202	107.5	0.90	0.84	0.79	0.83	2 to 4	
6	205	118.5	0.90	0.92	0.72	0.88	2 to 4	
1	171	17	0.77	0.63	0.82	0.74	4 mo.	MAP
2	206	57	0.82	0.63	0.85	0.76	4 mo.	
3	188	88	0.80	0.77	0.75	0.76	4 mo.	
4	181	113	0.88	0.82	0.81	0.81	4 mo.	
5	202	101	0.89	0.70	0.91	0.86	4 mo.	
6	205	126	0.89	0.83	0.82	0.82	4 mo.	
1	171	21	0.79	0.78	0.70	0.73	8 mo.	MAP
2	206	63	0.82	0.83	0.68	0.72	8 mo.	

3	188	67	0.77	0.51	0.88	0.75	8 mo.	
4	181	104	0.89	0.80	0.86	0.84	8 mo.	
5	202	97	0.89	0.74	0.92	0.88	8 mo.	
6	205	126	0.89	0.85	0.82	0.83	8 mo.	
1	171	16	0.80	0.66	0.81	0.76	~1 year	
2	206	82	0.90	0.87	0.84	0.86	~1 year	
3	188	88	0.89	0.82	0.81	0.82	~1 year	MAP
4	181	114	0.87	0.83	0.76	0.78	~1 year	
5	202	108	0.89	0.80	0.85	0.84	~1 year	
6	205	126	0.85	0.77	0.79	0.79	~1 year	
30th Percentile								
1	171	16.5	0.81	0.75	0.63	0.71	2 to 4	
2	206	44.5	0.94	0.91	0.83	0.89	2 to 4	
3	188	79.5	0.88	0.83	0.83	0.83	2 to 4	TOSREC
4	181	117.5	0.83	0.73	0.79	0.75	2 to 4	
5	202	115.5	0.87	0.79	0.77	0.79	2 to 4	
6	205	135.5	0.88	0.76	0.82	0.78	2 to 4	
1	171	31	0.78	0.84	0.57	0.74	4 mo.	
2	206	82	0.83	0.81	0.73	0.78	4 mo.	
3	188	85	0.77	0.57	0.86	0.66	4 mo.	MAP
4	181	128	0.82	0.84	0.59	0.74	4 mo.	
5	202	125	0.82	0.65	0.75	0.69	4 mo.	
6	205	144	0.82	0.76	0.75	0.75	4 mo.	
1	171	24	0.78	0.74	0.70	0.73	8 mo.	
2	206	82	0.78	0.77	0.54	0.67	8 mo.	
3	188	98	0.80	0.82	0.65	0.76	8 mo.	MAP
4	181	125	0.84	0.66	0.85	0.75	8 mo.	
5	202	128	0.86	0.80	0.73	0.76	8 mo.	
6	205	144	0.79	0.74	0.70	0.71	8 mo.	
1	171	22	0.83	0.76	0.77	0.76	~1 year	
2	206	82	0.91	0.72	0.89	0.76	~1 year	
3	188	104	0.85	0.63	0.92	0.71	~1 year	MAP
4	181	122	0.85	0.70	0.86	0.78	~1 year	
5	202	135	0.82	0.77	0.68	0.73	~1 year	
6	205	144	0.80	0.76	0.75	0.75	~1 year	

Note. N = sample size; AUC = area under the curve.

Further classification accuracy analyses were conducted using the Minnesota Comprehensive Assessment-III (MCA-III). Students were administered the MCA-III in Reading in grades 3, 4, and 5. Additionally, the same students completed three FAST™ CBMreading probes during the spring of the same school year and the median FAST™CBMreading score was compared with the MCA-III score. Only those students providing complete data were utilized in the classification accuracy analyses. ROC analysis was used to determine classification accuracy of FAST™ CBMreading probes, with spring MCA-III scale scores serving as the criterion measure. Students were disaggregated by grade level. Classification accuracy was computed for students identified as being at "High Risk" and those identified as "Somewhat at Risk" for reading difficulties using MCA-III Achievement Level Criteria (see Table 7).

Table 7. Classification Accuracy for FAST™ CBMreading and Minnesota Comprehensive Assessment-III

Grade	N	CBM-R M (SD)	MCA-III M (SD)	Correlation	Cut Score	AUC	Sensitivity	Specificity
				High Risk				
3	852	139	348 (20)	.76	132	.88	.80	.79
4	818	165	447 (15)	.71	154	.87	.80	.78
5	771	165	552 (16)	.70	152	.89	.80	.79
				Somewhat High Risk				
3	852	139	348 (20)	.76	142	.86	.78	.76
4	818	165	447 (15)	.71	165	.83	.75	.71
5	771	165	552 (16)	.70	164	.84	.77	.76

Note. N = sample size; M = Mean; SD = Standard Deviation; AUC = area under the curve.

The same classification accuracy analysis procedures were used with data from grades 3 through 5 in Georgia (Table 8).

Table 8. Classification Accuracy on FAST™ CBMreading with CRCT in Reading: Fall to Spring Prediction

Grade	N	CBM-R M (SD)	CRCT M (SD)	Correlation	Cut Score	AUC	Sensitivity	Specificity
				Some Risk				
3	329	115.96	848.65	.66	116.50	.79	.72	.71
4	320	137.64	848.18	.65	130.50	.81	.72	.73
5	353	149.27	841.22	.57	150.50	.71	.66	.66
				High Risk				
3	329	115.96	848.65	.66	80.50	.89	.82	.83
4	320	137.64	848.18	.65	100.50	.85	.83	.85
5	353	149.27	841.22	.57	128.50	.82	.79	.71

Note. N = sample size; M = Mean; SD = Standard Deviation; AUC = area under the curve.

Evidence for Use of FAST™ CBMreading as a Progress Monitoring Tool

Reliability of Slope

Traditional alternate-form reliability calculation methods may not accurately capture reliability of the slope data for the FAST™ CBMreading passages due to the small amount of variation in slope values, represented by low standard error of the estimate and standard error of the slope values. By using passage levels as groups instead of grades, FAST™ CBMreading passages may be reducing variability within grades, decreasing the reliability of slope estimates. The following analysis was conducted using HLM 7 software and used random slopes and random intercepts (see

Table 9).

Table 9. Reliability of the Slope for FAST™ CBMreading

Passage Level	Number of Weeks (range)	Number of Observations (range)	Median Coefficient
Level A	27-30	25-30	.95
	7-10	7-10	.78
Level B	27-30	25-30	.98
	7-10	7-10	.97
Level C	27-30	25-30	.98
	7-10	7-10	.97

Table 10 provides a summary for reliability of the slope by passage level. Reliability of the slope for multi-level analyses may be biased when standard error of the estimate and standard error of the slope is minimal. As noted, FAST™ CBMreading growth estimates are less prone to error than comparable progress monitoring materials. As a result, increased precision (less error) is paradoxically detrimental to multi-level reliability estimates (Raudenbush & Bryk, 2002). In such circumstances, the Spearman Brown correlation is more appropriate. The following information includes participants across three states: Georgia, Minnesota, and New York.

Table 10. Reliability of the Slope of FAST™ CBMreading using Spearman-Brown Split Half Correlation

Passage Level	N	Number of Weeks (range)	Coefficient	SEM
Level A	82	10	.71	.40
Level B	151	10 - 20	.74	.31
Level C	211	6 - 20	.65	.30
Level A	61	14 - 30	.95	.21
Level B	109	14 - 30	.70	.31
Level C	137	18 - 30	.66	.32

Note: N = sample size; SEM = standard error of measurement.

Evidence of reliability of the slope disaggregated by ethnicity in presented in Table 11. Participants were from urban, suburban, and rural areas of Minnesota. Slopes were estimated from 3 observations.

Table 11. FAST™ CBMreading Reliability of the Slope Disaggregated by Ethnicity

Grade	N (range)	Coefficient		Ethnicity
		Range	Median	
Grades 2-5	1308-1518	.25 - .43	.28	White
Grades 2-5	353-442	.32 - .60	.43	Black
Grades 2-5	197-210	.38 - .52	.40	Asian
Grades 2-5	247- 314	.21 - .52	.45	Hispanic

Note: N = sample size.

Validity of Slope

Validity of slopes for FAST™ CBMreading passages were examined using data from the AIMSweb DIBELS Next, MAP, and TOSREC assessments. Correlations between the slope and the achievement outcome are presented in Table 12. Coefficients provided in Table 12 were derived from progress monitoring data. Students were monitored with grade level passages for AIMSweb and DIBELS Next. Correlation coefficients in Table 12 may be underestimated due to differences in error between passage sets (see Ardoin & Christ, 2009 and Christ & Ardoin, 2009). The increased precision of FAST™ CBMreading passages may lead to less variable slopes compared to more error-prone progress monitoring passages. This in turn may deflate the measure of association between the two measures. Reported coefficients are partial Pearson correlation coefficients.

Table 12. Predictive Validity for the Slope for FAST™ CBMreading

Passage Level	Criterion	N	Observations (range)	Weeks (range)	Coefficient
Level A	AIMSWEB	33	10-30	30	.98
Level B		39	10-30	30	.84
Level C		70	10-30	30	.78
Level A	AIMSweb	59		10-30	.95
Level B		108		10-30	.85
Level C		116		10-30	.64
Level A	DIBELS Next	75		10-30	.76
Level B		253		10-30	.75
Level C		293		10-30	.50
Level A	DIBELS Next	57	10-30	10-30	.89
Level B		197	10-30	10-30	.82
Level C		152	10-30	10-30	.60
Level A	MAP	49	10	10	.21
Level B		71	10-20	10	.23
Level C		112	6-20	10	.21
Level A	MAP	33	14-30	30	.03

Level B		42	14-30	30	.41
Level C		78	18-30	30	.17
Level A	TOSREC Mid-Year	58			.43
Level B		98			.45
Level C		158			.36
Level A	TOSREC End of Year	58			.58
Level B		98			.22
Level C		158			.14
Level A	TOSREC	85		1-24	.46
Level B		130		1-29	.56
Level C		186		1-29	.16

Note: N = sample size.

FAST^{TM} earlyReading

FAST™ earlyReading Purpose and Use

The FAST™ earlyReading measure is designed to assess both unified and component skills associated with kindergarten and first grade reading achievement. FAST™ earlyReading is intended to enable screening and progress monitoring across four domains of reading (Concepts of Print, Phonemic Awareness, Phonics, and Decoding) and provide domain-specific assessments of these component skills as well as a general estimate of overall reading achievement. The current version of FAST™ earlyReading has an item bank with variety, including items with pictures, words, individual letters and letter sounds, sentences, paragraphs, and combinations of these elements.

FAST™ earlyReading consists of 13 different evidence-based assessments for screening and monitoring student progress. The FAST™ earlyReading subtests include: Concepts of Print, Onset Sounds, Letter Names, Letter Sounds, Word Rhyming, Word Blending, Word Segmenting, Decodable Words, Nonsense Words, Sight Words-Kindergarten (50 words), Sight Words-1st Grade (150 words), Sentence Reading and Oral Language (Sentence Repetition).

FAST™ earlyReading performance is an indicator of student reading development. It is designed to assess reading skills that predict successful broad reading proficiency. Not all FAST™ earlyReading subtests are given each screening period. Instead, there is a combination of subtests recommended for fall, winter, and spring screening aimed to optimize validity and risk evaluation. The scores from the screening subtests are used to provide a Composite score for each student. The broadest score available - and best estimate of your students' early reading skills - is the FAST™ earlyReading Composite score. The FAST™ earlyReading Composite is recommended for screening and is discussed in more detail regarding content and administration. Subtests that are not included in the composite can be used for progress monitoring or are considered supplemental. Supplemental subtests may be used to diagnose and evaluate skill deficits. Results from supplemental subtests provide guidance for instructional and intervention development.

FAST™ earlyReading is designed to accommodate quick and easy weekly assessments, which provide useful data to monitor student progress and evaluate response to instruction. The availability of multiple alternate forms for various subtests of FAST™ earlyReading make it suitable for monitoring progress between benchmark assessment intervals (i.e., fall, winter, and spring) for those students that require more frequent monitoring of progress. Onset Sounds has 13 alternate forms, and the following subtests have a total of 20 alternate forms: Letter Naming, Letter Sound, Word Blending, Word Segmenting, Decodable Words, Sight Words, and Nonsense Words. FAST™recommends Letter Sounds for monitoring progress in kindergarten and FAST™CBMreading for progress monitoring in first grade

FAST™ earlyReading Content Description

Concepts of Print

Concepts of Print is defined as the general understanding of how print works and how it can be used (Snow, Burns, & Griffin, 1998). Concepts of print is the set of skills used in the manipulation of text-based materials, which includes effective orientation of materials (directionality), page turning, identifying the beginning and ending of sentences, identifying words, as well as identifying letters, sentences, and sentence parts. Concepts of print are normally developed in the emergent literacy phase of development and enable the development of meaningful early reading skills: "Emergent literacy consists of skills, knowledge, and attitudes that are developmental precursors to conventional forms of reading and writing" (Whitehurst & Lonigan, 1998). These skills typically develop from preschool through the beginning of First Grade— with some more advanced skills that develop through second grade, such as understanding punctuation, standard spelling, reversible words, sequence, and other standard conventions of written and spoken language. Introductory level of logical and analytical abilities as in understanding the concepts of print has an impact on early student reading achievement (Adams, 1990; Clay, 1972; Downing, Ollila, & Oliver, 1975; Hardy et al., 1974; Harlin & Lipa, 1990; Johns, 1972; Johns, 1980; Lomax & McGee, 1987; Nichols et al., 2004; Tumner et al., 1988).

Phonemic Awareness

Phonemic Awareness involves the ability to identify and manipulate phonemes in spoken words (National Reading Panel [NRP], 2000). Phonemes are the smallest units of sound in spoken language. According to Adams, "to the extent that children have learned to 'hear' phonemes as individual and separable speech sounds, the system will, through the associative network, strengthen their ability to remember or 'see' individual letters and spelling patterns" (1990, p. 304). Hearing and distinguishing individual letter sounds comes last (Goswami, 2000). Children who manipulate letters as they are learning to hear specific sounds have been shown to make better progress in early reading development than those who do not (NRP, 2000). PA skills are centrally involved in decoding by processes of blending and segmenting phonemes (NRP, 2000). Phonemic awareness also helps children learn how to spell words correctly. Phonemic segmentation is required to help children retain correct spellings in memory by connecting graphemes (printed letters) to phonemes (NRP, 2000).

Phonics

Phonics is the set of skills readers use to identify and manipulate printed letters (graphemes) and sounds (phonemes). It is the correspondences between spoken and written language. This connection between letters, letter combinations, and sounds enable reading (decoding) and writing (encoding). Phonics skill development "involves learning the alphabetic system, that is, letter-sound correspondences and spelling patterns, and learning how to apply this knowledge" to reading (NRP, 2000b).

Decoding

"Decoding ability is developed through a progression of strategies sequential in nature: acquiring letter-sound knowledge, engaging in sequential decoding, decoding by recognizing word patterns, developing word accuracy in word recognition, and developing automaticity and fluency in word recognition" (Hiebert & Taylor, 2000, p. 467). When a child has a large and established visual lexicon of words in combination with effective strategies to decode unfamiliar words, he/she can read fluently— smoothly, quickly, and more efficiently (Adams, 1990; Snow et al., 1998). The reader can also focus his/her attention on monitoring comprehension: "If there are too many unknown words in the passage that require the child to apply more analytic (phonemic decoding) or guessing strategies to fill in the blanks, fluency will be impaired" (Phillips & Torgesen, 2006, p. 105). According to RAND, "readers with a slow or an inadequate mastery of word decoding may attempt to compensate by relying on meaning and context to drive comprehension, but at the cost of glossing over important details in the text" (2002, p. 104).

FAST™ earlyReading Content Development

FAST™ earlyReading development followed research recommendations for item and test development and included an iterative process of pilot testing, feedback and revisions. Subtests were created by reviewing the research literature in reading curriculum based assessment and early reading skill development. The FAST™ earlyReading subtests were created to measure each of the previously discussed components of early reading. The following is a list of all FAST™ earlyReading subtests and a brief description of the skills that they measure.

Concepts of Print. The FAST™ Concepts of Print subtest assesses a general understanding of how print is used so other reading skills can emerge. In this task, students are asked to complete basic tasks such as proper page orientation, accurate print tracking, and locating the beginning and ending of sentences.

Onset Sounds. The FAST™ Onset Sounds subtest assesses a student's ability to identify and manipulate the smallest units of sound in spoken language. Children are presented with a set of pictures and are asked to correctly identify the picture that begins with a sound or are asked to generate the initial sound for a picture.

Letter Names. The FAST™ Letter Names subtest assesses students' accuracy and automaticity naming uppercase and lowercase letters in isolation.

Letter Sounds. The FAST™ Letter Sounds subtest assesses students' ability and automaticity providing the sounds for lowercase letters in isolation.

Word Rhyming. The FAST™ Word Rhyming subtest requires students to identify pictures that rhyme with a given word or generate a rhyme for a pictured word

Word Blending. The FAST™ Word Blending subtest assesses students' ability to form a word from individually-spoken sounds or phonemes. Examiners say each phoneme in a word (e.g., /t/ /o/ /p/) and the student is expected to say the complete word (e.g., "top").

Word Segmenting. The FAST™ Word Segmenting subtest assesses students' ability to separate a spoken word into individual sounds, or phonemes. The examiner says a word and asks the student to say any sounds in the word.

Decodable Words. The FAST™ Decodable Words subtest assesses students' ability to read phonetically regular words (e.g., "pen"). More specifically, phonetically regular words have common phoneme-grapheme relationships that can be decoded. As the student becomes fluent with letter-sound correspondence, s/he will move from saying the sound of each letter to reading whole words.

Nonsense Words. The FAST™ Nonsense Words subtest assesses students' ability to read phonetically regular "words" (e.g., "vit"). It is called Nonsense Words because the "words" are decodable strings of letters that are not established words in the English language but are allowable letter sequences in English. The logic behind a nonsense word measure is that it assesses whether students can decode strings of letters and read them fluently while controlling for potential familiarity that students may have when decoding real words.

Sight Words-Kindergarten (50 words). The FAST™ Sight Words-50 subtest assesses a student's ability to recognize 50 of the most high-frequency words. The Sight Words-50 subtest is distinct from the decodable word measures; any high frequency words are not decodable (e.g., "your") and students must recognize them with automaticity rather than using decoding strategies.

Sight Words-1st Grade (150 words). The FAST™ Sight Words-150 subtest assesses a student's ability to recognize 150 of the most high-frequency English words. This subtest is similar to Sight Words-Kindergarten (50 words) but includes a more difficult range of words.

Sentence Reading. The FAST™ Sentence Reading subtest assesses students' reading rate and accuracy within connected text. This subtest is a precursor to CBMreading as students read individual sentences that are simple in structure and sentences are accompanied by pictures.

Oral Repetition. The FAST™ Oral Repetition subtest is a measure of students' receptive oral language, particularly students' knowledge of syntax (i.e., sentence structure). The examiner reads sentences out loud one at a time and asks the student to repeat each sentence verbally, word for word. The sentence structures become more complex as the student progresses on the task.

FAST™ earlyReading Administration

Administration time varies depending on which FAST™ earlyReading subtest is being administered. A timer is built into the software and is required for all subtests. For those subtests that calculate a rate-based score (i.e., number correct per minute), the default test duration is set to one minute. These subtests include Letter Names, Letter Sounds, Sight Words, Decodable Words, and Nonsense Words. For those subtests that do not calculate a rate-based score (number correct), the default test duration is set to open-ended. This includes Concepts of Print, Onset Sounds, Word Rhyming, Word Segmenting, and Word Blending subtests. FAST™ earlyReading is individually administered, and each subtest can take approximately 1 to 3 minutes to complete; administration of the composite assessments for universal screening takes approximately 5 minutes.

FAST™ earlyReading Scores and Scoring

Score Types

Each FAST™ earlyReading subtest produces a raw score. The primary score for each subtest is the number of items correct and/or the number of items correct per minute. These raw scores are used to generate percentile ranks and benchmarks.

The best estimate of students' early literacy skills is the FAST™ earlyReading composite score. The composite score consists of multiple subtest scores administered during a universal screening period. The FAST™ earlyReading composite scores were developed through regression and confirmatory factor analysis methodology as optimal predictors of spring broad reading achievement in kindergarten and first grade.

Table 13. Recommended Subtests for the FAST™ earlyReading Composite Score

Grade	Fall Composite	Winter Composite	Spring Composite
Kindergarten	Concepts of Print Onset Sounds Letter Names Letter Sounds	Onset Sounds Letter Sounds Word Segmenting Nonsense Words	Letter Sounds Word Segmenting Nonsense Words Sight Words-50
First grade	Word Segmenting Nonsense Words* Sight Words-150 Sentence Reading	Word Segmenting Nonsense Words* Sight Words-150 CBMreading	Word Segmenting Nonsense Words* Sight Words-150 CBMreading

*Decodable words can be substituted for Nonsense Words

A selected set of individual subtest scores were weighted to optimize the predictive relationship between FAST™ earlyReading and broad reading achievement scores (See Table 2 below). The weighting is specific to each season. It is important to emphasize that the weighting is influenced by the possible score range and the value of the skill. For example, Letter Sounds is an important skill with a score range of 0 to 60 or more sounds per minute. This represents a broad range of possible scores with benchmark scores that are high (e.g., benchmarks for fall, winter, and spring might be 10, 28, and 42, respectively). In contrast, Concepts of Print has a score range from 0 to 12 and benchmarks are relatively low in value (e.g., benchmarks for fall and winter might be 8 and 11, respectively). Because of both the score range *and* the relative value of Concepts of Print to overall early reading performance, the subtest score is more heavily weighted in the composite score. The weightings are depicted in Table 14. The high (H), moderate (M), and low (L) weights indicate the relative influence of a one point change in the subtest on the composite score.

The composite scores should be interpreted in conjunction with specific subtest scores. A variety of patterns might be observed. It is most common for students to perform consistently above or below benchmark on the composite and subtests; however, it is also possible to observe that a student is above benchmark on one or more subtest but below the composite benchmark. It is also possible for a student to be below benchmark on one or more subtests but above the composite benchmark. Although atypical, this phenomenon is not problematic. The recommendation is to combine the use of composite and subtest scores to optimize the decision-making process. Overall, composite scores are the best predictors of future reading success.

Table 14 Weighting Scheme for FAST™ earlyReading Composite Score

	Kindergarten			First Grade		
Subtest	Fall	Winter	Spring	Fall	Winter	Spring
Concepts of Print	H					
Onset Sounds	M	H				
Letter Names	L					
Letter Sounds	L	L	L			
Word Segmenting		L	M	L	L	L
Nonsense/Decodable Words		M	M	H	H	H
Sight Words			L	M	M	M
Sentence Reading				L		
CBMreading					L	L

Note. H - high weighting, M - moderate weighting, L - low weighting.

Benchmark Scores

Benchmark scores are available for each FAST™ earlyReading subtests and composite for the specific grade level and season for which they are intended for use (i.e., fall, winter, spring). Thus, a benchmark is purposefully not provided for every subtest, for each season. Benchmarks were established for FAST™ earlyReading to help teachers accurately identify students who are at risk or not at risk for academic failure. These benchmarks were developed from a criterion study examining FAST™ earlyReading assessment scores in relation to scores on the Group Reading Assessment and Classification Evaluation (GRADE; Williams, 2001). Measures of classification accuracy were used to determine decision thresholds using criteria related to sensitivity, specificity, and area under the curve (AUC). Specificity and sensitivity was computed at different cut scores in relation to maximum AUC values. Decisions for final benchmark percentiles were generated based on maximizing each criterion at each cut score (i.e., when the cut score maximized specificity ≥ .70, and sensitivity was also ≥ .70; see Silberglitt & Hintze, 2005). Based on these analyses, the values at the 40th and 15th percentiles were identified as the primary and secondary benchmarks for FAST™ earlyReading, respectively.

Normative Scores

Normative scores for FAST™ earlyReading reflect typical performance by percentile range. These data characterize typical performance for each subtest and composite, by season. FAST™ earlyReading measures have been normed on separate samples for kindergarten and first grade. FAST™ earlyReading reports include normative data compared to the group (e.g., class), school, district and national distributions. These data characterize typical performance for each grade level, by season.

FAST™ earlyReading Construct Validity

Content-Related Validity Evidence

The test specifications for FAST™ earlyReading subtests relate directly to their evidence of content validity. Each subtest was designed with the intent to address specific criteria aimed to maximize both utility and sensitivity. The Common Core State Standards for Reading were established in 2010. The standards alignment with FAST™ earlyReading subtests as well as the area of reading addressed are presented in the table below.

Table 15 Alignment of CCSS and FAST™ earlyReading Subtests

Subtest	Common Core State Standards	Reading Skill
Concepts of Print*	RF.K1, RF.K.1.a, RF.K.1.b, RF.K.1.c, RF.1.1, F.1.1.a	Concepts of Print
Letter Names*	RF.K.1.d	Alphabetic Principle (Phonics)
Letter Sounds*	RF.K.3.a	Alphabetic Principle (Phonics)
Decodable Words	R.F.K.3, RF.1.3, RF.1.3.b, RF.2.3, RF.3.3	Alphabetic Principle (Phonics)
Nonsense Words*	R.F.K.3, RF.1.3, RF.1.3.b, RF.2.3, RF.3.3	Alphabetic Principle (Phonics)

Sight Words (50)* Sight Words (150)*	RF.K.3.c, RF.1.3.g, R.2.3.f, RF.3.3.d	Fluency
Sentence Reading* (CBM W, S)	RF.K.4, RF.1.4, RF.1.4.b, RF.2.4, RF.2.4.b, RF.3.4	Fluency
Onset Sounds	RF.K.2.c, RF.K.2.D, RF.1.2.c	Phonemic Awareness
Rhyming	RF.K.2.a	Phonemic Awareness
Word Blending	RF.K.2.b, RF.K.2.c, RF.1.2.b	Phonemic Awareness
Word Segmenting*	RF.K.2.b, RF.K.2.d, RF.1.2.c, RF.1.2.d	Phonemic Awareness
Oral Repetition	SL.K.6, SL.1.6	Phonemic Awareness

Criterion-Related Validity Evidence

Criterion-related validity of FAST™ earlyReading subtests was examined using the Group Reading Assessment Classification Evaluaiton (GRADE). The GRADE is an untimed, group-administered, norm-referenced reading achievement test that is intended for children in preschool through grade 12. Comprised of 16 subtests categorized within five components, the GRADE utilizes subtest scores, depending on the testing level, to form the Total Test composite score. Evidence for the validity of FAST™ earlyReading is presented below on the external criterion measure of the GRADE Total Test composite score.

To establish criterion-related validity, students were recruited from school districts. In School District 1, three elementary schools participated. Kindergarten students from District 1 who participated in the study were enrolled in all-day or half-day Kindergarten. Most students within the school district were White (78%), with the remaining students identified as either Black (19%), or other (3%). Forty to fifty percent of students at each school were eligible for free and reduced lunch. In school District 2, most students within the school district were White (53%), with the remaining students identified as Black (26%), Hispanic (11%), Asian (8%), or other (2%). Forty to fifty percent of students at each school are on free and reduced lunch.

A summary of concurrent and predictive validity coefficients for the GRADE are presented (Table 4). The range is indicated of combinations of concurrent and predictive analyses across the school year (e.g., fall to fall, fall to spring, winter to spring…); not all season combinations were available for each subtest. As discussed previously, the composite score in kindergarten and first grade demonstrated the highest level of criterion validity; suggesting that it is the best estimate of current and later broad reading performance.

Table 16 Concurrent and Predictive Validity for FAST™ earlyReading

Subtest	Grade	N (range)	Coefficient (range)	Coefficient Median
Composite	K	173	.67 - .69	.68
	1	100	.72 - .83	.81
Onset Sounds	K	85 – 230	.03- .62	.58
Letter Names	K	85 – 230	.18 - .63	.44
Letter Sounds	K	85 – 230	.19 - .63	.49
Word Blending	K	213 – 230	.23 - .66	.41
	1	71 – 179	.12 - .56	.38
Word Segmenting	K	213 – 228	.25 - .58	.42
	1	71 – 179	.07 - .60	.41
Decodable Words	K	214	--	.27
	1	71 – 179	.22 - .78	.53
Sight Words-50	K	213	--	.19
Sight Words-150	1	71 – 179	.43 - .80	.66
Nonsense Words	K	105 – 215	.27 - .44	.36
	1	168 – 179	.43 - .67	.60

.

Reliability-Related Validity Evidence

Some FAST™ earlyReading subtests have fixed test lengths and are subject to typical internal consistency analyses. Some FAST™ earlyReading subtests, however, are timed. Internal consistency measures of reliability are inflated on timed measures because of the high percentage of incomplete items at the end of the assessment, which are those for which examinees did not respond (Crocker & Algina, 1986). As a solution to both illustrate the potential inflation and reduce it, estimates of internal consistency were run on the items completed by approximately 16% of students, the items completed by 50% of students, and items completed by approximately 84% of students. Items not completed were coded as incorrect. For both fixed test-length and inconsistent test-length analyses, data were derived from a random sample of students from the FAST™ database from the 2012-13 academic year. Reliability of measures with variable test length is reported in Table 17.

Table 17 Internal Consistency for FAST™ earlyReading Subtests

			Alpha		Split-Half	
Subtest	Grade	N	Range	Median	Range	Median
Concepts of Print	K	336	--	.75	--	.76
Onset Sounds	K	597	--	.87	--	.91
Letter Names	K	444	.95 - .98	.98	.96 - .99	.99
Letter Sounds	K	683	.93 - .98	.98	.93 - .99	.98
Word Blending	K-1	480	--	.90	--	.91
Word Segmenting	K-1	500	--	.95	--	.96
Rhyming	K	586	--	.94	--	.91
Decodable Words	K-1	434	.76 - .98	.95	.75 - .98	.96
Sight Words-50	K-1	505	.90 - .99	.97	.91 - .99	.98
Sight Words-150	1	678	.90 - .99	.99	.91 - .99	.99
Nonsense Words	K-1	501	.74-.96	.93	.73-.98	.95

Note. Data not available for sentence repetition or oral repetition

Test-retest reliability is a measure of the degree to which scores are stable across a short time period when the items, students, and testing conditions are constant. In fall 2012, data were collected to determine test-retest reliability for all FAST™ earlyReading screening measures. Participants included 85 kindergarten and 71 girst grade students from two elementary schools in a metropolitan area in the Midwest. Kindergarten students who participated in the study were enrolled in all-day kindergarten at two elementary schools within the same school district.

All first grade students who participated in the study were from a single school. Most students within the school district were White (78%), with the remaining students identified as either Black (19%), or other (3%). Forty to fifty percent of students at each school were on free and reduced lunch. Teachers randomly selected three to five students and sent home passive consent forms. The second administration took place two to three weeks after the termination of the initial screening period. Test-retest reliabilities are reported in .

Table 18.

Table 18 Test-Retest Reliability for FAST™ earlyReading

Subtest	Grade	N	Coefficient
Concepts of Print	K	39	.42
Onset Sounds	K	67	.79
Letter Names	K	45	.94
Letter Sounds	K	75	.92

Word Blending	K	70	.73
	1	67	.77
Word Segmenting	K	37	.86
	1	77	.83
Rhyming	K	39	.74
Decodable Words	K	29	.98
	1	73	.97
Nonsense Words	K	27	.94
	1	.64	.76
Sight Words-50	K	.34	.97
Sight Words-150	1	74	.94
Sentence Reading	1	37	.98
Composite	1	33	.97

Inter-rater reliability is a measure of the extent to which student scores are consistent across different examiners or scorers. FAST™ earlyReading subtests involve a small degree of subjectivity, given clear scoring guidelines and software-assisted scoring mechanisms. Unreliable scoring regarding FAST™ earlyReading may be the result of clerical errors or differences in the interpretation of a student's response. Evidence of inter-rater reliability is provided in Table 19. All coefficients represent Pearson product-moment correlation coefficients.

FAST™ earlyReading subtests were administered to students in nine elementary schools within three school districts in a metropolitan area in the Midwest. Students were administered five randomly selected progress monitoring forms. District A was about 56% White, 14% Black, 10% Hispanic, and 19% Asian/Pacific Islander. About 45% of students were eligible to receive free/reduced lunch and 13% were eligible for special education services. District B was about 93% White, 4% Black, 3% Hispanic, and 4% Asian/Pacific Islander. About 17% of students were eligible to receive free/reduced lunch and 10% were eligible for special education services. District C was about 80% White, 7% Black, 5% Hispanic, and 11% Asian/Pacific Islander. About 45% of students were eligible to receive free/reduced lunch and 10% were eligible for special education services.

Table 19 Inter-Rater Reliability by FAST™ earlyReading Subtests

Subtest	Grade	Correlation	N
Onset Sounds	K	.98	40
Letter Sounds	K	.99	47
Letter Names	K	.99	69
Word Blending	K	.98	95

Word Blending	1	.89	159
Word Segmenting	K	.85	90
Word Segmenting	1	.83	85
Nonsense Words	1	.99	51
Decodable Words	1	.99	120
Sight Words (50)	K	.99	9
Sight Words (150)	1	.97	125

Evidence of reliability is also available for alternate forms for all FAST™ earlyReading subtests (see Table 20). To effectively examine reliability coefficients, standard errors of measurement (SEM) have also been provided. The SEM is an index of measurement error representing the standard deviation of errors attributable to sampling.

Table 20 Alternate Form Reliability for FAST™ earlyReading

		Coefficient		
Grade	N (range)	Range	Median	SEM (SD)
Kindergarten				
Onset Sounds	25-29	.77-.89	.83	0.99 (.86)
Letter Naming	36-37	.82-.92	.88	5.07 (3.77)
Letter Sounds	34-36	.85-.94	.89	5.56 (4.89)
Word Blending	36-37	.59-.79	.71	0.97 (.82)
Word Segmenting	37-38	.68-.92	.82	8.07 (6.21)
Decodable Words	29	.96-.98	.97	2.93 (2.71)
Nonsense Words	28	.86-.96	.93	2.15 (1.91)
Sight Words (50)	24-28	.94-.99	.97	4.40 (4.13)
First Grade				
Word Blending	30-31	.15-.59	.26	
Word Segmenting	40	.67-.87	.82	9.83
Decodable Words	36-37	.97-.98	.98	2.98
Nonsense Words	26-27	.69-.96	.85	3.05 (3.04)
Sight Words (150)	37	.91-.96	.94	4.14

Note. N = sample size; SEM = standard error of measurement; SD = standard deviation.

To determine parallel form construction, one-way, within-subjects ANOVAs were also conducted to compare 5 randomly selected alternate forms for each individual subtest. There was not a significant effect ($p < .05$) for Onset Sounds [F (1,109) = 1.81, p =.18], Letter Names [F (1,146) = .71, p=.40], Letter Sounds [F (1,139) = .96, p =.33], Word Blending [F (1,121) = 1.60, p =.21], Word Segmenting [kindergarten = F (1,150) = 3.24, p=.07; first grade = F (1,121) = 1.60, p =.21], Decodable Words [F (1,145) = 1.72, p =.19], and Nonsense Words [kindergarten = F (1,107) =

.03, *p* =.86; first grade = *F* (1,106) = 2.34, *p* =.13]. This indicates that across all subtests, different forms did not result in significantly different mean estimates of correct responses.

Evidence for Use of FAST™ earlyReading as a Screening Tool

FAST™ earlyReading classification accuracy information is provided for both kindergarten and first grade, using the Group Reading Assessment Classification Evaluation (GRADE) as a criterion measure. Measures of classification accuracy were used to determine decision thresholds using criteria related to sensitivity, specificity, and area under the curve (AUC). Specifically, specificity and sensitivity were computed at different cut scores in relation to maximum AUC values. Decisions for final benchmark percentiles were generated based on maximizing each criterion at each cut score (i.e., when the cut score maximized specificity ≥ .70, and sensitivity was also ≥ .70; see Silberglitt & Hintze, 2005). In the scenario for which a value of .70 could not be achieved for either specificity or sensitivity, precedence was given to maximizing specificity.

Based on these analyses, the values at the 40th and 15th percentiles were identified as the benchmarks for FAST™ earlyReading. These values thus correspond with a prediction of performance at the 40th and 15th percentiles on the GRADE. Performance above the primary benchmark indicates the student is at low risk for long-term reading difficulties. Performance between the primary and secondary benchmarks indicates the student is at some risk for long-term reading difficulties. Performance below the secondary benchmark indicates the student is at high risk for long-term reading difficulties. These risk levels help teachers accurately monitor student progress using the FAST™ earlyReading measures.

Table 21 First Grade Classification Accuracy for FAST™ earlyReading Composite

Grade	AUC	Sensitivity	Specificity	Classification
	High Risk – Below 15th percentile			
Kindergarten				
Fall	.91	.88	.84	.84
Winter	.91	.94	.72	.77

Spring	.95	.75	.74	.74
First grade				
Fall	.98	1.0	.93	.93
Winter	.98	1.0	.82	.83
Spring	.99	.89	.90	.90
	Some Risk – Below 40th percentile			
Kindergarten				
Fall	.84	.80	.77	.78
Winter	.85	.84	.72	.75
Spring	.81	.75	.74	.74
First grade				
Fall	.93	.76	.84	.83
Winter	.97	1.0	.77	.81
Spring	.97	.92	.92	.92

Note. AUC = area under the curve.

Evidence for Use of FAST™ earlyReading as a Progress Monitoring Tool

Reliability of Slope

Data collected during a normative information-aimed study were used to determine reliability of the slope for FAST™ earlyReading subtests. Participants included kindergarten and first grade students from various elementary schools. Students were administered one or more FAST™ earlyReading subtests at three time points throughout the school year (i.e., fall, winter, spring). The results are presented in Table 22, and disaggregated by ethnicity in Table 23.

Table 22 Reliability of the Slope for FAST™ earlyReading Subtests

Subtest	Grade	N	Coefficient
Onset Sounds	K	2129	.91
Letter Names	K	1627	.81
Letter Sounds	K	2229	.88
Rhyming	K	904	.38

Word Blending	K	958	.73
Word Blending	1	824	.77
Word Segmenting	K	235	.60
Word Segmenting	1	824	.78
Decodable Words	K	52	.59
Decodable Words	1	918	.86
Sight Words (50)	K	167	.22
Sight Words (150)	1	624	.77
Nonsense Words	K	116	.75
Nonsense Words	1	664	.87

Note. N = sample size.

Table 23 Reliability of the Slope for FAST™ earlyReading Subtests Disaggregated by Ethnicity

Subtest	Grade	N	Coefficient	Ethnicity
Onset Sounds	K	342	.90	Black
		253	.89	Hispanic
		1253	.92	White
Letter Sounds	K	366	.93	Black
		247	.86	Hispanic
		1332	.89	White
Letter Names	K	256	.80	Black
		177	.76	Hispanic
		1049	.83	White
Nonsense Words	K	22	.70	Black
		89	.81	White
Word Blending	K	206	.77	Black
		125	.57	Hispanic
		515	.74	White
	1	156	.93	Black
		123	.74	Hispanic
		420	.77	White
Word Segmenting	K	156	.78	Black
		122	.77	Hispanic
		418	.73	White
	1	48	.60	Black
		15	.36	Hispanic
		157	.65	White
Nonsense Words	1	153	.93	Black
		92	.89	Hispanic
		328	.85	White

Decodable Words	1	199 136 449	.88 .91 .83	Black Hispanic White
Sight Words (150)	1	130 103 303	.71 .85 .79	Black Hispanic White

Note. N = sample size.

Validity of Slope

Validity of FAST™ earlyReading subtests were examined using the GRADE. The table below presents the correlation between the slope of performance using screening data (i.e., students were assessed three times per year, fall, winter and spring) and performance on the GRADE. All correlations account for initial level of performance.

Table 24 Predictive Validity of the Slope for All FAST™ earlyReading Subtest

Subtest	Grade	Criterion	N	Coefficient
Onset Sounds	K	$GRADE^K$	217	.29
Letter Names	K	$GRADE^K$	231	.44
Letter Sounds	K	$GRADE^K$	231	.54
Word Blending	K	$GRADE^K$	230	.48
Word Blending	1	$GRADE^1$	178	.16
Word Segmenting	K	$GRADE^K$	224	.49
Word Segmenting	1	$GRADE^1$	178	.23
Decodable Words	1	$GRADE^1$	179	.62
Sight Words (150)	1	$GRADE^1$	180	.59
Nonsense Words	1	$GRADE^1$	174	.61

Note. All coefficients were determined using the composite of the GRADE. Level is indicated in superscript.

FAST™ Adaptive Reading (FAST™ aReading)

FAST™ aReading Purpose and Use

The FAST™ Adaptive Reading (FAST™ aReading) assessment is a computer-adaptive measure of broad reading ability that is individualized for each student. FAST™ aReading provides a useful estimate of broad reading achievement from kindergarten through twelfth grade. The question-and-response format used in FAST™ aReading is multiple-choice, like many statewide, standardized assessments. Browser-based software adapts and individualizes the assessment for each child so that it functions at the child's developmental and skill level. The adaptive nature of the test makes it more efficient and more precise than paper-and-pencil assessments.

The design of FAST™ aReading has a strong foundation in both research and theory. During the early phases of student reading development, the component processes of reading are most predictive of future reading success (Stanovich, 1981, 1984, 1990; Vellutino & Scanlon, 1987, 1991; Vellutino, Scanlon, Small, & Tanzman, 1991). Indeed, reading disabilities are most frequently associated with deficits in accurate and efficient word identification. Those skills are necessary but not sufficient for reading to occur. After all, reading is comprehending and acquiring information through print. It is not merely rapid word identification or the "barking at words" (Samuels, 2007). As such, a unified reading construct is necessary to enhance the validity of reading assessment and inform balanced instruction throughout the elementary grades. FAST™ aReading was developed based on a skills hierarchy and unified reading construct (presented later in the technical manual).

FAST™ aReading assessment is individualized by the software and, as a result, the information and precision of measurement is optimized regardless of whether a student functions at, above, or significantly below grade level. As such, FAST™ aReading provides a useful estimate of broad reading achievement from kindergarten through twelfth grade. aReading is designed for universal screening to identify students at risk for academic delays and to differentiate instruction for all students.

FAST™ aReading Content Description

Concepts of Print

Concepts of print include the skills necessary for the general understanding of how print works and how it can be used (Snow, Burns, & Griffin, 1998). Together, concepts of print compose a set of skills used in the manipulation of text-based materials, which include accurate orientating of materials (directionality), page turning, as well as identifying the beginning and ending of sentences, words, letters, sentences, and sentence parts.

Phonological Awareness

Phonological awareness is a broad term involving the ability to detect and manipulate the sound structure of a language at the level of phonemes (i.e., smallest units of sound in spoken language), onset-rimes, syllables, and rhymes. It is used to refer to spoken language rather than letter-sound relationships, which are the focus of phonics. Most students, especially in preschool, kindergarten, and first grade, benefit from systematic and explicit instruction in this area (Adams, 1990; Carnine et al., 2009; NRP, 2000; Rayner et al., 2012; Snow, et al., 1998).

Phonemic awareness is a component of phonological awareness, and refers to the ability to know, think about, and use phonemes—individual sounds in spoken words. It is a specific type of phonological skill dealing with individual speech sounds that has been studied extensively and predicts success in reading development in languages that use alphabetic writing systems (Adams, 1990; NRP, 2000; Rayner, et al., 2012).

Phonics

Phonics is the mapping of the sounds in language to the symbols that represent them. For FAST™ aReading, we operationalize phonics as skills associated with the awareness and use of letter-sound (i.e., grapheme-phoneme) correspondence in relation to the development of successful reading and spelling using the language's orthography (e.g., alphabet). Assessment and instruction of phonics explores how these skills are applied to decode (read) and encode (spell/write) the language (NRP, 2000).

Orthography and Morphology

Measures of orthography and morphology assess readers' ability to recognize and decode or decipher words in isolation and during reading. The ability to quickly recognize words and access their meanings allows readers to focus their limited cognitive resources on meaning instead of decoding (e.g., Bear, Invernizzi, Templeton, & Johnston, 2012). These skills contribute substantively to vocabulary and reading comprehension development, therefore assessing students in these areas allows educators to determine if a student can accurately use and apply these skills.

Vocabulary

The assessment of vocabulary focuses on assessing word knowledge and vocabulary outlined in the state and national standards and based on relevant reading research for K-12 readers, including understanding and recognition of words in context that are appropriate for students at grade-level as well as appropriate for mature readers and writers to convey concepts, ideas, actions, and feelings (NAEP, 2011). These words include academic and content-specific words, word categories, word relations, and different parts of speech. The goal of vocabulary assessment should be to measure word knowledge in context rather than in isolation due to the integrated nature of reading comprehension in relation to vocabulary development.

Comprehension

Comprehension is the process of understanding what is heard and read. Comprehension, or constructing meaning, is the purpose of reading and listening. The NRP noted that "Comprehension has come to be viewed as the 'essence of reading' (Durkin, 1993), essential not only to academic learning but to lifelong learning as well" (NRP, 2000, p. 4-11). The assessment of reading comprehension in FAST™ aReading focuses on comprehension processes outlined in the state and national standards as well as relevant reading research for K-12 readers. FAST™ aReading includes items that evaluates the reader's development of an organized, coherent, and integrated representation of knowledge and ideas in the text. In addition, the items incorporate the use of inferential processes and identification of key ideas and details in the text as well as understanding its craft and structure.

FAST™ aReading Item Development

FAST™ aReading item development followed the process and standards presented by Schmeiser and Welch (2006) in the fourth edition of Educational Measurement (Brennan, 2006). Research assistants, teachers from each grade level (first through twelfth), and content experts in reading served as both item writers and reviewers for those items at the kindergarten through fifth grade level. Items for grades 6 through 12 were constructed to reflect the Common Core State Standards' (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) specifications for various skills of interest, as well as the National Assessment of Educational Progress' (NAEP, 2011) guidelines for reading assessment items. After items were written at all grade levels, they were reviewed for feasibility, construct relevance, and content balance. A stratified procedure was used to recruit a diverse set of item writers from urban, suburban and rural areas. The item writers wrote, reviewed, and edited assessment materials.

FAST™ aReading Computer Adaptive Test Development

There are three primary questions researchers should answer when developing a computer-adaptive test (CAT): (a) how the test is started; (b) how the system selects items; and (c) how does the test end (Nydick & Weiss, 2009). The purpose of this section of the manual is to detail how the termination criteria for FAST™ aReading were developed.

Traditional measures of reliability are not used when describing computer-adaptive tests. Rather, the level of precision across ability levels, and the number of items that must be administered to achieve that level are emphasized. Therefore, the first round of analysis was a process of estimating how many items had to be administered to attain an acceptable level of precision. The analysis to determine the optimum number of items per FAST™ aReading administration was conducted in multiple steps. First, a hybrid simulation was performed to generate responses for every participant. Next, cases were randomly selected and split into two groups. Then, simulations were conducted with the CATSim program to derive ability estimates and standard error of measurement (SEM) estimates for each participant for five conditions. Those conditions included a 20, 25, 30, and 40 item test as well as an administration for which every question in the item bank was administered to the hypothetical respondent.

The hybrid simulation was conducted using estimated item parameters and conducted in four steps (Nydick & Weiss, 2009) using CATsim: Use already calibrated items and item parameters; Use an already scored sparse response data matrix; Estimate ability of each examinee with the item parameters and responses; Use the ability estimates and parameters to impute missing data by simulating responses to items that examinees did not complete—using the same model with which the initial items were calibrated.

After the hybrid simulation, simulated participants were divided into two groups. CAT simulations with different test length termination criteria (20, 25, 30 and 40 items) were conducted on both groups. Mean ability and standard error estimates were calculated for each administration. In addition, a quadratic function was fitted for each simulation and plotted. Figures were generated that presented the average SEM across the range of FAST™ aReading scores for each group. The mean and standard deviation of FAST™ aReading Scores and SEM for each fixed-length CAT administration are presented in Table 25.

Table 25. Results of the FAST™ aReading CAT Simulation

Items	Group 1 (N=3,520)				Group 2 (N=3,519)			
	Score M	Score SD	SEM M	SEM SD	Score M	Score SD	SEM M	SEM SD
20	448	63	13	20	448	63	13	20
25	448	63	10	20	448	63	13	20
30	448	63	10	20	448	63	10	20
40	448	63	10	13	448	63	10	17
All	448	60	7	10	448	63	7	13

Note: N = sample size; M = mean; SD = standard deviation; SEM = standard error of measurement.

We estimated levels (from fitted quadratic models) of conditional standard error across ability levels for each group and each CAT simulation. Results indicate that different length FAST™ aReading CATs are similarly efficient and precise at the various test lengths. As scores deviate farther from the mean, estimates were less precise.

Next, the researchers of the FAST™ aReading project were interested in determining what the implications were when an FAST™ aReading test terminated sooner than 30 items. To do this, pre-existing data on FAST™ aReading from grade 1-5 administrations were analyzed under several different conditions. Ability estimates were analyzed as if the test was administered as usual, as well as when that same test was terminated after: (a) 10, (b) 15, (c) 20, and (d) 25 items.

The effects of manipulating test length on mean grade level FAST™ aReading scaled scores are presented in Table 26. The lowest variance was observed with the 30-item format.

Table 26. Change in FAST™ aReading Scaled Scores as a Function of Termination Criteria

	30 Items			25 Items		20 Items		15 Items		10 Items	
Grade	N	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	533	437	33	444	35	445	37	446	40	452	48
2	557	479	30	481	30	482	30	483	32	486	36
3	688	498	31	500	31	501	31	541	33	502	35
4	694	511	34	513	34	514	34	559	35	516	37
5	688	518	35	520	35	521	36	568	38	522	39

Note: N = sample size; SD = standard deviation.

The mean SEM for different test lengths of FAST™ aReading scaled scores are presented in Table 27. Ideal levels of SEM approximate .20, which translates to a value of 3 on the *original* FAST™ aReading scale. Like scaled scores, SEM estimates inflate as the number of items decrease. Considering the evidence from this table reaffirms the choice of a 30-item fixed test for FAST™ aReading

Table 27. Change in SEm as a Function of Terminating FAST™ aReading Tests

	30 Items			25 Items		20 Items		15 Items		10 Items	
Grade	N	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	533	10	2	11	2	12	3	15	3	18	4
2	557	10	1	11	1	13	2	15	3	19	4
3	688	11	1	12	1	13	1	15	2	19	3
4	694	11	1	12	1	14	1	16	1	20	2
5	688	11	1	12	1	14	1	16	2	20	3

Note: N = sample size; SD = standard deviation.

FAST™ aReading Administration

FAST™ aReading can be group administered in a classroom or computer lab setting, or a student can complete an administration individually with a computer or tablet device. The FAST™ aReading assessment terminates on its own, informing students they have completed all items. A typical FAST™ aReading administration is approximately 30 items. Students in grades K-5 take an average of 10-15 minutes to complete an assessment, and students in grades 6-12 take an average of 20-30 minutes. Administration time varies by student. Instructions for completing FAST™ aReading are provided via headphones to students. Before starting the test, students hear audible instructions via headphones and complete a practice item. General instructions about the aReading test that teachers can read aloud to students before they start are available in the FastBridge system.

Within an item-response theory based computer-adaptive test (CAT), items are selected based on the student's performance on all previously administered items. As a student answers each item, the item is scored in real time, and his or her ability is estimated. When a CAT is first administered, items are selected via a "step rule" (Weiss, 2004). That is, if a student answers an initial item correctly, his or her ability estimate increases by some value (e.g., .50). Conversely, if an item is answered incorrectly, the student's ability estimate decreases by that same amount. As testing continues, the student's ability is re-estimated, typically via maximum likelihood estimation. After an item is administered and scored, the student's ability is re-estimated and used to select the subsequent item. Items that provide the most information at that ability level that have not yet been administered are selected for the student to complete. The test is terminated after a specific number of items have been administered or after a certain level of precision is achieved. Subsequent administrations begin at the previous ability estimate and only present items that have not been previously administered to that student. Research using simulation methods and live data collections has been performed on FAST™ aReading to optimize the length of administrations, the level of the initial step size, and item selection algorithms to maximize the efficiency and psychometric properties of the assessment.

FAST™ aReading Scores and Scoring

Score Types

Scores generated by the FAST™ aReading computer-adaptive test yield scores based on an IRT logit scale. Because this type of scale is not often used in schools, the aReading IRT logit scale was converted to a scale like other educational measures. Such scales are arbitrarily created with predetermined basal and ceiling scores. Scores were scaled with a lower bound of 350 and a higher bound of 650. The mean value is 500 and the standard deviation is 50.

Benchmark Scores

Benchmark scores for FAST™ aReading are available for kindergarten through twelfth grade at three time points: fall, winter, and spring. Benchmarks were established for FAST™ aReading to help teachers accurately identify students who are at risk for not meeting the current grade level expectations as measured by future performance on important tests such as the state assessment.

Normative Scores

Normative scores for FAST™ aReading reflect typical performance by percentile range. FASTBridge reports include normative data compared to the group (e.g., class), school, district and national distributions. These data characterize typical performance for each grade level, by season.

Score Interpretations

To makes easy to interpret across all grades, FAST™ aReading scaled scores have an average of 500 and standard deviation of 50 across the range of kindergarten to twelfth grades. Scores should be interpreted regarding the published FAST™Bridge benchmarks and norms. In addition, FAST™ aReading has descriptions regarding the interpretation of a student's scaled score with respect to mastered, developing, and future skill development. These are intended to help teachers better understand the developmental progression and student needs. The skills listed as mastered, developing, or future for each student are based on the student's total aReading score.

FAST™ aReading Construct Validity

Content-Related Validity Evidence

FAST™ aReading was initially developed with a robust basis in reading research and theory. Items were created and revised by reading teachers and experts. Factor analysis of preliminary data provided evidence for a large primary factor (i.e., unidimensionality) and several smaller factors. Thus, FAST™ aReading is designed to provide both a unified and a component assessment of these dimensions, specifically focusing on five main areas (as put forth by the NRP, 2000a): (a) Concepts of Print, (b) Phonological Awareness, (c) Phonics, (d) Vocabulary, and (d) Comprehension.

The Common Core State Standards for English Language Arts and Literacy (Standards) were established in 2010. The cross walk between CCSS and FAST™ aReading domains is presented in Table 28. All items in the item bank are coded and aligned with specific domains, standards, and sub-standards by grade level.

Table 28. Cross-Referencing CCSS Domains and FAST™ aReading Domains

Print Concepts	Concepts of Print
Phonological Awareness	Phonemic Awareness
Phonetic Awareness	Phonetic Awareness
Vocabulary	Vocabulary
College and Career Readiness Reading Standards for Reading Literature and Informational Text	
Key Ideas and Details	Comprehension
Craft and structure	Comprehension & Vocabulary
Integration of Knowledge and Ideas	Comprehension & Vocabulary

Criterion-Related Validity Evidence

Criterion-related validity of FAST™ aReading was examined using the Gates MacGinitie Reading Tests-4th Edition (GMRT-4th; MacGinitie, MacGinitie, Maria, & Dreyer, 2000). The GMRT-4th is a norm-referenced, group administered measure of reading achievement. It is designed to provide guidance in planning instruction and intervention and is typically used as a classification tool for general reading achievement. The GMRT-4th was normed with students in the pre-reading stages through high school levels. The GMRT-4th was selected because of its strong criterion validity.

Five trained FAST™ aReading project team data collectors administered the GMRT-4th during February of 2011 at two separate schools. Students were administered the word decoding/vocabulary and comprehension subtests of the GMRT-4th during two separate testing sessions. Some students were administered the word decoding/vocabulary section first while other students were administered the comprehension subtest first. Participants included students in first through fifth grades. Three classrooms per grade at School A participated (n = 622); all students in first through fifth grades at School B participated (n = 760). See Table 29 for demographic information, disaggregated by school and Table 30for the validity coefficients.

Table 29. Demographics for Criterion-Related Validity Sample FAST™ aReading

Demographic Category	School A	School B
American Indian/Alaskan Native	3%	1%
Asian/Pacific Islander	13%	19%
Black	6%	5%
Hispanic	9%	6%
White	70%	69%
Free/Reduced Lunch	19%	14%
Limited English Proficiency	14%	14%
Special Education	11%	10%

Table 30. Correlation Coefficients between GMRT-4th and FAST™ aReading Scaled Score

Grade	Decoding	Vocabulary	Comprehension	Composite
1	.82 (131)	-	.73 (130)	.83 (125)
2	.68 (163)	-	.75 (215)	-
3	-	.79 (170)	.81 (168)	.84 (165)
4	-	.76 (182)	.72 (180)	.78 (175)
5	-	.65 (182)	.58 (187)	.64 (181)
1-5	.75 (348)	.74 (534)	.82 (881)	.86 (646)

Note. Sample size is denoted by ().

Overall, there appears to be a strong positive correlation between composite scores from the GMRT-4th and FAST™ aReading scaled scores. There is some variability between grades, with coefficient values between .64 and .83. Subtests showed greater variability. Specifically, comprehension correlation coefficients ranged from .58 to .81.

Reliability-Related Validity Evidence

Given the adaptive nature of FAST™ aReading test, a proxy for internal consistency is provided by Samejima (1994), based on the standard error of measurement of an instrument. Using this proxy, the internal consistency reliability coefficient for FAST™ aReading is approximately .95 (based on approximately 2,333 students).

Three-month test-retest reliability resulted in the following coefficients for 2,038 students in grades 1-5. Growth was measured four times over the academic year. The results by grade: one .71, two .87, three .81, four .86, five .75.

Evidence Related to Bias

Bias analyses of a sample of the items that comprise FAST™ aReading were conducted using data collected during the 2016-17 and 2017-18 academic years. Data for each year were analyzed separately. There were sufficient data to examine bias in relation to race/ethnicity. The race/ethnicity group comparisons examined were White versus Black, White versus Hispanic, White versus Asian, and White versus Native American. The results indicated that there is no or negligible DIF for all items examined in both years for all the race/ethnicity comparisons.

Bias was assessed using the logistic regression procedure for detection of uniform and non-uniform differential item functioning (DIF). The advantages of using the logistic regression procedure for DIF detection include being a model-based approach and having the capability to detect both uniform and non-uniform DIF with adequate and equal power; however, the procedure also tends to inflate Type I error rates. As such, an effect size measure developed by Jodoin and Gierl (2001) was computed and evaluated in addition to statistical significance. Jodoin and Gierl present a four-category framework for interpreting the effect size measure, where the four categories are indicative of no, negligible, moderate, and severe DIF.

Evidence for Use of FAST™ aReading as a Screening Tool

Cut scores for FAST™ aReading to predict students "At Risk" and "Somewhat at Risk" for reading difficulties were developed using the Gates-MacGinitie Reading Tests-Fourth Edition (GMRT-4th; MacGinitie, MacGinitie, Maria, & Dreyer, 2000) and the Measures of Academic Progress (MAP). Categories for the former were defined as students scoring below the 40th and 20th percentiles of the local sample and cut scores for each category developed by an adjacent school district for MAP were used on this sample.

At the beginning of the school year (October 2010) students completed an FAST™ aReading assessment. The measure was group administered via a mobile computer lab by a team of graduate students. Scaled scores were calculated for each student. In February 2011, the same students completed the GMRT-4th. Composite scores were available for all grades except second grade. Due to time constraints, one GMRT-4th subtest could not be administered to second grade students (the only grade that requires three subtests to yield a composite score). As a result, comprehension subtest scores were used for analysis. Test booklets were hand scored and inter-rater reliability was 100% across all subtest and composite scores. MAP scores for spring testing were provided to the FAST™ aReading team from school administrators.

FAST™ aReading classification accuracy was derived from a sample of 777 students in first through fifth grades from two suburban schools in the Midwest. The sample was 49% female and 51% male. Approximately 67% of students in the sample were White, 19% Asian/Pacific Islander, 5% Black, 5% Hispanic, 2% American Indian, and 2% unspecified. In addition, 10% of students were receiving special education services, and 10% of students were classified as having limited English language proficiency. Socioeconomic status information was not available for the sample, but the schools the students were drawn from had rates of free and reduced lunch of 13% and 23% in 2009-10.

The ROC curve analysis results for each grade for students at high risk and somewhat at risk with the GMRT-4th are presented in Table 31. Evaluation of the table below indicated that across grades, AUC statistics were extremely high, especially for students at high risk (median = .92) and values were still high for students at some risk (median = .87). In addition, sensitivity was higher for each grade when determining students at high risk compared to at some risk. positive predictive power was higher across grades when predicting students at some risk (median = .72 versus .56) while the opposite was true for negative predictive power (median = .82 versus .96).

Table 31. Classification Accuracy statistics for FAST™ aReading and GMRT-4th

Grade	N	FAST™ aReading Cut Score	Sensitivity	Specificity	PPP	NPP	AUC
High Risk - Below 20th Percentile[a]							
1	116	430	.88	.87	.66	.96	.94
2	188	461	.70	.93	.74	.92	.88
3	159	490	.97	.77	.50	.99	.92
4	156	495	.85	.92	.72	.96	.94
5	159	506	.85	.84	.59	.95	.87
Somewhat at Risk - Below 40th Percentile							
1	116	436	.76	.86	.81	.82	.91
2	188	477	.86	.71	.71	.86	.87
3	159	490	.82	.87	.78	.90	.89
4	156	506	.72	.82	.73	.81	.82
5	159	522	.83	.76	.71	.86	.85

Note. [a]The 20th percentile was used for this sample, which should approximate the 15th percentile. N = sample size; PPP = positive predictive power; NPP = negative predictive power; AUC = area under the curve.

A similar pattern of results emerged when predicting performance on the MAP (see Table 32). Compared to the GMRT-4th as a criterion, NPP was much higher when predicting MAP scores. This could be attributed to the fact that the base rate of students at risk was much lower for MAP scores.

Table 32. Classification Accuracy Statistics for FAST™ aReading and MAP

Grade	N	FAST™ aReading Cut Score	Sensitivity	Specificity	PPP	NPP	AUC
			High Risk - 20th Percentile				
2	188	497	1	.73	.14	1	.89
3	159	517	.95	.76	.21	1	.95
4	156	537	.96	.78	.30	1	.94
5	159	537	1	.82	.20	1	.93
			Somewhat at Risk - 40th Percentile				
2	188	490	.77	.84	.41	.96	.89
3	159	527	.89	.77	.46	.97	.89
4	156	537	.82	.87	.65	.94	.92
5	159	547	.93	.77	.39	.99	.88

Note. N = sample size; PPP = positive predictive power; NPP = negative predictive power; AUC = area under the curve.

Finally, classification accuracy analyses were conducted with FAST™ aReading and the Minnesota Comprehensive Assessment (MCA) to determine if FAST™ aReading predicted state reading assessments. The sample consisted of 1,786 students in third, fourth, and fifth grades from eight schools in the upper Midwest. Approximately 50% of students were female and 50% male. The ethnicity breakdown was approximately 45% White, 23% Black, 15% Hispanic, 8% Asian/Pacific Islander, 1% American Indian or Alaska Native, and 9% multiracial. In addition, 12% of students were receiving special education services. Socioeconomic status information was not available for the sample, but the schools the students were drawn from had rates of free and reduced lunch ranging from 16% to 83% in 2013.

Students completed FAST™ aReading assessment and MCAs during the spring of 2013. Students with incomplete data in FAST™ aReading, or those students with incomplete MCA Achievement Level Scores were excluded from analyses. ROC Analysis was used to determine classification accuracy of FAST™ aReading with Spring MCA scale scores serving as the criterion measure. Students were disaggregated by grade level. Classification accuracy was computed for students at “High Risk” and “Somewhat At Risk” on MCA Scale Scores. “High Risk” includes those students that did not meet standards. “Somewhat At Risk” includes those students who did not meet or only partially met standards. Classification accuracy statistics are provided in Table 33. Data collection is ongoing for all grade levels.

Table 33. Classification Accuracy for FAST™ aReading and MCA-III

Grade	N	FAST™ aReading M (SD)	MCA-III M (SD)	Correlation	Cut Score	AUC	Sensitivity	Specificity
High Risk (Does Not Meet Standards)								
3	629	534 (27)	347 (21)	.82	524.5	.90	.82	.81
4	615	549 (28)	447 (16)	.81	536.5	.92	.84	.84
5	516	564 (32)	553 (16)	.84	544.5	.96	.89	.86
Somewhat High Risk (Does Not Meet or Partially Meets Standards)								
3	629	534 (27)	347 (21)	.82	532.5	.90	.83	.82
4	615	549 (28)	447 (16)	.81	550.5	.89	.82	.82
5	516	564 (32)	553 (16)	.84	556.5	.93	.84	.84

Note: N = sample size; M = mean; SD = standard deviation; AUC = area under the curve.

FAST™ aReading evidence of classification accuracy is not limited to the Midwest. The following classification accuracy information was obtained from samples of students in other regions of the US. See Table 34 and Table 35 for results related to the Massachusetts Comprehensive Assessment, Table 36 and Table 37 for results related to the Georgia Criterion-Referenced Competency Tests.

Table 34. Classification Accuracy Fall FAST™ aReading with Massachusetts Comprehensive Assessment

Grade	N	FAST™ aReading M (SD)	MCA M (SD)	Correlation	Cut Score	AUC	Sensitivity	Specificity
Some Risk (“Warning” and “Needs Improvement”)								
3	93	485.91 (17.59)	241.89 (14.08)	.63**	478.5	.79	.73	.73
4	93	492.68 (18.52)	238.40 (15.14)	.69**	492.5	.85	.75	.78
5	72	506.93 (18.58)	243.63 (13.20)	.69**	507.5	.90	.85	.79
Some Risk (“Needs Improvement”)								
3	93	485.91 (17.59)	241.89 (14.08)	.63**	480.5	.71	.72	.63
4	93	492.68 (18.52)	238.40 (15.14)	.69**	494.5	.72	.71	.60
5	72	506.93 (18.58)	243.63 (13.20)	.69**	506.5	.88	.78	.87
High Risk (“Warning”)								
3	93	485.91 (17.59)	241.89 (14.08)	.63**	475.5	.80	.75	.72
4	93	492.68 (18.52)	238.40 (15.14)	.69**	476.5	.92	.89	.84
5	72	506.93 (18.58)	243.63 (13.20)	.69**	495.5	.85	1.00	.79

Note: N = sample size; M = mean; SD = standard deviation; AUC = area under the curve.

Table 35. Classification Accuracy of Winter FAST™ aReading with Massachusetts Comprehensive Assessment

Grade	N	FAST™ aReading M (SD)	MCA M (SD)	Correlation	Cut Score	AUC	Sensitivity	Specificity
Some Risk ("Warning" and "Needs Improvement")								
3	91	498.99 (18.66)	241.89 (14.08)	.76**	499.5	.85	.76	.74
4	94	504.33 (16.37)	238.40 (16.37)	.69**	505.5	.85	.76	.78
5	74	515.31 (14.42)	243.63 (14.42)	.61**	516.5	.83	.82	.78
Some Risk ("Needs Improvement")								
3	91	498.99 (18.66)	241.89 (14.08)	.76**	501.5	.71	.72	.64
4	94	504.33 (16.37)	238.40 (16.37)	.69**	506.5	.71	.78	.64
5	74	515.31 (14.42)	243.63 (14.42)	.61**	516.5	.81	.81	.76
High Risk ("Warning")								
3	91	498.99 (18.66)	241.89 (14.08)	.76**	476.5	.97	.88	.95
4	94	504.33 (16.37)	238.40 (16.37)	.69**	485.5	.94	.78	.88
5	74	515.31 (14.42)	243.63 (14.42)	.61**	506.00	.86	1.00	.80

Note: N = sample size; M = mean; SD = standard deviation; AUC = area under the curve.

Table 36. Classification Accuracy of Fall FAST™ aReading with Georgia Criterion-Reference Competency Tests

Grade	N	FAST™ aReading M (SD)	CRCT M (SD)	Correlation	Cut Score	AUC	Sensitivity	Specificity
Some Risk (Meets Standards)								
3	329	483.81 (18)	848.65 (28)	.73*	481.50	.83	.76	.76
4	320	491.37 (16)	848.18 (27)	.64*	490.50	.80	.73	.75
5	353	497.81 (16)	841.22 (25)	.64*	499.50	.75	.70	.68
High Risk (Does Not Meet Standards)								
3	329	483.81 (18)	848.65 (28)	.73*	466.50	.94	.82	.86
4	320	491.37 (16)	848.18 (27)	.64*	478.50	.89	.83	.76
5	353	497.81 (16)	841.22 (25)	.64*	485.00	.89	.79	.79

Note: N = sample size; M = mean; SD = standard deviation; AUC = area under the curve.

Table 37. Classification Accuracy of Winter FAST™ aReading with Georgia Criterion-Reference Competency Tests

Grade	N	FAST™ aReading M (SD)	CRCT M (SD)	Correlation	Cut Score	AUC	Sensitivity	Specificity
Some Risk (Meets Standards)								
3	327	495.67 (18)	848.64 (28)	.75*	498.50	.83	.76	.76
4	318	505.31 (16)	848.33 (27)	.71*	505.50	.82	.77	.78
5	351	512.19 (15)	841.14 (25)	.66*	516.50	.78	.71	.72
6	283	518.78 (13)	850.14 (23)	.67*	519.50	.87	.77	.80
7	322	521.77 (16)	842.50 (24)	.64**	512.5	.91	.82	.81
8	311	524.69 (22)	850.36 (54)	.33**	517.5	.95	1.00	.73
High Risk (Does Not Meet Standards)								
3	327	495.67 (18)	848.64 (28)	.75*	477.50	.95	.83	.86
4	318	505.31 (16)	848.33 (27)	.71*	487.50	.94	.83	.76
5	347	512.19 (15)	841.14 (25)	.66*	500.50	.92	.86	.85
6	283	518.78 (13)	850.14 (23)	.67*	NA	NA	NA	NA
7	322	521.77 (16)	842.50 (24)	.64**	509.5	.92	.86	.86
8	311	524.69 (22)	850.36 (54)	.33**	511.5	.92	.86	.84

Note: N = sample size; M = mean; SD = standard deviation; AUC = area under the curve.

FAST^{TM} CBMmath

FAST^{TM} CBMmath Purpose and Use

The research behind Curriculum-Based Measurement (CBM) for the purposes of assessing students' math computation skills has existed for well over three decades. Researchers from a variety of institutions nationally have contributed to what the education field knows today about CBM for mathematics. Through each iteration, researchers and teachers have gained important knowledge about what works well and what improvements can be made to efficiently and accurately assess students' computation skills to inform instruction. FastBridge Learning's FAST™ CBMmath offers the latest in the lineage of these research-based assessments. These tools are designed to better meet the needs of teachers, while still offering the measurement properties needed to accurately assess and monitor student performance over time.

FAST™ CBMmath consists of two types of assessments. Two of those assessments measure students' computation skills from grades 1 through 6 (i.e., FAST™ CBMmath Process and FAST™ CBMmath Automaticity). *Automaticity skills* are those which are automatized. They include rapid recall of math facts (i.e., multiplication time tables) that should take little or no cognitive effort. *Process skills* are those in which the student may have to solve multiple steps to reach a solution. Students are not expected to have these items memorized and would be given paper and a pencil to work out the solution. Because of the higher amount of cognitive effort, process skills are inherently more difficult than fluency skills.

In addition to the FAST™ CBMmath Automaticity and FAST™ CBMmath Process assessments, FAST™ CBMmath Concepts and Applications (FAST™ CBMmath CAP) is intended to measure applied and multi-step skills. These three assessments make up a suite of FAST™ CBMmath tools that covers the full range of elementary and middle school math skills.

The goal of FAST™ CBMmath is to serve as a tool to screen and monitor students' progress in math achievement for students in grades K through 8. Every skill created was based on skills outlined in the Common Core State Standards (2010). Together, these tools are used to screen and monitor the computation process (automaticity and process) and applied skills.

FAST™ CBMmath Content Description

FAST™ CBMmath Automaticity evaluates the degree to which basic facts and operations are accurate and automatic (fluent). It consists of a General Outcome Measure (GOM) for each level, plus optional single- and multi-skill subtests for further classification purposes. FAST™ CBMmath Automaticity is available for screening and progress monitoring. There are single skill and multiple skill subtests, which can be used for progress monitoring and further classification purposes.

FAST™ CBMmath Process evaluates the degree to which the primary processes (steps in a multi-step problem) are completed with accuracy. FAST™ CBMmath Process gives teachers the opportunity to track the types of errors students are making when solving single and multi-step math problems. FAST™ CBMmath Process consists of a General Outcome Measure (GOM) for each level, plus optional single- and multi-skill subtests for further classification purposes. FAST™ CBMmath Process is available for screening, progress monitoring or further classification purposes.

FAST™ CBMmath CAP evaluates the student's skills for solving complex and multi-step math problems. If is available for screening and progress monitoring. Assessment of student mathematics skills requires attention to multiple aspects of math proficiency. The FAST™ CBMmath CAP assessment, which measures math concepts and applications, includes items that cover skills from computation fact fluency to multi-step algebra problems.

FAST™ CBMmath Content Development

FAST™ CBMmath Automaticity and FAST™ CBMmath Process skills were based on computation skills outlined in the Common Core State Standards (CCSS 2010). For each skill, 21 forms were created. The first form was a screener and the following 20 were progress monitoring forms. FAST™ CBMmath CAP items were developed from an existing bank of items, including those used in the FAST™ aMath and FAST™ Standards Based Math assessments. All items were developed in accordance with the mathematics learning hierarchy in the Common Core State Standards, which includes: Counting and Cardinality, Operations and Algebraic Thinking, Number and Operations in Base 10, Measurement and Data, and Geometry.

FAST™ CBMmath Automaticity

First Grade

1x1 addition to 10. These items are one digit by one digit addition with a sum less than or equal to 10.

1x1 subtraction to 10. These items are subtraction of a one digit number from a number less than or equal to 10, with the difference being greater than or equal to 0.

1x1 addition to 18. These items are one digit by one digit addition with a sum less than or equal to 18.

2x1 subtraction from 20. These items are subtraction of a one digit number from a number less than or equal to 20, with the difference being greater than or equal to 0 and no regrouping necessary.

1x1 and 1x2 addition to 20. These items are both two one-digit numbers and a one-digit and a two-digit number with a sum less than or equal to 20.

1x1 addition and subtraction to 10. Items created for 1x1 addition and 1x1 subtraction to 10 were randomized and used in the creation of 30-item forms consisting of both 1x1 addition items with a sum less than or equal to 10, and 1x1 subtraction items with a product greater than or equal to 0.

1x1 and 2x1 addition and subtraction to 20. Addition items are both two one digit numbers and a one digit and a two-digit number with a sum less than or equal to 20. Subtraction items are both two one digit numbers and a two-digit number and a two-digit number with a product equal to or greater than 0.

General Outcome Measure (GOM). Items for the Grade 1 Fluency General Outcome Measure (GOM) included: 1x1 addition to 10, 1x1 subtraction from 10, 1x1 subtraction to 18, 1x1 and 2x1 addition to 20, and 2x1 subtraction. No maintenance items were included.

Second Grade

2 x 1 addition to 100 without regrouping. Items are two digit by one digit addition with a sum less than or equal to 100 and do not require regrouping.

2 x 2 addition to 100 without regrouping. Items are two digit by two-digit addition with a sum less than or equal to 100 and do not require regrouping.

2 x 1 subtraction from 100 without regrouping. Items are subtraction of a one digit number from a two-digit number less than or equal to 100, with the difference being greater than or equal to 0, and re-grouping is necessary.

2 x 2 subtraction from 100 without regrouping. Items are subtraction of a two-digit number from a two-digit number less than or equal to 100, with the difference being greater than or equal to 0, and re-grouping is not necessary.

2x1 and 2x2 addition to 100 and 2x1 and 2x2 subtraction from 100 without regrouping. Items are two digit by one digit and two digit by two-digit addition with a sum less than or equal to 100 that do not require re-grouping and two digit by one digit and two digit by two-digit subtraction with a difference less than or equal to 100 with no regrouping.

Grade 2 Fluency GOM. Items for the Grade 2 Fluency General Outcome Measure (GOM) included: 2x1 addition to 100, 2x2 addition to 100, 2x1 subtraction from 100, and 2x2 subtraction from 100 without regrouping. Maintenance items from Grade 1 were included.

Third Grade

Fact Families. Items included multiplication fact families from the digits 1 to 12.

1x2 multiplication to 12. Items were a one digit factor multiplied by a one or two-digit factor between 0 and 12.

2 x 1 division from 100 without remainder. Items are two digit numbers less than or equal to 100 by one digit division with the quotient being a whole number, and no remainder.

1x2 multiplication to 12 and 2x1 Division. Items were a one digit factor multiplied by a one or two-digit factor between 0 and 12 and two digit numbers less than or equal to 100 by one digit division with the quotient being a whole number, and no remainder.

Addition, subtraction, multiplication, and division (all skills combined). Items included: 1x2 to 12 multiplication; 3x2 and 3x3 addition to 1000; 3x2 and 3x3 subtraction from 1000; and 2x1 and 2x2 division from 100 with no remainder.

Grade 3 Fluency GOM. Items for the Grade 3 Fluency General Outcome Measure (GOM) included: 1x2 multiplication to 12 and 2x1 and 2x2 division from 100 with no remainder. Maintenance items from Grade 2 were included.

FAST™ CBMmath Process

Second Grade

2x1 and 2x2 addition to 100 with regrouping. Items are two digit by one digit and two digit by two-digit addition with a sum less than or equal to 100 that require regrouping.

2x1 and 2x2 subtraction from 100 with regrouping. Items are two digit by one digit and two digit by two-digit subtraction with a difference less than or equal to 100 that require regrouping.

2x1 and 2x2 addition to 100 and 2x1 and 2x2 subtraction from 100 with regrouping. Items are two digit by one digit and two digit by two-digit addition with a sum less than or equal to 100 that require re-grouping and two digit by one digit and two digit by two-digit subtraction with a difference less than or equal to 100 that require regrouping.

Grade 2 Process GOM. Items for the Grade 2 Process General Outcome Measure (GOM) included: 2x1 addition to 100, 2x2 addition to 100, 2x1 subtraction from 100, and 2x2 subtraction from 100, all with regrouping.

Third Grade

3x2 and 3x3 addition to 1000. Items were three digit by two digit or three-digit addition with a sum less than or equal to 1000 that require regrouping.

3x2 and 3x3 subtraction from 1000. Items were three digit by two digit or three-digit subtraction with a difference less than or equal to 1000 that require regrouping.

3x2 and 3x3 addition and subtraction to 1000 with regrouping. Items were three digit by two digit or three-digit addition and subtraction with a sum or difference less than or equal to 1000 that require regrouping

Grade 3 Process GOM. Items for the Grade 3 Process General Outcome Measure (GOM) included.3x2 and 3x3 addition to 1000 and 3x2 and 3x3 subtraction from 1000. Maintenance items from Grade 2 were included.

Fourth Grade

3 x 3 x 3 or 3 x 3 x 2 addition. Items are adding three numbers (all three digits) or three numbers (2 three digits and 1 two digit) with no upper limit.

3x1 and 4 x 1 multiplication. Items are 3- and 4-digit multiplication by one digit (no upper limit).

2 x 2 multiplication. Items are two digit by two-digit multiplication with no upper limit.

2 x 1 division from 100 with and without remainder. Items are two digit by one digit division with and without remainders.

4 x 1 and 3 x 1 division with and without remainder. Items are three or four digit numbers divided by one digit numbers with and without remainders.

3 x 1, 4 x 1, and 2 x 2 multiplication and 2 x 1 division. Multiplication items were either three digit by one digit, four digits by one digit, or two digits by two digits. Division problems were two-digit divisor and one digit divisor with a quotient equal to or greater than zero.

3 x 3 x 2 and 3 x 3 x 3 addition; 1 x 3, 1 x 4, and 2 x 2 multiplication, and 2 x 1 division, and 3 x 1 and 4 x 1 division. Addition items were three digit by three digit by either two or three digit with no upper limit. Multiplication items were either three digit by one digit, four digits by one digit, or two digits by two digits. Division problems were two digit, three digit, or four-digit divisor and one digit divisor with a quotient equal to or greater than zero.

Grade 4 Process GOM. Items for the Grade 4 Process General Outcome Measure (GOM) included: 1x3 and 1x4 multiplication; 2x2 multiplication; 3x3x2 and 3x3x3 addition; 2x1 division from 100 with and without a remainder; and 4x1 and 3x1 with and without a remainder. Maintenance items from Grade 3 were included.

Fifth Grade

2 x 3 and 3 x 3 multiplication. Items are two or three digit by three digit multiplication items with no upper limit.

3 x 2 and 4 x 2 division without remainder. Items are three or four digit dividends with two digit divisors, without remainder.

3 x 2 and 4 x 2 division with remainder. Items are three or four digit dividends with two digit divisors, with a remainder.

3 x 2 and 3 x 3 multiplication; 3 x 2 and 4 x 2 division with and without a remainder (all skills combined). Items are two or three digit by three digit multiplication with no upper limit and three or four digit dividends with two digit divisors, with or without remainder.

3 x 2 and 4 x 2 division with and without a remainder. Items are three or four digit dividends with two digit divisors, with or without remainder.

5th Grade GOM. Items for the Grade 5 Process General Outcome Measure (GOM) included: 2x3 and 3x3 multiplication; 3x2 and 4x2 division without a remainder; and 3x2 and 4x2 division with a remainder. Maintenance items from Grade 4 were included.

Sixth Grade

Decimal addition with regrouping. Items are two or three digit numbers with one, two, or three decimal places.

Decimal subtraction with regrouping. Items are two or three digit numbers with one, two, or three decimal places added to a two or three-digit number with the equal number of decimal places.

Decimal multiplication with regrouping. Items are one or zero digit numbers with one, two, or three decimal places multiplied by a zero or one digit number with one, two, or three decimal places.

FAST™ CBMmath Administration

FAST™ CBMmath Automaticity items are administered electronically (i.e., on a computer or tablet) or using a paper-and-pencil form. Assessments are timed for up to 4 minutes. Forms contain between 30-40 items. The General Outcome measure (GOM) is the grade level screener. Each subtest has its own screener which is used as the starting point for progress monitoring.

FAST™ CBMmath Process forms consist of a General Outcome Measure (GOM) per grade level and benchmark period for Screening plus optional multiple and single-skill subtests for further classification purposes and progress monitoring. This assessment is administered via paper-pencil forms. This is a group administered assessment and forms contain between 10 and 24 items.

FAST™ CBMmath CAP forms are computer-administered and include about 20 questions per test. Times for grade levels range from 15 to 30 minutes.

FAST™ CBMmath Scores and Scoring

For screening, FAST™ CBMmath Process uses a rapid scoring method, so teachers can quickly enter assessments and get a general idea of error types. For progress monitoring, teachers use the error analysis method, attaching specific error types to specific problems. Scores are calculated and reported as items_correct_per_ten_minutes. FAST™ CBMmath Process items are weighted by the total number of possible process errors within the item. For example, there are more possible process errors within a multiplication item than within an addition item. Therefore, a multiplication item would be worth more possible points. Incorrect answers will be analyzed to determine which of the potential errors led to the incorrect response.

When applicable, intermediary steps were included on answer keys to facilitate the scoring of process items. For example, within multi-digit multiplication and long division problems that have not been automatized, there is a certain order of operations to solve the problems. For multiplication items, the student must multiply each digit by another and add numbers together. For division, students divide numbers and subtract in the intermediary steps. Formulas to calculate these intermediary steps were created and used to make answer keys. Thus, teachers can identify at which intermediary step the student made an error. There are 9 total possible errors in process items. The possible errors are as follows:

> *Regrouping or carrying error:* A student made an error in adding numbers with a sum greater than 10 or subtracting numbers with a difference less than 0. The student made an error in "borrowing" from the next placeholder to solve the addition or subtraction problem.

Calculation-addition error: A student made an error in adding two numbers; a student wrote an incorrect answer for an addition problem.

Calculation-subtraction error: A student made an error in subtracting two numbers; a student wrote an incorrect answer for a subtraction problem.

Calculation-multiplication error: A student made an error in multiplying two numbers; a student wrote an incorrect answer for a multiplication problem.

Calculation-division error: A student made an error in dividing two numbers; a student wrote an incorrect answer for a division problem.

Misread operation sign error: A student made an error due to misreading an operation sign and performing a different operation than specified in the item.

Placeholder/maintain value error: A student made an error in maintaining place value or using a placeholder in calculating a multiplication or division item.

Remainder error: A student made an error in a division problem regarding the remainder; the remainder in the solution is incorrect.

Unknown error: The error a student made is unclear or unknown, or it appears that the student guessed. The error does not fit any of the above categories, but the student did not leave the item completely blank.

Blank/skip: The student left the item blank and did not attempt to solve the item.

Non-errors: If a student reverses a digit, this is not reported as an error.

Score Types

FAST™ CBMmath Automaticity scores are reported in items correct per 10 min. Although the administrations are only one to two minutes in duration, the use of a 10-minute scale helps avoid decimals and provides a more sensitive scale (i.e., 1.3 items correct per minute versus 13 items correct per 10 minutes, IC10).

FAST™ CBMmath Process scores are based on the multiple steps required to solve a problem. Although the administration is timed to 10 to 15 minutes, these are not considered fluency- or automaticity-type assessments. Items are weighted by the total number of possible process errors within the item. For example, there are more possible process errors within a multiplication item than within an addition item. Therefore, a multiplication item would be worth more possible points. Incorrect answers will be analyzed to determine which of the potential errors led to the incorrect response.

FAST™ CBMmath Construct Validity

Content-Related Validity Evidence

The Common Core State Standards for Mathematics were established in 2010. The standards alignment with FAST™ CBMmath assessments are presented in the tables below.

Table 38 FAST™ CBMmath CCSS Alignment

	Grade	Skill(s)	Common Core Standard
FAST™ CBMmath Automaticity: Single Skill	Grade 1	1x1 Addition to 10	1.OA.6
		1x1 Addition to 18	1.OA.6
		1x1 and 2x1 Addition to 20	1.OA.6
		1x1 Subtraction from 10	1.OA.6
		1x1 Subtraction from 20	1.OA.6
	Grade 2	2x1 Addition to 100	2.NBT.5
		2x2 Addition to 100	2.NBT.5
		2x1 Subtraction from 100	2.NBT.5
		2x2 Subtraction from 100	2.NBT.5
	Grade 3	Fact Families: 1-12	3.OA.7
		2x1 Multiplication to 12	3.OA.7
		2x1 Division from 100	3.OA.7
FAST™ CBMmath Automaticity: Multi-Skill	Grade 1	1x1 Addition to 10 1x1 Subtraction to 10	1.OA.6
		1x1 and 2x1 Addition to 20 1x1 and 2x1 Subtraction to 20	1.OA.6
	Grade 2	Addition to 100 Subtraction to 100	2.NBT.5
	Grade 3	2x1 Multiplication to 12 2x1 Division to 12	3.OA.7
		Addition, Subtraction, Multiplication, and Division	3.OA.7
FAST™ CBMmath Process: Single-Skill	Grade 2	2x1 and 2x2 Addition to 100	2.NBT.5
		2x1 and 2x2 Subtraction from 100	2.NBT.5
	Grade 3	3x2 and 3x3 Addition to 1000	3.NBT.2
		3x2 and 3x3 Subtraction from 1000	3.NBT.2
	Grade 4	3x1 and 4x1 Multiplication	4.NBT.5
		2x2 Multiplication	4.NBT.5
		3x3x2 and 3x3x3 Addition	4.NBT.4
		2x1 Division from 100 with and without remainder	4.NBT.6

		3x1 and 4x1 Division with and without remainder	4.NBT.6
	Grade 5	2x3 and 3x3 Multiplication	5.NBT.5
		3x2 and 4x2 Division without remainder	5.NBT.6
		3x2 and 4x2 Division with remainder	5.NBT.6
	Grade 6	Decimal Addition	6.NS.3
		Decimal Subtraction	6.NS.3
		Decimal Multiplication	6.NS.3
FAST™ CBMmath Process: Multi-Skill	Grade 2	2x1 and 2x2 Addition 2x1 and 2x2 Subtraction	2.NBT.5
	Grade 3	3x2 and 3x3 Addition to 1000 3x2 and 3x3 Subtraction from 1000	3.NBT.2
	Grade 4	3x1, 4x1, and 2x2 Multiplication 2x1 Division	4.NBT.5-6
		3x3x2 and 3x3x3 Addition 3x1, 4x1, and 2x2 Multiplication 2x1, 3x1, and 4x1 Division	4.NBT.4-6
	Grade 5	3x2 and 3x3 Multiplication 3x2 and 4x2 Division with and without remainder	5.NBT.5-6
		3x2 and 4x2 Division with and without remainder	5.NBT.6
	Grade 6	Decimal Addition Decimal Subtraction Decimal Multiplication 3x2 and 3x3 Multiplication 3x2 and 3x3 Division	6.NS.3

Criterion-Related Validity Evidence

Criterion-related validity of FAST™ CBMmath Automaticity and FAST™ CBMmath Process were examined using FAST™ aMath scaled scores. Students completed FAST™ CBMmath GOM measures and FAST™ aMath during Fall of the 2014-15 school year. Validity coefficients are presented for both FAST™ CBMmath Automaticity and FAST™ CBMmath Process in Table 39.

Table 39 Criterion-Related Validity for FAST™ CBMmath GOMs

Assessment	N	Correlation
FAST™ CBMmath Automaticity		
Grade 1 GOM	326	.60
Grade 2 GOM	612	.53
Grade 3 GOM	670	.41
FAST™ CBMmath Process		
Grade 2 GOM	549	.59
Grade 3 GOM	463	.63
Grade 4 GOM	673	.56
Grade 5 GOM	565	.65
Grade 6 GOM	38	.69

Note: N = sample size.

FAST^{TM} earlyMath

FAST^{TM} earlyMath Purpose and Use

The objective of the FAST™ earlyMath measures is to extend and improve on the development of curriculum based measures for early numeracy. More specifically, FAST™ earlyMath was developed to be an efficient, instructionally relevant, and technically adequate assessment to identify students who may have difficulties in mathematics and monitor student progress. The suite of 17 FAST™ earlyMath subtests are used to screen and monitor a student's progress in foundational math skills. These subtests are designed for students in the early primary grades and are designed for use in kindergarten and first grade.

The FAST™ earlyMath subtests include: Numeral Identification-K, Numeral Identification-1, Subitizing, Match Quantity, Quantity Discrimination-Most, Quantity Discrimination-Least, Number Sequence-K, Number Sequence-1, Composing, Decomposing-K, Decomposing-1, Counting Objects, Equal Partitioning, Verbal Addition, Verbal Subtraction, Story Problems and Place Value. Some of the subtests have both kindergarten and first grade versions. This is to better represent developmental trajectories; the types of math skills that students can typically do at each of these grades is very different.

FAST™ earlyMath performance is an indicator of student math development. It is designed to assess math skills that predict successful broad mathematics proficiency. Not all FAST™ earlyMath subtests are given each screening period. Instead, three subtests are given at each screening period. The scores from the screening subtests given are used to provide a composite score for each student. The broadest score available - and best estimate of your students' early math skills - is the FAST™ earlyMath composite score. This is intended to optimize validity and risk evaluation. Subtests that are not included in the composite are considered supplemental. Supplemental subtests may be used to diagnose and evaluate skill deficits. Results from supplemental subtests provide guidance for instructional and intervention development.

FastBridge recommends weekly progress monitoring in FAST™ earlyMath and teachers have the option to monitor student progress using the Numeral Identification-K, Numeral Identification-1, Number Sequence K, Decomposing-1, Match Quantity, Quantity Discrimination Most and Least, and Place Value subtests. For students in kindergarten, it is recommended to use the Numeral Identification-K subtest. For students in first grade, we recommend progress monitoring with the Decomposing-1 subtest. This may vary on the area of intervention and instruction that is being targeted

FAST™ earlyMath Content Description

FAST™ earlyMath assesses a variety of skills including Number, Operations and Relations. These skill domains are important in the development of early numeracy (National Research Council [NRC], 2009).

Number. The Number domain involves the ability to perceive, say, describe, and construct numbers (NRC, 2009). It also involves the understanding of the rules and processes of the counting sequence (Purpura & Lonigan, 2013). Tasks that measure a child's understanding of Number knowledge typically include verbal counting, counting forward and backward, counting error identification, structured counting (i.e., counting with one-on-one correspondence), knowledge of cardinality, subitizing, and estimation (NRC, 2009; Purpura & Lonigan, 2013). Many aspects of Number knowledge are developed before a child enters formal schooling, and are part of a child's informal mathematical knowledge. Informal mathematics knowledge is the foundation for children learning formal mathematics skills once entering school, and research has shown that informal mathematics skills are strong predictors of later formal mathematics ability (Jordan, Kaplan, Nabors Oláh, & Locuniak, 2006; Lembke & Foegen, 2009; Mazzocco & Thompson, 2005).

Relations. The Relations early numeracy domain involves having the knowledge of how two or more numbers or objects are connected to each other and possessing an understanding of the mental line. Relation skills involve developing an understanding of connections between sets of quantities (e.g., set comparison and matching) and quantities and numerals (e.g., set to numerals; Purpura, Baroody, et al. 2013), relationships between two or more numerals (e.g., ordinality, number order), and numeral names (NRC, 2009; Purpura & Lonigan, 2013). The specific number Relation skills mediate the relationship between informal and formal mathematics knowledge (Purpura, Baroody, et al. 2013). In other words, developing skills such as matching a quantity to a numeral and numeral identification are necessary steps before children can apply mathematics knowledge to formal concepts.

Operations. The Operations domain, also referred to as Arithmetic Operations, is defined as understanding how groups of numbers, objects, or a combination of the two are composed and decomposed (Purpura & Lonigan, 2013). Proficiency of skills in this domain are assessed using tasks such as verbal addition and subtraction, story problems, and decomposing (NRC, 2009; Purpura & Lonigan, 2013). Although some children may acquire skills in the Operations domain before entering formal schooling, such as simple addition, procedures to develop the Operations domain skills are formally taught to Kindergarten and First Grade students. For example, according to the Common Core State Standards by the end of First Grade students should acquire skills such as adding and subtracting within 20 and solving word problems (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

FAST™ earlyMath Content Development

FAST™ earlyMath development followed research recommendations for item and test development and included an iterative process of pilot testing, feedback and revisions. Subtests were created by reviewing the research literature in math curriculum based assessment and early numeracy skill development. The FAST™ earlyMath subtests were developed with experts with a variety of perspectives on mathematics instruction and assessment. This includes university faculty, research assistants, experience math teachers, math specialists, math interventionists, school district administrators and experts in assessment. The FAST™ earlyMath subtests were created to measure each of the three established domains of early numeracy with alignment to the CCSS. The following is a list of all FAST™ earlyMath subtests and a brief description of the skills that they measure.

Match Quantity. The Match Quantity subtest assesses the student's ability to correctly match a quantity of dots to a numeral, given a choice of four numerals. In this task, students are required to make a connection between quantity and numeral.

Decomposing - K. The Decomposing subtest assesses the student's ability to automatically decompose (take apart) fives and tens by using "parts" and a "whole." The Decomposing - K assessment is designed to measure if students automatically know how to decompose numbers to five and ten and are therefore not allowed to use any counting strategies.

Subitizing. The Subitizing subtest measures a student's ability to recognize the correct quantity of dots when presented with an image for 1 second. The measure includes both subitizing (the ability to instantly and accurately recognize groups of objects between one and three items) and array identification (mental counting and can be assessed with arrays consisting of four or more dots).

Counting Objects. The Counting Objects subtest assesses a student's ability to count a set of dots with one-to-one correspondence, and his/her ability to recognize that the last number counted in a sequence represents the overall quantity. The student is shown pages with arrays of dots and told to count the dots and say how many total there are.

Equal Partitioning. The Equal Partitioning subtest assesses the student's ability to recognize if two groups of objects are equal quantities, the ability to distinguish which group has more or less, and his/her ability to equally divide a set of manipulatives into two and three groups.

Composing. The Composing subtest assesses the student's ability to automatically compose pairs of five and ten. Composing requires an understanding about how numbers can be put together to create different numbers. The Composing assessment is designed to measure the student's ability to automatically compose numbers to five and ten and are therefore not allowed to use any counting strategies.

Quantity Discrimination - Most. The Quantity Discrimination Most subtest measures a student's ability to select the largest of four visually presented numerals between one and ten.

Quantity Discrimination - Least. The Quantity Discrimination Least subtest assesses a student's ability to select the smallest of four visually presented numerals between one and ten.

Numeral Identification. The Numeral Identification subtests assess the student's ability and automaticity at naming written numerals. There are two versions of Numeral Identification, one for kindergarten and one for first grade. The kindergarten version includes numerals up to 31. The first grade version includes numerals up to 120.

Number Sequence. The Number Sequence subtests assess oral counting and the understanding of the mental number line. Types of items include: Count Sequence (measures the student's ability to count forward, and also counting backward), Number After (items of various difficulty level which assess the understanding of "number after," "one more than," and "two more than."), Number Before (items of various difficulty level assess understanding of "number before," "one less than," and "two less than."), and Number Between (measures the student's understanding of the concept "between"). There are two versions of Number Sequence, one for kindergarten and one for first grade to represent expectations as outlined in national and state standards.

Decomposing - 1. The Decomposing subtest assesses the student's ability to put together (compose) and take apart (decompose) numbers by using "parts" and a "whole." Numbers are represented both as quantities (i.e., dots) and numerals. Composing and decomposing is a vital step for students towards understand base-ten reasoning and form strategies for addition facts summing greater than 10 (Baroody, 2006).

Place Value. The Grouping and Place Value subtest assesses the student's ability to correctly produce the numeral that corresponds with a set of base-10 blocks, and his/her ability to select the correct grouping of base-10 blocks when presented with a numeral.

Story Problems. The Story Problems subtest assesses the student's ability to represent and solve story problems involving addition and subtraction. The examiner verbally presents story problems (two with an accompanying image, and four without). Students are asked to both identify the correct written expression and solve the problem.

Verbal Addition. The Verbal Addition subtest assesses a student's ability and fluency in responding verbally to basic addition facts that are presented orally. Facts include partitions of 5 and 10, doubles, adding through 10, 10 plus, 2 digit by 1 digit addition to 20, and partitions of 20.

Verbal Subtraction. The Verbal Subtraction subtest assesses a students' ability and fluency in responding verbally to basic subtraction facts that are presented orally. Facts include combinations up to 5 and 10, doubles, subtraction through 10, subtracting 10, 2-digit by 1-digit subtraction, and subtracting from 20.

FAST™ earlyMath Administration

Administration time varies depending on which FAST™ earlyMath subtest is being administered. A timer is built into the software and is required for all subtests. For subtests that calculate a rate-based score (e.g., number correct per minute), the default test duration is set to one minute or 30 seconds depending on the measure. For those subtests that do not calculate a rate-based score (number correct), the default duration is set to open-ended. Each individual subtest can take approximately 1 to 3 minutes to complete. For universal screening, administration of the composite assessments takes approximately 5 to 7 minutes per student.

FAST™ earlyMath Scores and Scoring

Score Types

Each FAST™ earlyMath subtest produces a raw score. The primary score for each subtest is the number of items correct and/or the number of items correct per minute. These raw scores are used to generate percentile ranks and benchmarks. The best estimate of students' early mathematics skills is the FAST™ earlyMath composite score. The composite score consists of multiple subtest scores administered during a universal screening period. The FAST™ earlyMath composite scores were developed through regression and confirmatory factor analysis methodology as optimal predictors of spring broad math achievement in kindergarten and first grade. The FastBridge default FAST™ earlyMath Composite score includes different subtests depending on the grade level and screening period. Below is a table that shows the recommended subtests for each grade and screening period.

Table 40 Recommended subtests for each screening period for the Composite score

Grade	Fall Composite	Winter Composite	Spring Composite
Kindergarten	Match Quantity Number Sequence-K Numeral ID-K	Decomposing - K Number Sequence-K Numeral ID-K	Decomposing - K Number Sequence-K Numeral ID-K
First Grade	Decomposing-1 Number Sequence-1 Numeral ID-1	Decomposing-1 Number Sequence-1 Place Value	Decomposing-1 Place Value Story Problems

A select set of individual subtest scores were weighted to optimize the predictive relationship between FAST™ earlyMath and broad math achievement scores. Subtests were selected to assess skill in all three domains of early numeracy (i.e., Number, Relations, and Operations) in each screening period. The individual subtest scores were weighted to optimize the predictive relationship between FAST™ earlyMath and broad math achievement (see Table 41). The weighting is specific to each season. It is important to emphasize that the weighting is influenced by the possible score range, as well as the value of the skill at the point in the developmental trajectory. For example, Number Identification is an important skill with a score range of 0 to 60 (or more) numbers per minute. In contrast, Match Quantity has a score range from 0 to 12 and benchmarks are relatively low in value (e.g., benchmarks might be 7, 10 and 12, respectively). Because of both the score range *and* the relative value of Match Quantity to overall early math performance, the subtest score is more heavily weighted in the composite score. The high (H), moderate (M), and low (L) weights indicate the relative influence of a one point change in the subtest on the composite score. A one point change for an H weighting is highly influential. A one point change in an L weighting has low influence in the composite score.

The composite scores should be interpreted in conjunction with specific subtest scores. A variety of patterns might be observed. It is most common for students to perform consistently above or below benchmark on the composite and subtests; however, it is also possible to observe that a student is above benchmark on one or more measures, but below the composite benchmark. It is also possible for a student to be below benchmark on one or more subtests, but above the composite benchmark. Although atypical, this phenomenon is not problematic. The recommendation is to combine the use of composite and subtest scores to optimize the decision-making process. Overall, composite scores are the best predictors of future math success.

Table 41 Weighting Scheme for FAST™ earlyMath Composite Scores

	Kindergarten			First Grade		
FAST™ earlyMath Subtests	Fall	Winter	Spring	Fall	Winter	Spring
Match Quantity	M					
Decomposing (K)		H	H			
Numeral Identification	L	L	L	L		
Number Sequence	H	H	H	M	M	
Decomposing (1)				M	M	M
Place Value					M	M
Story Problems						H

Note. The weighting of subtests for the composite is represented above. H - high weighting, M - moderate weighting, L - low weighting.

Benchmark Scores

Benchmarks were established for FAST™ earlyMath to help teachers accurately identify students who are at risk or not at risk for academic failure. These benchmarks were developed from a criterion study examining FAST™ earlyMath assessment scores in relation to scores on the Group Mathematics Assessment and Classification Evaluation (GMADE) which is discussed later in this manual. Measures of classification accuracy were used to determine decision thresholds using criteria related to sensitivity, specificity, and area under the curve (AUC). Based on these analyses, the values at the 40th and 15th percentiles were identified as the primary and secondary benchmarks for FAST™ earlyMath, respectively. These values thus correspond with a prediction of performance at the 40th and 15th percentiles on the GMADE, a nationally normed assessment of early mathematics skills. Performance above the primary benchmark indicates the student is at low risk for long term mathematics difficulties. Performance between the primary and secondary benchmarks indicates the student is at some risk for long term mathematics difficulties. Performance below the secondary benchmark indicates the student is at high risk for long term mathematics difficulties. These risk levels help teachers accurately monitor student progress using the FAST™ earlyMath subtests. More information about how to understand benchmarks is found in the Benchmarks and Norms Interpretation and Use Guidelines.

Normative Scores

Normative scores for FAST™ earlyMath are intended to establish a baseline distribution for FAST™ earlyMath. The FAST™ earlyMath subtests have been normed on separate samples for kindergarten and girst grade. Normative scores for FAST™ earlyMath reflect typical performance by percentile range. FastBridge reports include normative data compared to the group (e.g., class), school, district and national distributions. These data characterize typical performance for each grade level, by season.

Student Strategies and Errors

In addition to raw scores, composite scores, benchmarks, and norms, FAST™ earlyMath provides educators and professionals a summary of student strategies and errors. A checklist allows administrators to easily note student strategies during testing. Common strategies/error types are displayed differently for each FAST™ earlyMath subtest and were determined by experienced test administrators, test developers, content specialists, math interventionists, and teachers. For example, possible strategies/errors for the Number Sequence subtest include inability to cross decade, count sequence partially omitted, stated “number before” for “number after”, and stated “number after” for “number before”. Taking notes and indicating student errors or strategies during test administration provides additional information for making educational decisions, designing interventions, and targeting instruction.

FAST™ earlyMath Construct Validity

Content-Related Validity Evidence

The test specifications for FAST™ earlyMath subtest relate directly to their evidence of content validity. Each subtest was designed with the intent to address specific educational standards (CCSS; see below) and domains of mathematics (Number, Relations, Operations; see above section on content for more evidence). FAST™ earlyMath subtests were also developed with substantial feedback from teachers, content specialists, and testing experts to ensure that the content of each subtest was instructionally relevant for students in each grade. The Common

Core State Standards for Mathematics were established in 2010. The standards alignment with FAST™ earlyMath subtests are presented in the table below.

Table 42 Alignment of CCSS and FAST™ earlyMath subtests

FAST™ earlyMath Subtest	CCSS
Match Quantity*	K.CC.3, 4, 5
Numeral Identification-K*	K.CC.3
Number Sequence-K*	K.CC.2
Decomposing-K*	K.NBT.1, K.OA.3,5
Subitizing	K.CC.5
Counting Objects	K.CC.4, 5
Equal Partitioning	K.CC.6
Composing	K.NBT.1, K.OA.3,5
Quantity Discrimination	K.CC.6, 7
Numeral Identification-1*	1.NBT.1
Number Sequence-1*	1.NBT.5, 1.OA.5
Decomposing-1*	1.OA.6, 8
Place Value*	K.CC.5
Story Problems*	1.NBT.2
Verbal Addition	1.OA.6
Verbal Subtraction	1.NBT.1, 1.OA.1

Note. Alignment to specific state standards are available upon request
*indicates subtests included in any of the composite scores

Criterion-Related Validity Evidence

Criterion-related validity of FAST™ earlyMath subtests was examined using two measures: (a) Measures of Academic Progress for Primary Grades (MAP), and (b) Group Mathematics Assessment and Classification Evaluation (GMADE™; Level R and 1). The MAP (Northwest Evaluation Association, 2005) is a classification and computer adaptive assessment designed to measure mathematics ability and progress. In kindergarten and first grade, MAP measures mathematics achievement in problem solving, algebra, computation, measurement, statistics, and number sense. MAP data were collected in the winter and spring for kindergarten, and in the fall, winter, and spring for first grade students at two schools located within one district in a suburban town as part of the district's regular screening process. The sample consisted of 221 students in kindergarten and 195 students in first grade. Approximately 50% of students were female. The ethnicity breakdown across both grades were 84-91% White, 3-4% Black, 2-4% Hispanic, 2-7% Asian/Pacific Islander, and 1-3% American Indian or Alaska Native. In addition, 7-12% of students were receiving special education services and 30-35% of students were eligible for free and/or reduced lunch.

The Group Mathematics Assessment and Classification Evaluation (GMADE™) is a norm-referenced assessment of mathematics skills. Kindergarten students were administered Level R forms, which are composed of two subtest measures: Concepts and Communication, and Process and Applications. First Grade students were administered Level 1 forms, which are composed of three subtest measures: Concepts and Communication, Process and Applications, and Operations and Computation. The assessment was administered to students in three school districts located in the Midwest near a metropolitan city. The GMADE was administered to classrooms in two testing sessions by trained graduate students. The assessment was administered two to four weeks after the spring FAST™ earlyMath assessment was administered. The sample consisted of 155 students in kindergarten and 170 students in first grade. Approximately 44% of students were female. The ethnicity breakdown across both grades were 74-80% White, 8-14% Black, 6% Hispanic, 4-5% Asian/Pacific Islander, and 1-3% American Indian or Alaska Native. In addition, 8-13% of students were receiving special education services and 29-34% of students were eligible for free and/or reduced lunch.

A summary of concurrent and predictive validity coefficients for the MAP and GMADE are presented for both kindergarten (Table 43) and first grade (Table 44). As discussed previously, the Composite measure in kindergarten and first grade demonstrated the highest level of criterion validity; suggesting that it is the best estimate of current and later broad mathematics performance.

Table 43 Concurrent and Predictive Validity for Kindergarten FAST™ earlyMath Subtests

Subtest	Criterion	Correlation Range	Correlation Median
Composite	MAP	.59 - .72	.69
	GMADE	.56 - .63	.56
Match Quantity	MAP	.39 - .57	.52
	GMADE	.34 - .47	.44
Numeral Identification	MAP	.43 - .64	.58
	GMADE	.46 - .51	.47
Number Sequence	MAP	.53 - .70	.61
	GMADE	.48 - .54	.49
Decomposing	MAP	.46 - .52	.47
	GMADE	.43 - .44	.44
Counting Objects	MAP	.36 - .37	.37
	GMADE	--	.25
Equal Partitioning	MAP	.40 - .45	.43
	GMADE	--	.39
Composing	MAP	.38 - .46	.42
	GMADE	--	.46
Quantity Discrimination - Most	MAP	--	.30
	GMADE	--	.40
Quantity Discrimination - Least	MAP	--	.37
	GMADE	--	.44

Table 44 Concurrent and Predictive Validity for First Grade FAST™ earlyMath Subtests

Subtest	Criterion	Correlation Range	Correlation Median
Composite	MAP	.59 - .69	.66
	GMADE	.67 - .69	.68
Numeral Identification	MAP	.42 - .57	.50
	GMADE	.45 - .61	.59
Number Sequence	MAP	.58 - .64	.64
	GMADE	.56 - .65	.60
Decomposing	MAP	.51 - .59	.56
	GMADE	.56 - .63	.59
Place Value	MAP	.35 - .63	.51
	GMADE	.55 - .58	.57
Story Problems	MAP	.34 - .49	.45
	GMADE	.48 - .52	.52

Reliability-Related Validity Evidence

Some FAST™ earlyMath subtests have fixed test lengths and are subject to typical internal consistency analyses. Some FAST™ earlyMath subtests, however, are timed. Internal consistency measures of reliability are inflated on timed measures because of the high percentage of incomplete items at the end of the assessment, which are those for which examinees did not respond (Crocker & Algina, 1986). As a solution to both illustrate the potential inflation and reduce it, estimates of internal consistency (reliability) were run on the items completed by approximately 16% of students, the items completed by 50% of students, and items completed by approximately 84% of students. Items not completed were coded as incorrect. For timed tests, a range of coefficients and the median are provided. For both fixed test length and inconsistent test length analyses, data were derived from a random sample of students from the FAST™ database from the 2013-14 academic year. Reliability coefficients are presented in Table 45 and Table 46.

Table 45 Internal Consistency for Kindergarten FAST™ earlyMath Subtests

		Alpha		Split-Half	
Subtest	N	Range	Median	Range	Median
Match Quantity	144	.74 - .80	.76	.76 - .87	.78
Numeral Identification	45	.89 - .97	.96	.93 - .98	.98
Number Sequence	598	--	.76	--	.87
Decomposing	601	--	.80	--	.83
Counting Objects	76	--	.83	--	.99
Equal Partitioning	73	--	.52	--	.75
Composing	603	--	.82	--	.83
Quantity Discrimination - Most	40	.05 - .61	.34	.23 - .78	.58
Quantity Discrimination - Least	40	.49 - .73	.64	.15 - .81	.39

Table 46 Internal Consistency for First Grade FAST™ earlyMath Subtests

		Alpha		Split-Half	
Subtest	N	Range	Median	Range	Median
Numeral Identification	45	.88 - .92	.90	.96 - .98	.98
Number Sequence	572	--	.85	--	.91
Decomposing	573	.79 - .88	.87	.82 - .90	.88
Place Value	91	.81 - .85	.82	.85 - .90	.87
Story Problems	138	--	.81	--	.88

Test-retest reliability is a measure of the degree to which scores are stable across a short time period when the items, students, and testing conditions are constant. Test-retest correlations between student scores on two different testing sessions were calculated in January 2014. Experienced graduate students who attended a 2-hour training collected the data. After the first testing session, the same data collectors then administered the FAST™ earlyMath subtests to the same individual students one to three weeks later. We collected Place Value subtest data using group administration procedures. Examiners tested students by classrooms to be more efficient and to limit distractions to the teachers' classrooms.

Reliability coefficients across subtests were moderate to high ranging from r = .62 to .87 in kindergarten and r = .71 to .91 in first grade (Table 47). In both grades, the test-retest reliability coefficients for the FAST™ earlyMath composite score was the highest, indicating that it may be the most stable measure of student early numeracy skills. The sample consisted of 39 students in kindergarten and 39 students in first grade. Approximately 65% of students in kindergarten and 49% of students in first grade were female. The ethnicity breakdown across both grades were 59-82% White, 10-21% Black, 5-8% Hispanic, 0-13% Asian/Pacific Islander, and 0-3% American Indian or Alaska Native. In addition, 3-5% of students were receiving special education services and 15-38% of students were eligible for free and/or reduced lunch.

Table 47 Test-Retest Coefficients for FAST™ earlyMath Assessments

Subtest	Correlation	N
Kindergarten		
Composite	.87	37
Match Quantity	.76	36
Numeral Identification	.85	38
Number Sequence	.80	35
Decomposing	.68	38
Counting Objects	.68	33
Equal Partitioning	.71	32
Composing	.62	38
Quantity Discrimination - Most	.73	38
Quantity Discrimination - Least	.73	38
First Grade		
Composite	.91	30
Numeral Identification	.91	36
Number Sequence	.71	36
Decomposing	.83	36
Place Value	.77	57
Story Problems	.77	39

Inter-rater reliability is a measure of the extent to which student scores are consistent across different examiners or scorers. Estimates of inter-rater reliability are based on two independent scorers, and the coefficients represent the level of agreement between examiners or raters. FAST™ earlyMath measures involve a small degree of subjectivity, given the clear scoring guidelines. However, unreliable scoring may be the result of timing discrepancies between administrators or scorers or differences in the interpretation of a student's response. A summary of inter-rater reliability is presented in Table 48. When low agreement was observed across raters it was typically the result of hesitation or discontinue rules. Overall, the median inter-rater reliability across kindergarten was high (93% to 100% agreement).

Table 48 Summary of Inter-Rater Reliability for FAST™ earlyMath Assessments

Subtest	Average Correlation	Median Correlation	Correlation Range	N
Kindergarten				
Match Quantity	.93	.93	.58 - 1.00	45
Numeral Identification	.96	.98	.83 - 1.00	45
Number Sequence	.99	1.00	.85 - 1.00	45
Decomposing	.95	1.00	.75 - 1.00	45
Composing	.95	1.00	.63 - 1.00	45
Quantity Discrimination - Most	.91	.93	.47 - 1.00	45
Quantity Discrimination - Least	.91	.93	.47 - 1.00	45
First Grade				
Numeral Identification	.98	.98	.92 - 1.00	45
Number Sequence	.97	1.00	.71 - 1.00	45
Decomposing	.93	1.00	.67 - 1.00	45
Place Value	.98	1.00	.75 - 1.00	62
Story Problems	.94	1.00	.67 - 1.00	45

Evidence for Use of FAST™ earlyMath as a Screening Tool

Evidence of classification accuracy was derived from a sample of 155 students in kindergarten and 170 students in first grade. Approximately 44% of students were female. The ethnicity breakdown across both grades were 74-80% White, 8-14% Black, 6% Hispanic, 4-5% Asian/Pacific Islander, and 1-3% American Indian or Alaska Native. In addition, 8-13% of students were receiving special education services and 29-34% of students were eligible for free and/or reduced lunch. The Group Mathematics Assessment Classification Evaluation (GMADE™) was used as the criterion measure for the ROC analyses that resulted in the classification accuracy information. High risk and some risk was defined by performance at the 15th and 40th percentile ranks on the GMADE. The following analysis was used to guide the development of benchmark scores.

The ROC curve analysis results for each grade for students at high risk and some risk using the Youden Index with the GMADE are presented in Table 49. Classification accuracy was calculated for up to three months per academic year (e.g., fall to spring, winter to spring, spring to spring). The optimal results were presented according to the largest AUC values. The decision thresholds were determined using criteria related to sensitivity, specificity, and area under the curve (AUC). Sensitivity is the proportion of students identified as being at risk in mathematics who were also found to be at risk on the GMADE. Specificity is the proportion of students who were identified as being not-at risk in mathematics in who were also found to be not at risk in mathematics on the GMADE. The overall correct classification accuracy rate is also provided, which suggests that overall proportion of students who were correctly identified.

Table 49 Classification Accuracy for FAST™ earlyMath

Grade	N	AUC	Sensitivity	Specificity	Classification
High Risk – Below 15th percentile					
Kindergarten	121	.88	.82	.83	.83
First grade	122	.88	.69	.97	.94
Some Risk – Below 40th percentile					
Kindergarten	121	.86	.81	.82	.82
First grade	122	.85	.79	.75	.76

Note. AUC = Area under the curve.

Evidence for Use of FAST™ earlyMath as a Progress Monitoring Tool

Alternate-form reliability represents the extent to which test results generalize to different item samples. To determine alternate-form reliability, students are tested with alternate, but equivalent forms of the test and scores from these forms are correlated. Alternate-form reliability was evaluated for a select number of FAST™ earlyMath subtests that are available for progress monitoring (i.e., Numeral Identification-K, Numeral Identification-1, Match Quantity, Number Sequence-K, Quantity Discrimination, Decomposing-K, and Place Value). Reliability coefficients are presented in Table 50.

FAST™ earlyMath assessments were administered to kindergarten (N = 64) and first grade (*N* = 68) students from one elementary school in a metropolitan area in the Midwest. Approximately 43-44% of students were female across grades. The ethnicity breakdown across both grades were 58-71% White, 15-23% Black, 9-11% Hispanic, 1-5% Asian/Pacific Islander, and 3-4% Multiracial. In addition, 3-6% of students were receiving special education services. Of the 20 progress monitoring forms available for each subtest, five forms were chosen at random for use in the study. Five forms were administered to students daily during a one-week period. Each student did not take five forms of every subtest, but rather was assigned one to two subtests to receive daily. If students were absent or unable to be tested one day, they were administered two forms simultaneously the next day.

Pearson correlations between each of the five forms for each FAST™ earlyMath subtests were calculated. The correlation represents the degree of association between the forms. Standard error of measurement (SEM) was also calculated and is reported (Table 50).

Table 50 Alternate Form Reliability for FAST™ earlyMath Subtests

		Coefficient		
Grade	N (range)	Range	Median	SEM (SD)
Kindergarten				
Number Identification	38-42	.88-.95	.92	4.51 (15.94)
Match Quantity	34-38	.44-.68	.61	1.72 (2.76)
Number Sequence	39-41	.67-.82	.75	1.30 (2.59)
Quantity Discrimination	39-42	.75-.85	.80	2.16 (4.84)
First Grade				
Number Identification	42-48	.89-.95	.91	3.12 (10.40)
Decomposing	41-44	.76-.87	.84	1.67 (4.18)
Place Value	39-43	.63-.86	.82	1.39 (3.28)

One-way, within-subjects (or repeated measures) ANOVAs were conducted to compare the effect of alternate forms (N = 5) across students (N = 41 to 46) on the number of mean correct responses within individuals for each progress monitoring form. There were non-significant effects for form across Match Quantity, $F(4, 114) = 1.66$, $p = 0.16$, Number Sequence, $F(4, 99) = 0.32$, $p = 0.86$, Numeral Identification-K, $F(4, 215) = 0.29$, $p = 0.89$, Numeral Identification-1, $F(4, 206) = 0.12$, $p = 0.98$, Decomposing, $F(4, 130) = 1.23$, $p = 0.30$, and Place Value, $F(4, 109) = 0.53$, $p = 0.71$. These results indicate that different forms did not result in significantly different mean correct response and provides evidence for the use of the alternate forms as progress monitoring tools.

FAST^{TM} Adaptive Math (FAST^{TM} aMath)

FAST™ aMath Purpose and Use

FAST™Adaptive Math (FAST™aMath) is designed to address issues of instructional relevance. FAST™aMath is a simple, efficient, computer-adaptive measure of both broad and component math skills from kindergarten through eighth grade. FAST™aMath is highly researched and based on the recommendations of the National Math Panel (2008) and National Common Core Standards (2010). FAST™aMath is designed to identify those students with deficits in math achievement in need of additional instruction and predict performance on state accountability measures. FAST™aMath includes fully automated administration and scoring of individualized assessments for purposes of universal screening and instructional leveling.

The objective of FAST™aMath is to extend and improve on the quality of currently available assessments. At present, research and validation of early intervention measures to screen for student proficiency in math is in its infancy (Gersten, Jordan, & Flojo, 2005). While some measures do show promise (i.e., Oral Counting, Number Identification, Quantity Discrimination, and Missing Numbers), these measures have insufficient reported reliability and validity evidence for use in early identification and formative assessment. The goal of FAST™aMath is instructional efficacy.

FAST™aMath is used two to three times a year to evaluate annual growth. FAST™aMath is designed to identify those students with deficits in math achievement in need of additional instruction and predict performance on state accountability measures. FAST™aMath is intended for use from kindergarten through eighth grade.

FAST™aMath Content Description

The Common Core State Standards Initiative (CCSS, 2010) resulted in national standards for math instruction, which may inform instructionally relevant assessments. The CCSS outline six domains—each comprised of clearly stated objectives—to be covered from Kindergarten to Fifth Grade: Counting & Cardinality, Operations & Algebraic Thinking, Number & Operations in Base Ten, Number & Operations—Fractions, Measurement & Data, and Geometry. These domains also map neatly onto the structure of instructional domains and focal points identified by the National Council for Teachers of Mathematics (2000). The domains of math achievement measured by FAST™aMath are directly linked with the CCSS and the six domains listed above are described in more detail below.

Counting & Cardinality

The Counting & Cardinality (CC) domain is confined to kindergarten and addresses students' basic knowledge of numbers. For example, students are expected to know number names, count to tell the number of objects, and compare numbers. The CC domain serves as a fundamental building block for the development of more complex math skills. For example, students may first be able to count a series of objects, later recognize the count of small groups without explicit counting, and still later, group large numbers of objects into meaningful groups (e.g., by tens) to arrive at a total.

Operations & Algebraic Thinking

The Operations & Algebraic Thinking domain extends from kindergarten through fifth grade and deals largely with the representation and solution of basic math facts. In kindergarten, students are expected to begin parsing out the differences in meaning between "addition" and "subtraction." As students progress they are expected to solve increasingly complex problems that may require addition, subtraction, multiplication, or division. Upon reaching grades four and five, students are expected to be familiar with the concepts of multiples and factors. In addition, students may be asked to interpret numerical expressions or analyze relationships using knowledge of the four operations developed in grades K-3.

Number & Operations in Base Ten

The Number & Operations in Base Ten domain extends from kindergarten through fifth grade and includes knowledge of place value and its applications. In kindergarten through second grade, students are expected to gain knowledge of place value and apply it to counting and basic operations involving addition and subtraction of whole numbers. At higher grade levels, students extend this knowledge to interpret the relationships between the digits of a single number. Students are eventually expected to do multi-digit operations involving whole numbers and decimal numbers.

Number & Operations - Fractions

The Number & Operations – Fractions domain is a part of the standards for students in third through fifth grade. In third grade, fractions are introduced to students as a new set of numbers in addition to whole numbers. Students are expected to understand fractions as partitions and compare fractions by reasoning about their size. Students at this level use math models involving equal parts or partitions to develop their understanding of fractions. For students in fourth grade, the ability to compare fractions is required. Students may also need to convert between decimal numbers and fractions. In fourth and fifth grade, students continue to extend their knowledge on operations of fractions with whole numbers and of fractions with fractions. By the end of fifth grade, students are expected to solve real-world problems with operations including multiplication, division, addition, and subtraction.

Measurement & Data

The Measurement & Data domain extends from kindergarten through fifth grade and addresses conversion of units, as well as the interpretation of data. At kindergarten, students are tested on classifying and comparing objects with measurable attributes. In first and second grades, students develop their ability to work with variables such as time, length, and volume. By fourth grade, students are required to convert various units in a measurement system. Through all grade levels, students are expected to develop an understanding of data on diagrams. By the end of fifth grade, students may be asked to complete tasks such as creating a line plot of data or using different operations to calculate measurements.

Geometry

The Geometry domain extends from kindergarten through fifth grade and covers knowledge ranging from comparison of shapes to the interpretation of coordinate planes. Through all grade levels, students are expected to build on their ability to classify and create shapes and solids by understanding the attributes of each category. As students reach higher grade levels, they are asked to work with more specific categories and more abstract figures. For example, students may be tested on the differences between an obtuse angle and an acute angle.

As noted, the representation of the CCSS domains differs by grade. In some cases, such as the Counting & Cardinality domain, standards from a domain are only present in one grade. The representation of grades by domain (as evidenced by the total number of standards) is provided in Figure 2.

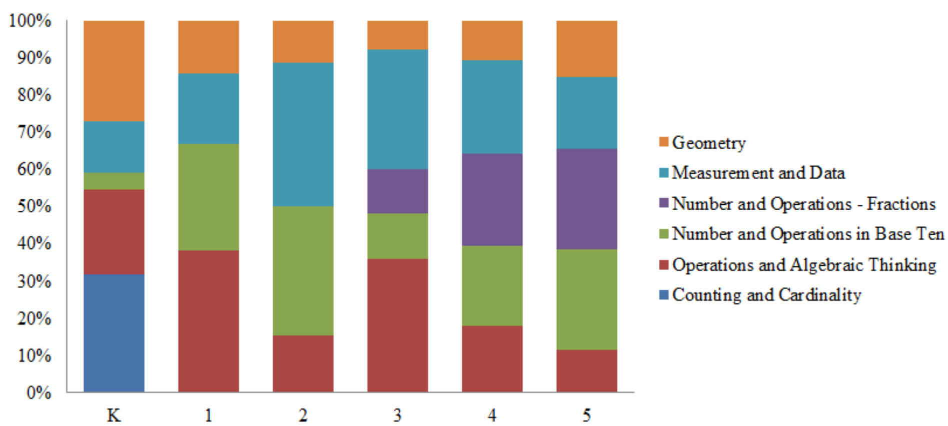

Figure 2. FAST™aMath Representation of Domains by Grade in the CCSS

FAST™aMath Item Development

FAST™aMath item development followed the process and standards presented by Schmeiser and Welch (2006) in the fourth edition of Educational Measurement (Brennan, 2006). Research assistants, teachers from each grade level (1-5), and content experts in math served as both item writers and reviewers. After items were written they were reviewed for feasibility, construct relevance, and content balance. A stratified procedure was used to recruit a diverse set of item writers from urban, suburban and rural areas. The item writers wrote, reviewed, and edited assessment materials.

FAST™aMath Computer Adaptive Test Development

FAST™aMath used a research-based skills hierarchy and unified construct of broad math achievement to establish an instructionally relevant assessment. The importance and emphasis on each component skill (domain) varies across children. Each assessment is individualized by the FAST™aMath software and built-in assessment algorithms. As a result, the information and precision of measurement is optimized regardless of whether a student functions at, above, or below grade level (i.e., same age and grade peers). The model presented in Figure 3 depicts how the assessment would likely function for the *typical* child; however, the grade labels and content balancing that are proposed in the a-priori model derive from the recommendations of expert panels and are subject to empirical evaluation and refinement.

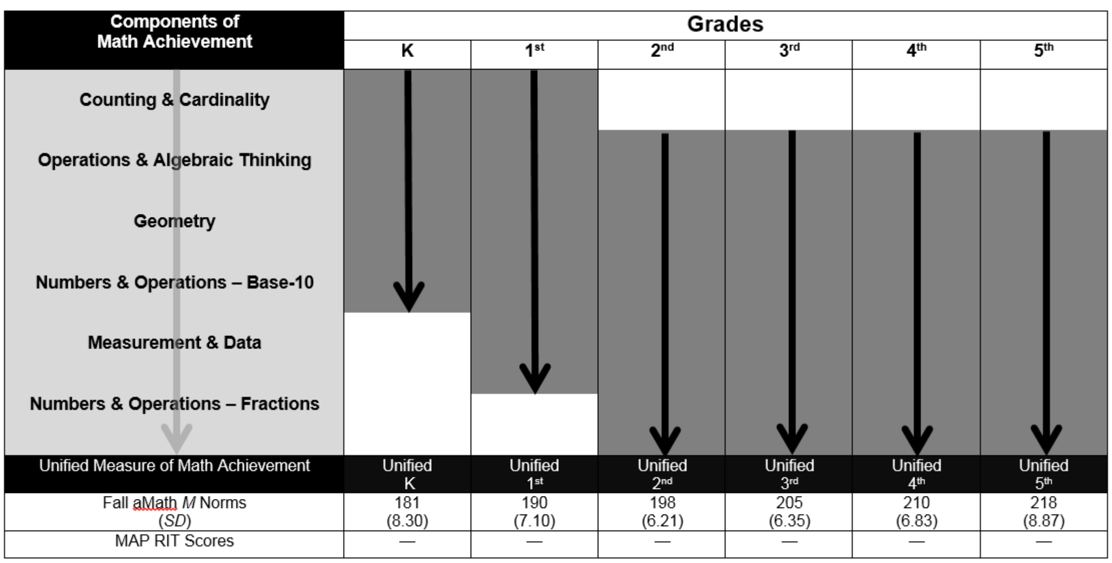

Figure 3. A Priori Model of the Construct: Specific and Unified Measurement of Math Achievement

Each FAST™aMath item was written to align with a specific domain and standard specified within the CCSS. Thus, in addition to the item-writing guidelines provided by Haladyna et al. (2007), the extent to which items assessed the intended standard was also considered during the item review process. The first round of precision analyses determined how many items had to be administered to attain an acceptable level of precision. Ideally the level of standard error across theta estimates approximates .20 (which equates to 3 on the FAST™aMath scale). The analysis to determine the optimum number of items per FAST™aMath administration was conducted in multiple steps. First a hybrid simulation (described below) was performed to generate responses for every participant. Next, simulations were conducted with the CATSim program to derive theta estimates and SE*m* estimates for each participant across five conditions. These conditions were based on the number of items that comprise each test and included: 20, 25, 30, and 40 items.

One issue with conducting a P-H simulation is the necessity that responses exist for every item from every single participant. Thus, the sparse data matrix of responses used to calibrate items for FAST™aMath could not be used in P-H simulations. After the hybrid simulations, computer adaptive testing (CAT) simulations with different test length termination criteria (20, 25, 30, 35 and 40 items) were conducted on both groups. Mean ability and standard error estimates were calculated for each administration (see Table 51). In addition, a polynomial function was fitted for each simulation and plotted. That is figures were generated that presented the average SE*m* across the range of FAST™aMath scores for each group. The table below presents the mean and standard deviation of FAST™aMath Scores and SE*m* for each fixed length CAT administration.

Table 51. Results of the FAST™aMath CAT Simulation

Items	Mean Score	SD Score	Mean SEM	SD SEM
20	192	16	3	2
25	192	16	3	2
30	192	16	3	2
35	192	16	3	2
40	192	16	3	2

An evaluation of Table 51 indicates that different length FAST™ aMath CATs are similarly efficient at measuring ability levels near 192. Ideally, SEM should approximate .20 (Weiss, personal communication), which translates to a value of 3 on the FAST™ aMath scale. Conditional SEM estimates are teased apart by grade in the table and figure below based on theta estimates from 30 item tests. From Table 52, it follows that the most precise estimates (across ability levels) are available for students in grades four and five. The largest difference in the mean SEM exists between kindergarten and first grade; however, this may be largely due to the small sample size available for students in kindergarten.

Table 52. Means and Standard Deviation for FAST™aMath Scaled Scores and SEM Values across Grades (Test Length = 30 Items)

		Scaled Score		SEM	
Grade	N	Mean	SD	Mean	SD
K	672	174.2413	9.100414	4.61075	2.576738
1	1415	181.7719	9.901944	3.382788	1.889782
2	1292	187.5579	10.8515	2.911892	1.539083
3	1542	194.8509	11.76315	2.683112	1.585994
4	1425	197.1197	13.57372	2.692789	1.124524
5	1323	207.0675	14.36281	2.550285	0.721794

Note. N = Sample Size; M = Mean; SD = Standard Deviation.

FAST™aMath Administration

FAST™aMath can be group administered in a computer lab setting, or a student can complete an administration individually at a computer terminal set up in a classroom or with the use of a tablet device. Test sessions for FAST™aMath typically last 15 to 30 minutes, depending on grade, student ability, and other factors. The test terminates on its own informing students they have completed all items. FAST™aMath administrations are typically completed following 30 items.

FAST™aMath Scores and Scoring

Score Types

Scores generated by the FAST™aMath computer-adaptive test (CAT) yield scores based on an IRT logit scale of -3 to 3. As emphasized previously, there are several shortcomings in reporting logit scores to educational professionals. Given these shortcomings, researchers for FAST™aMath chose to adopt an arbitrary scale for reporting theta estimates, like aReading. The FAST™aMath scale yields scores that are transformed from logits using the following formula:

$$Y = 200 + (15 * \text{Logit Score})$$

where, Y is the new ™aMath scaled score, and θ is the initial FAST™aMath logit theta estimate. Scores were scaled with a lower bound of 150 and a higher bound of 250. The mean value is 200 and the standard deviation is 15. Details on interpreting FAST™aMath scaled scores for instructional purposes is delineated in the following section.

Benchmark Scores

Benchmark scores for FAST™aMath are available for kindergarten through eigth grade at three time points: fall, winter, and spring. Benchmarks were established for FAST™aMath to help teachers accurately identify students who are at risk or not at risk for academic failure.

Normative Scores

Normative scores for ™aMath reflect typical performance by percentile range. These data characterize typical performance for each grade level.

Score Interpretations

FAST™aMath scaled scores have an average of 200 and standard deviation of 12 across the range of kindergarten to eighth grades. Scores should be interpreted with reference to the norms and benchmarks. In addition, the FAST™aMath has descriptions regarding the interpretation a student's scaled score with respect to mastered, developing, and future skill development. These are intended to help teachers better understand the developmental progression and student needs. FAST™ generates individual reports to describe the reading skills that a student has mastered, is developing, and will develop based on the student's scaled score.

FAST™aMath Construct Validity

Construct-Related Validity Evidence

Criterion-related validity of FAST™aMath was examined using the Measures of Academic Progress (MAP) and the Group Mathematics Assessment and Classification Evaluation (GMADE; Williams, 2004). The MAP is a computer-adaptive test for students in grades two through five. Each test requires an administration time of approximately 40 to 140 minutes. The GMADE is a comprehensive, norm-referenced assessment of mathematical skills. Students complete the GMADE using paper and pencils. The total time required to complete the GMADE varies from 60 to 90 minutes. One large Midwestern elementary school participated in the validity study (see Table 53). Teachers collected data on both occasions. Project personnel supervised FAST™aMath data collections. Project personnel analyzed data from one school in a Midwestern school district (*N* = 432).

Table 53. Demographics for Criterion-Related Validity Sample for MAP, GMADE™, and FAST™aMath

Category	Total	K	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
N	496	89	77	91	89	74	76
Gender (Male)	49%	40%	47%	48%	51%	53%	54%
White	88%	89%	90%	89%	85%	88%	90%
Black	6%	6%	3%	7%	6%	7%	8%
Hispanic	3%	2%	3%	2%	5%	3%	3%
Asian	2%	1%	4%	1%	5%	0%	0%
American Indian	1%	0%	0%	1%	0%	2%	0%
Special Education	15%	5%	10%	11%	24%	22%	18%
Free/Reduced Lunch	8%	2%	1%	0%	20%	18%	11%

Note. N = sample size.

Descriptive data for each measure are provided in Table 54 and correlations are available in Table 55.

Table 54. Mean, Standard Deviation, and Sample Size of FAST™aMath Scaled Scores

	FAST™aMath			GMADE			MAP		
Grade	N	M	SD	N	M	SD	N	M	SD
K	89	189	6	81	106	15	89	167	14
1	77	198	6	72	101	14	77	186	8
2	91	206	7	67	102	17	91	198	11
3	89	211	7	86	109	13	89	212	11
4	74	218	12	60	105	15	74	223	11
5	76	224	12	42	106	16	53	229	14

Note. M = Mean; SD = Standard Deviation; N = Sample Size.

Table 55. Correlation Coefficients between MAP and FAST™aMath Scaled Scores

Grade	N	MAP	N	GMADE™
K	89	.76	81	.62
1	77	.71	72	.66
2	91	.81	67	.67
3	89	.76	86	.76
4	74	.84	60	.67
5	76	.88	65	.84

Overall, the strongest correlations were observed between FAST™aMath and MAP scores. This is likely due to the similar nature and purpose of the two assessments. Correlations between FAST™aMath and the GMADE were slightly lower but generally provide adequate criterion-related validity evidence. In both cases, correlations varied across grades, with the strongest correlations occurring in fifth Grade. Recently, additional data has been collected to support criterion-related evidence of FAST™aMath.

Table 56. Criterion Validity Evidence of May FAST™aMath and Spring MCA in Math

Grade	N	FAST™ aMath M (SD)	MCA M (SD)	Correlation	Cut Score	AUC	Sensitivity	Specificity
Somewhat at Risk (Does Not Meet or Partially Meets Standards)								
3	155	210.55 (8.83)	354.70 (16.50)	.87	209.5	.95	.86	.88
4	90	217.31 (11.05)	456.83 (17.57)	.87	214.5	.94	.83	.85
5	67	226.79 (11.97)	550.85 (13.21)	.88	226.5	.98	.93	.92
High Risk (Does Not Meet Standards)								
3	155	210.55 (8.83)	354.70 (16.50)	.87	205.5	.97	.89	.88
4	90	217.31 (11.05)	456.83 (17.57)	.87	208.5	.98	.94	.92
5	67	226.79 (11.97)	550.85 (13.21)	.88	220.00	.94	.85	.87

Reliability-Related Validity Evidence

Given the adaptive nature of FAST™aMath test, a proxy for internal consistency and alternate forms is provided by Samejima (1994), based on the standard error of measurement of an instrument. Using this proxy, internal consistency and alternate forms reliability coefficient for FAST™aMath is approximately .95 (based on over 2,000 students).

Evidence Related to Bias

Bias analyses of a sample of the items that comprise FAST™aMath were conducted using data collected during the 2016-17 and 2017-18 academic years. Data for each year were analyzed separately. There were sufficient data to examine bias in relation to race/ethnicity. The race/ethnicity group comparisons examined were White versus Black, White versus Hispanic, White versus Asian, and White versus Native American. The results indicated that there is no or negligible DIF for all items examined in both years for all the race/ethnicity comparisons.

Bias was assessed using the logistic regression procedure for detection of uniform and non-uniform differential item functioning (DIF). The advantages of using the logistic regression procedure for DIF detection include being a model-based approach and having the capability to detect both uniform and non-uniform DIF with adequate and equal power; however, the procedure also tends to inflate Type I error rates. As such, an effect size measure developed by Jodoin and Gierl (2001) was computed and evaluated in addition to statistical significance. Jodoin and Gierl present a four-category framework for interpreting the effect size measure, where the four categories are indicative of no, negligible, moderate, and severe DIF.

Evidence for Use of FAST™aMath as a Screening Tool

FAST™aMath classification accuracy was derived from a sample of 432 students in kindergarten through fifth grade from one large Midwestern elementary school. In the sample, 89 students were in kindergarten, 77 in first grade, 91 in second grade, 89 in third grade, 74 in fourth grade, and 76 in fifth grade. Gender of the sample was approximately 49% male and 51% female. Approximately 88% of the students in the sample were White, 6% Black, 3% Hispanic, 2% Asian, and 1% American Indian. In addition, approximately 8% of students were receiving free or reduced lunch and 15% were receiving special education services. Cut scores for FAST™aMath to predict students at risk and somewhat at risk for math difficulties were developed for the GMADE. Categories were defined as students scoring below the 40th and 10th percentiles, respectively. Students completed the GMADE as well as the FAST™aMath. Teachers collected data on both occasions. Project personnel supervised the FAST™aMath data collections.

Table 57 below presents the ROC curve analysis results for each grade for students at high risk and somewhat at risk using the Youden Index with the GMADE. ROC curves across grades and risk levels were far from the diagonal line indicating that FAST™aMath predicts math difficulties at much greater level than chance. AUC statistics were variable but generally adequate at the 40th percentile and 10th percentile. AUC values were especially high for 5th grade Somewhat at Risk (.97) and 3rd grade At Risk (.98). Sensitivity was somewhat larger for each grade when determining students at High Risk compared to students Somewhat at Risk. In general, NPP was very high (.84 to .90) while PPP was low (.29 to .92). This is likely due to the low base rate for GMADE scores (very few students were at risk).

Table 57. Classification Accuracy Statistics for FAST™aMath and the GMADE

Grade	N	FAST™aMath Cut Score	Sensitivity	Specificity	PPP	NPP	AUC
High Risk - 10th Percentile GMADE							
K	81	187	0.71	0.75	0.12	0.98	0.75
1	72	190	0.95	0.83	0.62	0.98	0.83
2	67	202	0.88	0.80	0.53	0.96	0.92
3	86	201	0.95	1.00	0.33	1.00	0.98
4	60	206	0.95	0.75	0.50	0.98	0.88
5	42	218	0.84	0.80	0.40	0.97	0.88
Somewhat At Risk - 40th Percentile GMADE							
K	81	187	0.72	0.80	0.65	0.98	0.81
1	72	196	0.79	0.90	0.67	0.98	0.88
2	67	206	0.80	0.92	0.67	0.98	0.82
3	86	207	0.96	1.00	0.43	1.00	0.89
4	60	215	0.67	0.80	0.64	0.97	0.74
5	42	225	0.85	0.88	0.81	0.97	0.97
Criterion			.70	.70	-	-	.85

Note: N = sample size; PPP = positive predictive power; NPP = negative predictive power; AUC = area under the curve.

A similar pattern of results emerged when predicting performance on the MAP (Table 58). However, only cut scores for those students "Somewhat At Risk" were identified. As with the GMADE, NPP was much higher than PPP when predicting MAP scores. This is also likely due to the low base rate of risk status among students in the sample.

Table 58. Classification Accuracy Statistics for FAST™aMath and MAP

Grade	N	FAST™aMath Cut Score	Sensitivity	Specificity	PPP	NPP	AUC
Somewhat At Risk - MAP Specified Cut Score							
K	89	188	0.84	0.91	0.77	0.94	0.91
1	77	193	0.84	1.00	0.20	1.00	0.93
2	91	196	0.97	1.00	0.62	1.00	0.98
3	89	208	0.77	1.00	0.21	1.00	0.91
4	74	206	0.91	1.00	0.50	1.00	0.98
5	53	220	0.86	1.00	0.62	1.00	0.94

FAST™ Social, Academic and Emotional Behavior Risk Screener

FAST™ SAEBRS Purpose and Use

The FAST™ Social, Academic and Emotional Behavior Risk Screener (FAST™ SAEBRS**)** was designed to be a brief and contextually relevant screener of students risk for emotional and behavioral problems. The FAST™ SAEBRS evaluates general student behavior, as well as behavior within the social, academic, and emotional domains. The FAST™ SAEBRS is a brief behavior rating scale. A teacher completes the FAST™ SAEBRS for an individual student with whom the teacher has a history of interactions. Ratings correspond to the frequency with which the teacher has observed various maladaptive and adaptive behaviors in the previous month. It is estimated that it takes approximately 3 to 5 minutes to complete the measure for each student.

FAST™ SAEBRS data can be used to assess students' general, social, academic, and emotional behaviors. Data can also be useful in program evaluation and in determining how students may be best supported at Tier 1. For instance, the data can be used to indicate whether a school should invest in the support of teacher classroom management practices, given the prevalence of social behavioral concerns, or in the instruction of academic enabling skills given the noted extent of academic behavioral difficulties.

FAST™ SAEBRS can be used to identify students who are at risk for general, social, academic, and emotional behaviors at least three times a year. By evaluating in which of the three specific domains a student is at risk, educators may determine what type of supports are most appropriate and which problem behaviors should be prioritized through intervention. For instance, if a student is only at risk for emotional problems, then a school may decide to target emotional behaviors via the application of social-emotional learning programs. The FAST™ SAEBRS is designed for universal screening of students at risk for social-emotional and behavior problems in kindergarten through twelfth grade.

FAST™ SAEBRS Content Description

FAST™ SAEBRS includes items from three domains. Each domain is defined as follows. Social Behavior (6 items) is defined as behaviors that promote (e.g., social skills) or limit (e.g., externalizing problems) one's ability to maintain age appropriate relationships with peers and adults. Academic Behavior (6 items) is defined as behaviors that promote (e.g., academic enablers) or limit (e.g., attentional problems) one's ability to be prepared for, participate in, and benefit from academic instruction. Finally, Emotional Behavior (EB; 7 items) is defined as actions that promote (e.g., social-emotional competencies) or limit (e.g., internalizing problems) one's ability to regulate internal states, adapt to change, and respond to stressful/challenging events. In accordance with the principles of prevention science, each factor corresponds to various risk and protective factors suggested by developmental psychological research to predict the development of emotional/behavioral disorders.

FAST™ SAEBRS Content Development

The FAST™ SAEBRS was developed for use in universal screening for behavioral and emotional risk. The measure falls within a broad class of highly efficient tools, suitable for teacher use in evaluating and rating all students on common behavioral criteria (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). The FAST™ SAEBRS is grounded within a conceptual model, which states that a student's success in school is not only related to his or her academic achievement, but also success within multiple behavioral domains. Research suggests the FAST™ SAEBRS may be used to evaluate student functioning in terms of general behavior, as assessed by a broad Total Behavior scale. Research further suggests the SAEBRS may be used to evaluate student behavior within multiple inter-related narrow domains, as assessed by the Social Behavior, Academic Behavior, and Emotional Behavior subscales (Kilgus, Chafouleas, & Riley-Tillman, 2013; Kilgus, Eklund, von der Embse, & Taylor, 2014; Kilgus, Sims, von der Embse, & Riley-Tillman, 2014).

FAST™ SAEBRS Administration

General or special education classroom teachers serve as the most appropriate FAST™ SAEBRS informants. Teachers chosen to complete the FAST™ SAEBRS should have interacted extensively with each target student during the month FAST™ SAEBRS preceding FAST™ SAEBRS completion. A teacher may complete the SAEBRS following an approximately 30-minute training session available via online training modules. It is estimated that it takes approximately 1 to 3 minutes to complete the measure for each student. Teachers complete the FAST™ SAEBRS once for each student in their classroom. Therefore, if 15 students are enrolled in a teacher's classroom, the teacher will fill out the FAST™ SAEBRS 15 times.

Once a teacher is ready to rate a student, he/she should complete the FAST™ SAEBRS subscales deemed by the school to be pertinent to their decision making. To complete each FAST™ SAEBRS item, the teacher indicates how frequently the student in question has displayed each behavior (as described within each item) during the previous month. The teacher is to ONLY consider the behavior exhibited by the student during the month prior to FAST™ SAEBRS completion. No other behaviors outside of this period should be taken into consideration during item completion. It is common for teachers to request a definition of the behaviors represented within each FAST™ SAEBRS item. For instance, many seek additional clarification regarding what should be considered a 'temper outburst.' However, as part of standard administration, FAST™ SAEBRS users are not to be provided with such definitions. Rather, teachers are to use their best judgment in considering what actions are representative of each behavior.

A nomination form is used to determine which students who are at risk and would benefit most from being evaluated by the FAST™ SAEBRS. These students should frequently display inappropriate social, academic, and or emotional behaviors and rarely display appropriate behaviors. The teacher selects 3 to 5 of his/her students who fit the definition for risk for social problems during the previous month.

FAST™ SAEBRS Scores and Scoring

To score the FAST™ SAEBRS, negatively worded items are first reverse scored. Item scores are then then summed within each subscale and the overall scale. FAST™ reports a student's overall performance on each FAST™ SAEBRS scale as a sum of item scores within each scale. Scores range from 0-18 for Social Behavior, 0-18 for Academic Behavior, 0-21 for Emotional Behavior, and 0-57 for Total Behavior. The Total Behavior score is calculated by summing the Social Behavior, Academic Behavior, and Emotional Behavior subscale scores. Although FAST™ SAEBRS scores can often be used as continuous variables, it is sometimes convenient to classify scores as at risk and not at risk. Using the ranges in Table 59, subscale and scale score can be dichotomized in terms of risk categories within the domains. FAST™ SAEBRS risk ranges have been established based on comparison of the FAST™ SAEBRS to multiple criterion gold standard behavior rating scales, including the Social Skills Improvement System (Gresham & Elliot, 2008) and the Behavioral and Emotional Screening System BASC-2 (Kamphaus & Reynolds, 2007).

Table 59 FAST™ SAEBRS Score Ranges for Risk and No Risk

SAEBRS Scale/Subscale	Not At Risk	At Risk
General Behavior	37-57	0-36
Social Behavior	13-18	0-12
Academic Behavior	10-18	0-9
Emotional Behavior	18-21	0-17

FAST™ SAEBRS Construct Validity

Current research for the FAST™ SAEBRS is indicative of the screener's content validity and concurrent criterion-related validity. The FAST™ SAEBRS was originally subjected to expert content validation, within which experts considered the extent to which items corresponded to the various constructs, were relevant to those constructs, and were fair/appropriate/unbiased indicators of each construct. The FAST™ SAEBRS has gone on to be refined and validated with large, heterogeneous samples that are representative of the broader US population. Data collection is ongoing in elementary, middle, and high schools throughout the United States.

Content-Related Validity Evidence

To ensure content validity, FAST™ SAEBRS developers met to discuss the purpose of the FAST™ SAEBRS, as well as the variables to which it should correspond. Current literature regarding behavioral assessment measures, direct behavior ratings, and positive behavioral interventions and supports were considered and informed FAST™ SAEBRS item development. Once a pool of items was generated, three school psychology professors and one doctoral student served as content experts, completing a series of steps as part of a content validation process. First, experts placed each item in the latent variable category to which they felt the item most closely corresponded. Second, using a 3-point Likert scale, experts indicated how certain they were of their category placement (1 = not sure, 2 = somewhat sure, and 3 = very sure). Third, experts rated how relevant each item was to their chosen category using a 3-point Likert scale (1 = not relevant, 2 = somewhat relevant, 3 = highly relevant). Fourth, experts provided open-ended feedback regarding which items should be removed, added, or revised.

The Content Validity Index (CVI) and the Factorial Validity Index (FVI) were used to determine which items should be removed based on rater feedback. The CVI represented expert opinions regarding item representativeness, and was calculated for each item by dividing the number of experts who rated the item as somewhat or highly relevant, and dividing this number by the overall number of experts. The FVI represented the extent to which experts assigned each item to its appropriate category, and was computed for each item by dividing the number of experts who assigned the item in accordance with expectations by the overall number of experts. All FAST™SAEBRS items exhibited CVI and FVI values equal to or greater than .80.

Criterion-Related Validity Evidence

FAST™ SAEBRS criterion-related validity has been supported by several investigations. Results support the correspondence between each broad and narrow FAST™ SAEBRS scale and multiple gold standard criterion behavior rating scales, including the Social Skills Improvement System (SSIS; Gresham & Elliott, 2008) and the Behavioral and Emotional Screening System (BESS: Kamphaus & Reynolds, 2007). Table 60 and Table 61 show the correlations between the FAST™ SAEBRS and the criterion measures, representing concurrent and predictive criterion-related validity coefficients. The following measures were used as criterion measures: (a) Social Skills Improvement System (SSIS; Gresham & Elliott, 2008), (b) BASC-2 BESS, (c) Student Risk Screening Scale (SRSS; Drummond, 1994), and (d) Student Internalizing Behavior Screener (SIBS; Cook et al., 2011).

Table 60 Criterion Validity Evidence for FAST™ SAEBRS Total Score

Type of Validity	Test or Criterion	N	Correlation	Confidence Interval
Concurrent	SSIS-Social Skills	276	.88	.85, .90
Concurrent	SSIS-Problem Behaviors	276	-.89	-.91, -.86
Concurrent	SSIS-Academic Competence	276	.61	.53, .68
Concurrent	BESS	567	-.93	-.94, -.92
Concurrent	SRSS	346	-.84	-.87, -.81
Concurrent	SIBS	346	-.67	-.73, -.61
Predictive	BESS	1243	-.76	-.73, -.78
Predictive	BESS	1243	-.74	-.71, -.76

Table 61 Concurrent Criterion-Related Validity for FAST™ SAEBRS

Measure	Grade	Criterion	N	Correlation
Social	Elementary	SSIS	243	.82
Social	Elementary	BESS	219	.80
Social	Middle	BESS	359	.83
Academic	Elementary	SSIS	243	.76
Academic	Elementary	BESS	219	.90
Academic	Middle	BESS	359	.89
Emotional	Elementary	BESS	219	.71
Emotional	Middle	BESS	359	.73
General	Elementary	SSIS	243	.88
General	Elementary	BESS	219	.93
General	Middle	BESS	359	.95

The presented data speak to the validity of nomological net upon which the FAST™ SAEBRS is founded. The FAST™ SAEBRS theoretical framework specifies that the measure should be capable of predicting a student's broader social-emotional and behavioral functioning. It further specifies the measure should be capable of predicting a student's behavior within the social, academic, and emotional domains. Evidence to date has supported this, as evidence by: (a) concurrent and predictive correlations with the BESS, an indicator of broad and general functioning, (b) correlations with the SSIS-Social Skills and SRSS, indicators of student social competence and externalizing behavior, respectively (both of which are theoretically captured through the FAST™ SAEBRS Social Behavior subscale), (c) correlations with the SSIS-Academic Competence scale, an indicator of student academic functioning (which is theoretically captured through the FAST™ SAEBRS Academic Behavior subscale), and (d) correlations with the SIBS, an indicator of student internalizing behavior (which is theoretically captured through the FAST™ SAEBRS Emotional Behavior subscale). When taken together, existing validity evidence supports all elements of the theoretical framework upon which the FAST™ SAEBRS is founded.

Reliability-Related Validity Evidence

Multiple statistics were used in evaluating FAST™ SAEBRS internal reliability. First, a series of omega coefficients were used as part of a model-based approach to the evaluation of FAST™ SAEBRS Total Behavior scale internal reliability. Second, alpha coefficients were used as a separate, non-model-based approach. Internal reliability is considered relevant given the presumption that all FAST™ SAEBRS items are related to the broader construct of general behavioral functioning. Consideration of a model-based coefficient like omega is considered particularly relevant given the presumption that the FAST™ SAEBRS is founded upon a bifactor model, wherein all items are related to both the general behavior factor and one of three narrow factors.

Internal reliability was evaluated across two studies. The first study was conducted in four urban elementary schools (K-5) located in the Midwestern United States. All general education teachers in each school chose to participate in this study. The teachers screened all students in their classroom, resulting in a sample of 68 teacher participants and 1,243 students. The sample was characterized by a diverse student population regarding ethnicity, including sizeable subsamples of White (54.5%), Black (28.6%), Hispanic (5.3%), and Multiracial (8.4%) students. The free/reduced-price lunch rate across the four schools was equal to 65.1%. Second, alpha coefficients were evaluated as part of Kilgus, Eklund, von der Embse, Taylor, and Sims (2016). This study was conducted with 567 elementary students (52.9% female) and 34 classroom teachers. The sample was characterized by a diverse student population, including sizeable subsamples of White (50.1%), Black (34.4%), Hispanic/Latina(o) (11.3%), and multiracial (3.7%) students.

Omega (ω) coefficients represent the proportion of variance to all factors common to an item set of interest. Hierarchical omega (ω_H) coefficients represent the proportion of variance attributable to a factor after controlling for all other factors. These latter statistics are particularly informative when examining measures corresponding to bifactor structures (such as *the FAST™ SAEBRS*), as items are presumed to be multidimensional and driven by both general and specific factors. Beyond omega, coefficient alphas were also calculated in evaluating FAST internal reliability.

Coefficient alphas are presented below in Table 62. Omega, amongst the sample of 1,243 elementary students was .98. Hierarchical omega was .87.

Table 62 Internal Consistency of FAST™ SAEBRS

Measure	Grade	N	Coefficient Alpha
Social	Elementary	243	.89-.94
Social	Middle	359	.93
Social	High	488	.89
Academic	Elementary	243	.90-.92
Academic	Middle	359	.92
Academic	High	488	.93
Emotional	Elementary	219	.83
Emotional	Middle	359	.77
Total	Elementary	243	.93
Total	Middle	359	.94
Total	High	488	.93

A series of correlation analyses were used to evaluate the association between FAST™ SAEBRS data administered at two different time points within two weeks. Interest in test-retest reliability was founded in the assumption that most students within a school should maintain their social-emotional and behavioral risk status across the school year (Dever, Dowdy, Raines, & Carnazzo, 2015). Accordingly, it was anticipated there should be some degree of consistency in scores. With that said, it was expected such consistency would be tempered by the inherent variability of behavior and the delivery of intervention and supports to a subsample of students in the school.

Test-retest reliability was evaluated with internal FastBridge Learning data. This sample included 53 students (42% female), all of whom were evaluated with the FAST™ SAEBRS twice within 14 days. In terms of ethnicity, the sample consistent of 89% Black, 6% White, 2% Hispanic/Latina(o), and 4% multiracial students. FAST™ SAEBRS scores at Time 1 spanned the range of expected performance, ranging from 13 to 57. This suggested that students from all performance levels (i.e., low, moderate, and high risk) were represented within this sample, despite its restricted size.

Pearson product-moment correlation (*r*) coefficients were used to evaluate the association between FAST™ SAEBRS administrations. In a sample of 53 elementary students, Test-retest reliability was 1.00.

Evidence for FAST™ SAEBRS inter-rater reliability is provided in Table 63. Data were collected only for high school. All coefficients represent Pearson product-moment correlation coefficients.

Table 63 Inter-Rater Reliability for FAST™ SAEBRS

Measure	Grade	Correlation	N
Social	High School	.41	488
Academic	High School	.47	488
General	High School	.48	488

Evidence Related to Bias

Further, to examine the validity and bias across different groups of students, multi-group confirmatory factor analysis (MG-CFA) was used to examine measurement equivalence/invariance for FAST™ SAEBRS (Pendergast, von der Embse, Kilgus, & Eklund, 2017). Specifically, analyses considered the extent to which the FAST™ SAEBRS (inclusive of only Social Behavior and Academic Behavior items) was invariant across ethnic categories. Participants from two racial groups (White n = 412, and Black n = 323) were included in analyses. Participants from other racial groups were excluded because the sample sizes were too small (<100).

CFAs were conducted in Mplus 6.2 using WLSMV estimation (Beauducel & Herzberg, 2006). Overall model fit was evaluated based on the RMSEA and the CFI (Kline, 2010; Tanaka, 1993). Criteria for evaluating minimally acceptable model fit were established a priori: RMSEA values ≤ 0.08 and CFA values ≥ 0.90 (Browne & Cudeck, 1993; Hu & Bentler, 1995; Markland, 2007). The ME/I of the two-factor SAEBRS structure was assessed by applying increasingly restrictive equality constraints across groups to examine (a) configural invariance, (b) metric invariance, and (c) scalar/threshold invariance. Nested models (i.e., models with increasingly restrictive invariance tests) were compared using the change in SB $\chi2$ (ΔSB $\chi2$), change in (ΔCFI), and change in RMSEA (ΔRMSEA) values. Within the current study, each nested model was compared to its parent model, the latter of which possessed increasingly restrictive invariance specifications. As the models grew more restrictive, non-significant $\Delta\chi2$ ($p > 0.05$), ΔCFI < 0.01 (Cheung & Rensvold, 2002), and ΔRMSEA < 0.015 indicated that the more restrictive model had a comparable fit to the data as less restrictive one (Byrne, 2011; Meade et al., 2008; Satorra & Bentler, 2001).

In the first step, configural invariance was established. Fit indices for the configural model fell within specified ranges (CFI = 0.993; RMSEA = 0.077). Next, a metric invariance model was tested wherein factor loadings were constrained to be equal across racial groups. The model had adequate fit based on the fit criteria (CFI = 0.994; RMSEA = 0.069). The change-in-model fit indices suggested that the fit of the metric invariance model was not significantly worse, and, in fact, was slightly better relative to the configural invariance model (Δχ2 was non-significant, ΔCFI was < 0.01, and was ΔRMSEA < 0.015). Subsequently, a scalar/threshold invariance model was tested whereby factor loadings and thresholds were constrained to be equal across groups. The model had adequate fit (CFI = 0.994; RMSEA = 0.062), and the change-in-model fit indices indicated that the fit of the scalar/threshold model was not significantly different from that of the metric model (Δχ2 was non-significant, ΔCFI was < 0.01, ΔRMSEA was < 0.015). Therefore, scalar/threshold invariance was supported.

Evidence for Use of FAST™ SAEBRS as a Screening Tool

Receiver operating characteristic (ROC) curve analysis has been used across multiple studies to identify FAST™ SAEBRS cut scores (Kilgus et al., 2013; Kilgus, Eklund, et al., 2014). FAST™ SAEBRS cut scores are used in applied decision making to differentiate between students who are at risk or not at risk for behavioral and emotional problems. Cut scores were selected using a linear algorithm, which prioritized sensitivity greater than or equal to .80 and specificity at or above .70 (Kilgus et. al. 2013). The overall classification accuracy of each FAST™ SAEBRS scale has also been evaluated using area under the curve (AUC) statistics. AUC statistics correspond to the likelihood that a randomly selected at risk student will have a lower FAST™ SAEBRS score than a randomly selected not at risk student. Researchers have suggested that AUC values between .50 and .70 be considered low, .70 to .90 moderate, and .90 to 1.00 high (Steiner & Cairney, 2007).

FAST™ SAEBRS classification accuracy was derived from samples from elementary schools (N = 243 and N = 219) and a sample of 359 students from middle schools. Table 64 and Table 65 show the classification accuracy results using two different criterion measures collected across two investigations. Results are overall promising, with sensitivity, specificity, and AUC values falling in acceptable ranges.

Table 64 Classification Accuracy Statistics for FAST™ SAEBRS and SSIS

Grade	Measure	N	Cut Score	Sensitivity	Specificity	AUC
Elementary	Social	243	12	.87	.83	.93
Elementary	Academic	243	9	.84	.84	.90
Elementary	Total	243	23	.93	.94	.96

Note: N = sample size; AUC = area under the curve.

Table 65 Classification Accuracy Statistics for FAST™ SAEBRS and BESS

Grade	Measure	N	Cut Score	Sensitivity	Specificity	AUC
Elementary	Social	219	12	.87	.83	.93
Elementary	Academic	219	9	.84	.84	.90
Elementary	Emotional	219	17	.93	.94	.96
Elementary	Total	219	36	.88	.75	.88
Middle	Social	359	12	.87	.86	.94
Middle	Academic	359	9	.90	.77	.94
Middle	Emotional	359	17	.89	.70	.88
Middle	Total	359	36	.94	.91	.98

Note: N = sample size; AUC = area under the curve.

FAST™ Developmental Milestones

FAST™ DevMilestones Purpose and Use

The FAST™ Developmental Milestones (FAST™ DevMilestones) measure is designed to assess both unified and component skills associated with broad developmental expectations for kindergarten success. FAST™ DevMilestones is intended to enable screening and progress monitoring across six domains of early development (language, literacy, and communications; cognitive development; social emotional development; creativity and the arts; approaches to learning; and physical and motor development). FAST™ DevMilestones provides domain-specific assessments of component skills over time, as well as a general estimate of overall readiness for kindergarten success. The objective of the FAST™ DevMilestones is to improve on the quality of available assessments by providing an efficient tool that is closely aligned to learning standards that span the transition from preschool to kindergarten, while reflecting key developmental skills empirically linked to readiness to learn.

FAST™ DevMilestones is an observation-based rating scale that is completed by teachers and other adults to identify skill proficiency relative to theoretically and empirically derived criteria of expected performance across the duration of the kindergarten year. For each item, ratings are provided with observable indicators demonstrating a hierarchy of skill performance ranging in complexity and sophistication.

FAST™ DevMilestones is intended for use as a universal screener and progress monitoring tool, allowing for differentiation between students who are (a) developing at a typical rate, evidenced by achieving milestones at expected times, or (b) developing at an atypical rate, evidenced by not achieving certain milestones as expected. FAST™ DevMilestones is well suited for use as a kindergarten entry screener based on ratings of skill proficiency with anchors that reflect expected performance as a student progresses through kindergarten.

FAST™ DevMilestones supports data-based decisions by informing the selection of interventions matched to developmental needs and performance monitoring as interventions are applied. FAST™ DevMilestones also offers a unique approach to scoring that supports easy identification of developmental strengths that may be leveraged to enhance skill performance in areas of need. It is suggested that the FAST™ DevMilestones screening be conducted three times per year, including once in the early fall, winter, and spring. Administration at these time points will facilitate formative assessment of students' performance across increasingly sophisticated demonstrations of key developmental milestones. FAST™ DevMilestones is intended for use with all students enrolled in or immediately ready to be enrolled in kindergarten.

FAST™ DevMilestones Content Description

Language, Literacy, and Communications

Early language and literacy development are measured by six items within the language, literacy, and communications subscale of FAST™ DevMilestones. Skills within this domain reflect a student's ability to acquire and use language in support of academic tasks and social interactions. There has been growing consensus in the empirical literature that successful reading skills are facilitated through broad and deep understandings and uses of vocabulary that support verbal reasoning and verbal expression (Dikinson, McCabe, & Essex, 2006; Scarborough, 2001). Engaging in frequent and varied social conversations and reciprocal interactions support not only the expansion of vocabulary knowledge and refinement of categorically related concepts (Neuman & Roskos, 2005), but also enhance abstract reasoning (Snow, 1991) and abilities to discriminate units of language, such as words, segments, and phonemes (Goswami, 2001). In combination, these skills play essential roles in language and literacy development that are highly predictive of later reading achievement.

Cognitive Development

Acquisition and use of mathematical knowledge for problem solving, reasoning, and attitudes towards learning are measured by eight items in the cognitive development subscale of FAST™ DevMilestones. There are several specific skills have been identified as predictive of later proficiency in mathematics, scientific reasoning, and general knowledge. These skills include, understanding patterns and relations between objects and numbers based on numeracy skills (Clements, Sarama, & DiBiase, 2004), awareness of a wide array of attributes that may be used in evaluating relations between objects (Kilpatrick, Swafford, & Findell, 2001; NCTM 2000), as well as problem solving that reflects multiple ways of reasoning and multiple types of strategies to reach a solution (Miller, 2004). To complement empirical evidence for specific skills supporting cognitive development, there is research to suggest that when examining early development, skill acquisition is interdependent and best influenced by an emphasis on broad knowledge and conceptual understandings that are embedded in many different types of learning experiences (NAEYC/NCTM, 2002). Therefore, measurement in this domain emphasizes more global performance relative to students' demonstrations of reasoning about relations between objects and problem solving.

Social Emotional Development

Social emotional competencies are measured with eight items in the social emotional development subscale of FAST™ DevMilestones. A student's social and emotional competence is integrally linked to cognitive and academic competence as manifested by his/her ability to learn and be successful at school (Raver & Knitzer, 2002). Evaluating social emotional development includes examining awareness and regulation of one's own and others', cognitions, emotions, and behaviors (Raver & Knitzer, 2002). Emotion regulation skills have long been associated with students' academic outcomes based on proficiency in exhibiting behavioral control

(Neuenschwander et al., 2012) and managing attention to learning (Kuhl & Kraska, 1989). Skills associated with social emotional development also include a student's sense of self and ability to establish and maintain relationships with adults and peers. Relationships have been shown to be impacted through proficiency with emotional regulation through students' abilities to appropriately engage and disengage others in their environment (Porges, 2003), as well as students' abilities to develop good interpersonal skills that are absent of externalizing problem behaviors (Dunn & Brown, 1994; Rydell et al., 2003). In addition to the impact emotional regulation proficiency may have on interpersonal skills, interpersonal skills are also impacted through other skill deficits demonstrated by some students. Researchers have long explored the dynamic interaction between deficits in language, literacy, cognitive skills (Beitchman et al., 2001; Raver & Knitzer, 2002) and problem behaviors that disrupt formation of normative interpersonal skills. To effectively evaluate and understand performance in one domain, it is important to evaluate and understand performance in other domains.

Creativity and the Arts

A student's ability to appreciate and explore a variety of creative expressions is evaluated with three items in the creativity and the arts subscale of FAST™ DevMilestones. Researchers have suggested that children's experiences in their daily lives provide the basis for creativity and imagination through the processes of disassociation and association (Eckhoff & Urback, 2008). Experiences with different media create opportunities for students to explore/modify/change elements of their experience (disassociate) which may then be followed by creating new associations between elements in ways that were different from the student's original experience. These types of experiences have been demonstrated to improve students' ability to identify new, more sophisticated ways to problem solve (Eisner, 2002). Further, students who have learned to participate in a variety of creative expressions, such as art and music, have shown greater positive emotions such as interest and enthusiasm, as well as enhanced emotion regulation skills (Brown & Sax, 2012).

Approaches to Learning

Attitudes and behaviors that influence a student's ability to effectively engage and benefit from instruction are measured by five items in the approaches to learning subscale of FAST™ DevMilestones. How students engage in learning opportunities in classrooms has been linked to school achievement and adjustment (Fredricks, Blumenfeld, & Paris, 2004). Children who demonstrate greater persistence in their engagement and maintain engagement through transitions have been shown to make stronger gains in self-regulation skills and expressive language skills (WIlliford et al., 2013). Further, children who are highly motivated and self-confident tend to be more engaged in learning, thereby demonstrating improved outcomes, through their willingness to try new things and show curiosity in learning (Deci & Ryan, 1985; Dominguez et al., 2010; Eisenberg et al., 1998; Rydell et al., 2005).

Physical and Motor Development

Students' performance and understanding of physical health and well-being are evaluated with three items in the physical and motor development subscale of FAST™ DevMilestones. According to NEGP (1995), inclusion of this domain in recommendations for learning standards highlighted emerging research at the time that healthy children were more able to actively engage in learning opportunities than students who experienced some form of physical or emotional ailment. Research following the original recommendations of NEGP have demonstrated positive impacts from gross motor activities including physical well-being (Kirkcaldy, Shephard & Siefen, 2002), psychological well-being with improved self-esteem (Ekelund, Heian, & Hagen, 2004), and enhanced social skills and self-confidence (Gendron, Royer, Bertrand & Potrin, 2004). Improvements in these types of outcomes have been linked to greater resistance to addiction and generally improved well-being (Kirkcaldy et al., 2002). In addition to benefits observed relative to gross motor performance, many of which have direct relation to skills in other developmental domains, proficiency with fine motor skills also has important implications for school success. Students' abilities to demonstrate use of fine motor skills with precision and dexterity are related to how students get needs met and actively participate in learning activities.

FAST™ DevMilestones Content Development

The development of FAST™ DevMilestones followed an iterative process that began with identification of developmental domains to be included in FAST™ DevMilestones, moved to specification of domain-related skills, and concluded with evaluation of sequencing and alignment with learning standards. Guidance for selection of developmental domains was taken from the National Education Goals Panel (NEGP, 1995), which has since been used to guide the development of early learning standards by state agencies across the United States (Scott-Little, Kagen, & Frelow, 2006). NEGP recommendations provided the basis for domain selection given research supporting the need to consider the impact of development in domains outside of those traditionally examined in school settings. Currently, in response to federal regulations, most learning standards developed for use in K-12 school systems do not include standards addressing social emotional development and dispositions for learning (Logue, 2007), despite empirical support for the interaction between skills in those domains and children's learning (Fox, Dunlap, & Powell, 2002). Hence, the six developmental domains included in DevMilestones were selected to span recommendations for learning standards for children prior to Kindergarten while also providing appropriate emphasis on domains that require evaluation once children reach school age.

Item Specification, Construction, and Scaling of Ratings

Following identification of the developmental domains, a writing team was assembled to specify key skills within domains and construct items. Specification of skills within each domain was guided by empirical evidence supporting relations between skill proficiency and later school achievement. Table 66 provides a summary of key skills and a sample of the empirical evidence that guided item construction.

Table 66. Overview of Empirical Support Guiding Identification of FAST™ DevMilestones Key Skills

Developmental Domain	Key Skills	Empirical Support
Language, Literacy, and Communications	Motivation and enthusiasm for literacy activities. Engaging in social conversation and reciprocal interactions. Understanding and use of vocabulary and expressive communication.	Dikinson, McCabe, & Essex (2006); Goswami (2001); Neuman & Roskos (2005)
Cognitive Development	Understanding patterns and relations between objects. Using concrete and abstract strategies for mathematical problems. Problem solving.	Clements, Sarama, & DiBiase (2004); Kilpatrick, Swafford, & Findell (2001); Miller (2004); NCTM (2000)
Social Emotional Development	Understanding relationships. Recognizing and regulating emotions. Establishing and maintaining reciprocal relationships with others.	Bowman, Donovan, & Burns (2000); Neuenschwander, Rothlisberger, Cimeli, & Roebers (2012); Raver & Knitzer (2002); Skiba & Peterson (2000); Sugai, Horner, & Gresham (2002)
Creativity and the Arts	Showing interest and preferences for a variety of media and creative expressions. Using a variety media to participate in creative expressions.	Brown & Sax (2012); Eckhoff & Urback (2008); Eisner (2002)
Approaches to Learning	Showing curiosity and taking risks. Demonstrating imagination in ways of participating and problem solving. Showing persistence. Building on past experiences and learning.	Dominguez et al. (2010); Fredricks, Blumenfeld, & Paris, (2004); Williford, Maier, Downer, Pianta, & Howes (2013)
Physical and Motor Development	Engaging in a variety of gross and fine motor activities. Understanding physical health and well-being.	Ekelund, Heian, & Hagen, (2004); Gendron, Royer, Bertrand & Potrin, (2004); Kirkcaldy, Shephard & Siefen, (2002); Sugden (1986)

The writing team was comprised of experts in child development and measurement who combined information from the empirical literature with reviews of existing measurement tools and existing learning standards to create a pool of items representing each developmental domain. Item construction was based heavily on recommended language for use in existing sets of early learning standards for preschool-age children (i.e., Bowman et al., 2000; Logue, 2007; NEGP, 1995; Neuman & Roskos, 2005) and language used in describing Common Core State Standards for school-age children. Items were then refined to reflect only observable skills and rely on only minimal inferences within the rating process.

Refining items to reflect only observable skills was based on an item-scaling process that was guided by use of behaviorally anchored ratings as the basis for scoring items within FAST™ DevMilestones. Behaviorally anchored ratings were initially developed to improve the objectivity and reliability of rating scales by limiting inferences needed when assigning a performance rating to specific items (Christ & Boice, 2009). This approach to development of the scale for FAST™ DevMilestones lead the writing team to identify observable indicators for each item that aligned along a hierarchy of complexity and sophistication that were consistent with expectations for performance in a school environment. This process produced a complete matrix of indicators within categorical ratings that ranged from inquiring to mastering.

Alignment with Learning Standards

As part of the development process, item alignment with an existing set of learning standards was explored. An independent reviewer with expertise in child development was recruited to evaluate the alignment between DevMilestones and two sets of learning standards used in Minnesota: The Early Childhood Indicators of Progress (ECIPs) and the kindergarten standards that are aligned with the Common Core State Standards (CCSS). The process for exploring alignment involved three passes through linking of FAST™DevMilestones items with each set of learning standards. Pass one was based on a blind read of each item from the FAST™ DevMilestones tool and placing each with an item on the ECIPS or Kkindergarten CCSS Standards for which there was some theoretical linkage. Pass two involved sorting the items from FAST™ DevMilestones into the intended developmental domains to be represented by each item and then refining the item associations to reflect those FAST™ DevMilestones items that were intended to measure traits within the corresponding domain of the ECIPs or CCSS. The final pass involved refinement of item linkages and verification that performance indicators for each item accurately represented hierarchies of performance that were consistent with learning standards. This process revealed an 88% overlap in DevMilestones items and these two specific sets of learning standards.

FAST™ DevMilestones Administration

Completion times for FAST™ DevMilestones vary depending on the teacher's knowledge of individual students and the overall number of students receiving ratings from a teacher. It takes approximately 5 to 10 minutes to complete the FAST™ DevMilestones per student. When a teacher prepares to complete the FAST™ DevMilestones, preparation should be for the entire classroom of students. FAST™ DevMilestones is designed to be completed for all students at the same time. For each item, the teacher will be prompted to provide a rating for each student listed for the classroom before moving to the next item. Each item includes ratings that provide unique examples of observable performance indicators to consider when assigning a rating. Performance indicators are arranged to present a hierarchy of skill proficiency that may be used to support assignment of ratings across children and across time.

FAST™ DevMilestones Scores and Scoring

FAST™ DevMilestones is intended for use in evaluating student achievement of developmental milestones within and across multiple developmental domains. The 33 items comprising FAST™ DevMilestones correspond to six subscales, each of which represents a different developmental domain. These include Language, Literacy, and Communications (6 items), Cognitive Development (8 items), Social and Emotional Development (8 items), Creativity and the Arts (3 items), Approaches to Learning (5 items), and Physical and Motor Development (3 items).

Score types

Each item rating is converted to a numerical item score. Ratings of "unable to rate" do not produce an item score and therefore are also not counted toward any subscale scores. The remaining item ratings are scored in the following fashion:

Item Rating	Timing Expectations	Score
Not yet at first level	PreK	0
Inquiring	K-Entry	1
Emerging	Fall	2
Incorporating	Winter	3
Mastering	Spring	4

Subscale scores are calculated by summing item scores within each subscale and dividing the sum by the total number of items scored. This yields a mean subscale score. If any item is rated as, "unable to rate," that item is not counted within the total number of items available for a subscale. The total score is calculated by summing item scores across all subscales and dividing the sum by the total number of items that were scored. This yields an overall mean FAST™ DevMilestones score.

Score Interpretations

Higher FAST™ DevMilestones scores relative to expectations for each administration period suggest the student's performance is above expectations for that skill. Lower FAST™ DevMilestones scores indicate below expected performance of that skill. The finding of below expected performance would suggest the skill should potentially be targeted for intervention. The student support team, in collaboration with each student's classroom teacher, should determine the ultimate necessity of intervention.

FAST™ DevMilestones

Higher FAST™ DevMilestones item scores, mean FAST™ DevMilestones subscale scores, and mean FAST™ DevMilestones total scores correspond to greater proficiency and sophistication in demonstrating developmental milestones. To interpret each individual student's scores, it is recommended that his or her scores be compared to the original FAST™ DevMilestones ratings with attention to the period in which the ratings were provided. This approach allows educators to determine how the student is developing relative to standards for each individual skill, domain of skills, and overall functioning.

FAST™ DevMilestones Construct Validity

Content-Related Validity Evidence

As part of the development process, item alignment with an existing set of learning standards was explored. An independent reviewer with expertise in child development was recruited to evaluate the alignment between FAST™ DevMilestones and two sets of learning standards used in Minnesota, the Early Childhood Indicators of Progress (ECIPs) and the kindergarten standards that are aligned with the Common Core State Standards (CCSS). The process for exploring alignment involved three passes through linking of DevMilestones items with each set of learning standards. Pass one was based on a blind read of each item from the FAST™ DevMilestones tool and placing each with an item on the ECIPs or kindergarten CCSS for which there was some theoretical linkage. Pass two involved sorting the items from FAST™ DevMilestones into the intended developmental domains to be represented by each item and then refining the item associations to reflect those FAST™ DevMilestones items that were intended to measure traits within the corresponding domain of the ECIPs or CCSS. The final pass involved refinement of item linkages and verification that performance indicators for each item accurately represented hierarchies of performance that were consistent with learning standards. This process revealed an 88% overlap in FAST™ DevMilestones items and these two specific sets of learning standards (see Table 67 and Table 68 below).

Table 67 FAST™ DevMilestones Alignment Study: Standards 1

Domains	GOLD	DRDP	FAST™	WSS-MN
Social Emotional	19	19	19	7
Approach to Learning	12	9	11	4
Language/ Literacy	21	21	16	13
Creativity/ Arts	6	1	7	7
Cognitive	27	23	23	4
Physical	11	0	10	5
TOTAL Indicators	96	73	86	40
Percent Coverage	98%	74%	88%	41%

Note. This information is from an independent study by the Minnesota Department of Education, which was presented at the Minnesota Assessment Conference (M. Cox, July 2015)

Table 68 FAST™ DevMilestones Alignment Study: Standards 2

Claims & Proposed Uses	GOLD	DRDP	FAST™	WSS-MN
Domains represent basic components of early development in assessment domains.	Strong	Strong	Strong	Strong
Provides information of school readiness or future school performance.	Moderate	None	Moderate	N/A
Domains and items are research based.	Strong	N/A	Strong	Strong
Alignment Studies re: CCSS	Yes	Yes	Yes	Yes
Alignment Studies re: MN State Standards	Yes	N/A	Yes	N/A

Note. This information is from an independent study by the Minnesota Department of Education, which was presented at the Minnesota Assessment Conference (M. Cox, July 2015)

Construct-Related Validity Evidence

Confirmatory factor analysis via Mplus 7.3 was employed to examine the fit of items to the associated constructs. Weighted least squares means- and variance- adjusted (WLSMV) was used to estimate parameters. The CFA for the 8-factor model yielded $\chi^2(751)$=2283.394*, with p=.000. It should be noted that the chi-square value is not used for the difference testing when using MLSWV. For the goodness-of-fit, the following cutoff values for good fit were suggested by Hu and Bentler (1999): CFI > .95, RMSEA < .06, and WRMR < .90. Root Mean Square Error of Approximation (RMSEA) was .09, which was higher than the recommended cutoff; CFI was .92, which was lower than the recommended cutoff; Weighted Root Mean Square Residual (WRMR) was 1.55, which was higher than the recommended cutoff. Overall, the model is mediocre fit. All the factor loadings on each latent variable are significant with p values smaller than .001. Results are shown on the following page.

Reliability-Related Evidence

Evidence of FAST™ DevMilestones internal consistency across items and developmental domains is provided in Table 69 based on existing data from multiple states in the Midwest, Northeast, and Southeast.

Table 69. Internal Consistency for DevMilestones Scales

Constructs	Fall		Winter		Spring	
	N	α	N	α	N	α
Language, Literacy, and Communications (6 items)	87	.904	74	.941	14	.758
Cognitive Development (8 items)	90	.884	74	.951	-	-
Social and Emotional Development (8 items)	92	.915	73	.941	-	-
Creativity and the Arts (3 items)	92	.817	73	.827	-	-
Approaches to Learning (5 items)	92	.834	73	.883	-	-
Physical & Motor Development (3 items)	93	.727	73	.698	-	-
Full Scale	83	.974	73	.984	-	-

Evidence of the reliability of the full scale and subscale scores is presented in Table 70. Data were gathered from multiple states in the Midwest, Northeast, and Southeast. The reliability range coefficients have been computed with a 95% confidence interval. In addition, the time lag in mean number of weeks between data collections is reported.

Table 70. Test-Retest Reliability for DevMilestones Scales

Constructs	N	Coefficients
Language, Literacy, and Communications	72	.786
Cognitive Development	72	.766
Social and Emotional Development	71	.677
Creativity and the Arts	72	.471
Approaches to Learning	71	.714
Physical & Motor Development	72	.584
Full Scale	72	.773

Note. Used fall-to-winter data, with students tested in both periods. N = sample size.

References

Adams, M. J. (1990). *Beginning to read.* Cambridge, MA: MIT Press.

AERA, APA, NCME. (1999). *Standards for educational and psychological testing.* Washington DC: American Educational Research Association.

AIMSWeb Benchmark and Progress Monitoring System for Grades K-8. Pearson Education Inc.

Ardoin, S. P., & Christ, T. J. (2009). Curriculum based measurement of oral reading: Estimates of standard error when monitoring progress using alternate passage sets, *School Psychology Review*, 38, 266-283.

Ardoin, S. P., Carfolite, J., Christ, T. J., Roof, C. M., & Klubnick, C. (2010). Examining readability estimates' predictions of students' oral reading rate: Spache, Lexile, and Forcast. *School Psychology Review, 39*(2), 277-285.

Ardoin, S. P., Suldo, S. M., Witt, J. C., Aldrich, S., & McDonald, E. (2005). Accuracy of readability estimates' predictions of CBM performance. *School Psychology Quarterly, 20*(1), 1-22.

Armbruster, B. (2002). *Put reading first* DIANE Publishing. Aud, S., Wilkinson-Flicker, S., Kristapovich, P., Rathbun, A., Wang, X., & Zhang, J. (2013). *The Condition of Education 2013* (NCES 2013-037). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved [November 21st, 2013] from http://nces.ed.gov/pubsearch.

Atkin, J. M., Black, P., & Coffey, J. (2001). Classroom Assessment and the National Science Education Standards. Washington, DC: National Academy Press.

Baroody, A. J. (2006). Why children have difficulties mastering the basic number combinations and how to help them. *Teaching Children Mathematics, 13*, 22 - 31.

Bear, D.R., Invernizzi, M.A., Templeton S.R., & Johnston, F.R. (2012). *Words Their Way: Word Study for Phonics, Vocabulary, and Spelling Instruction, 5th ed.* Upper Saddle, NJ: Pearson Education Inc.

Beitchman J, Wilson B, Johnson CJ, Atkinson L, Young A, Adlaf E, et al. (2001). Fourteen-year follow-up of speech/language-impaired and control children: psychiatric outcome. Journal of the American Academy of Child and Adolescent Psychiatry, 40, 75-82.

Berch, D. B., & Mazzocco, M. M. (Eds) (2008). Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities. Baltimore, MD: Paul H. Brookes.

Betts, J., Pickart, M., & Heistad, D. (2009). An investigation of the psychometric evidence of CBM-R passage equivalence: Utility of readability statistics and equating for alternate forms *Journal of School Psychology*, *47*, pp. 1-17.

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

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The Effects of Morphological Instruction on Literacy Skills A Systematic Review of the Literature. *Review of Educational Research*, *80*(2), 144-179.

Bowman, B., Donovan, S., & Burns, M.S. (Eds.). (2000). Eager to learn: Educating our preschoolers. National Research Council Commission on Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.

Brennan, R. L. (Ed.). (2006). *Educational Measurement* (4th ed.). Westport, CT: American Council on Education and Praeger Publishers.

Brown, E.D., Sax, K.L. (2013). Arts enrichment and preschool emotions for low-income children at risk. Early Childhood Research Quarterly, 28, 337-346.

Burt, J. S. (2006). What is orthographic processing skill and how does it relate to word identification in reading? *Journal of Research in Reading*, *29*(4), 400-417.

Cain, K., Oakhill, J., & Bryant, P. E. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, 96, 31-42.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development*, *65*, 684-698.

Carnine, D.W., Silbert, J., Kem'ennui, E.J., & Tarver, S.G. (2009). *Direct instruction reading* (5th ed.). , OR: Pearson Education Ltd.

Carroll, J. B. (1993). Human cognitive abilities: A survey of factor-analytic studies. New York:
Cambridge University Press.

Carlisle, J.F. (1995). Morphological awareness and early reading achievement, L.B. Feldment, (ed.), *Morphological aspects of language processing,* (pp. 189-209). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Carlisle, J. F. (2011). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, 45 (4), 464-487.

Case, S.M., & Swanson, D.B. (2002). *Constructing written test questions for the basic and clinical sciences, 3rd edition.* Philadelphia, PA: National Board of Medical Examiners.

Cawley, J. F., Parmar, R. S., Lucas-Fusco, L. M., Kilian, J. D., & Foley, T. E. (2007). Place value and mathematics for students with mild disabilities: Data and suggested practices. *Learning Disabilities: A Contemporary Journal, 5,* 21-39.

Chall, J. (1987). Two vocabularies for reading: Recognition and meaning. In M. McKeown & M. Curtis (Eds.). The nature of vocabulary acquisition (pp. 7-17). Hillsdale, NJ: Lawrence Erlbaum.

Christ, T. J. (2006). Short term estimates of growth using curriculum-based measurement of oral reading fluency: Estimates of standard error of the slope to construct confidence intervals. *School Psychology Review*, 35, 128-133.

Christ, T. J., & Ardoin, S. P. (2009). Curriculum-based measurement of oral reading: Passage equivalence and probe-set development. *Journal of School Psychology*, 47, 55-75.

Christ, T. J., & Boice, C. H. (2009). Rating scale items: A brief review of nomenclature, components and formatting. *Assessment for Effective Intervention, 34*(4), 242-250.

Christensen, C. A., & Bowey, J. A. (2005). The efficacy of orthographic rime, grapheme-phoneme correspondence, and implicit phonics approaches to teaching decoding skills. *Scientific Studies of Reading*, *9*(4), 327-349.

Clements, D., Sarama, J., & DiBiase, A. M. (Eds.). (2004). Engaging young children in mathematics: Findings of the 2000 national conference on standards for preschool and Kindergarten mathematics education. Mahwah, NJ: Erlbaum.

References

Compton, D. L., Appleton, A. C., & Hosp, M. K. (2004). Exploring the relationship between text-leveling systems and reading accuracy and fluency in second grade students who are average and poor decoders. *Learning Disabilities Research & Practice*, 19(3), 176-184.

Clay, M. (1972). Reading, the Patterning of Complex Behavior. Auckland, New Zealand: Heinemann.

Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory* (Vol. 6277). New York: Holt, Rinehart and Winston.

Deacon, S. H., Parrila, R., & Kirby, J. R. (2008). A review of the evidence on morphological processing in dyslexics and poor readers: A strength or weakness. *The Sage handbook of dyslexia*, 212-237.

Deno, S. L. (1985). Curriculum-Based measurement: The emerging alternative. *Exceptional Children*, 52, 219-232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review*, 1(3), 358-374.

Deno, S. L. (2002). Problem solving as best practices. In A. Thomas & J. Grimes (Eds.), Best practices in school psychology IV. (pp. 37-56). Bethesda, MD: National Association of School Psychologists.

Deno, S. L. (2003). Developments in Curriculum-Based measurement. *The Journal of Special Education*, 37(3), 184-192.

Deno, S. L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), Assessment for intervention: A problem-solving approach (pp. 10-42). New York: Guilford Press.

Deno, S. L. & Mirkin, P. K. (1977) Data-based Program Modification: A Manual.

Reston VA: Council for Exceptional Children

De Ayala, R. J. (2009). The theory and practice of item response theory. Guilford Press.

Dickinson, D., McCabe, A., & Essex, M. (2006). A window of opportunity we must open to all: The case for high-quality support for language and literacy. In D. K. Dickinson & S. B. Neuman (Eds.), Handbook of early literacy research (pp. 11-28). New York: Guilford Press.

Dominguez, X., Vitiello, V.E., Maier, M.F., & Greefield, D.B. (2010). Longitudinal examination of young children's behavior: Child-level and classroom-level predictors of change throughout the preschool year. School Psychology Review, 39, 29-47.

Downing, J., Ollila, L., & Oliver, P. (1975). Cultural differences in children's concepts of reading and writing. *British Journal of Educational Psychology*, *45*, 312-316.

Dunn, J. & Brown, J. (1994). Affect expression in the family, children's understanding of emotions, and their interactions with others. Merrill-Palmer Quarterly, 40, pp. 120-137

Durkin, D. (1993). *Teaching them to read* (6th ed.). Boston, MA: Allyn and Bacon.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS Next). Eugene, OR: Institute for the Development of Educational Achievement. Available: https://dibels.org/next/index.php.

Eckoff, A. & Urback, J. (2008). Understanding imaginative thinking during childhood: Sociocultural conceptions of creativity and imaginative thought. Early Childhood Education Journal, 36, 179-185.

Eisenberg, N., Shepard, S. A., Fabes, R. A., Murphy, B. C, & Guthrie, I. K. (1998). Shyness and children's emotionality, regulation, and coping: Contemporaneous, longitudinal, and across-context relations. *Child Development, 69,* 767-790.

Eisner, E. W. (2002). The arts and the creation of mind. New Haven, CT: Yale University Press.

Ekeland, E., Heian, F. & Hagen, K.B. (2004). Can exercise improve self-esteem in children and young people? A systematic review of randomized control trials. *British Journal of Sports Medicine, 39,* 792-798.

Flesch, R. F. (1949). A new readability yardstick. Journal of Applied Psychology, 32(3), 221-233.

Francis, D. J., Santi, K. L., Barr, C., Fletcher, J. M., Varisco, A., & Foorman, B. R. (2008). Form effects on the estimation of students' oral reading fluency using DIBELS. *Journal of School Psychology*, 46, 315-342.

Fry, E. B., & Kress, J. E. (2006). *The reading teacher's book of lists*. San Francisco: Jossey-Bass.

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

Fuchs, L. S., & Fuchs, D. (2002) Curriculum-Based Measurement: Describing competence, enhancing outcomes, evaluating treatment effects, and identifying treatment non-responders. *Peabody Journal of Education*, 77:2, 64-84, doi: 10.1207/S15327930PJE7702_6

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly*, *27*(4), 216-227.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, *9*(2), 20-28.

Gajdamaschko, N. (2005). Vygotsky on imagination: Why an understanding of the imagination is an important issue for schoolteachers. Teaching Education, 16, 13-22

Geary, D.C. (1999). Mathematical Disabilities: What we know and don't know. Retrieved

August 23, 2011 from the LD Online Web site: http://www.ldonline.org/article/5881.

Geary, D.C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37,*

(1), 4-15.

Gendron, M., Royer, E., Bertrand, R. & Potrin, P. (2004). Behavior disorders, social competence, and the practice of physical activities among adolescents. *Emotional and Behavioral Difficulties, 9,* 249-259.

Gersten, R., Clarke, B., Jordan, N. C., Newman-Gonchar, R., Haymond, K., & Wilkins, C. (2012). Universal screening in mathematics for the primary grades: Beginnings of a research base. *Exceptional Children, 78*, 423-445.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of learning disabilities*, *38*(4), 293-304.

Goswami, U. (2000). Phonological and lexical processes. In M. J. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.). *Handbook of reading research, Volume III* (pp. 251-267) Mahwah, NJ: Lawrence Earlbaum.

Goswami, U. (2001). Early phonological development and the acquisition of literacy. In S. B. Neuman & D. Dickinson (Eds.), *Handbook of Early Literacy Research* (pp. 111-125). New York: Guilford Press.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and special education*, *7*(1), 6-10.

Goldman, S. R., & Varnhagen, C. K. (1986). Memory for embedded and sequential episodes in stories. *Journal of Memory and Language*, 25, 401-418.

Group Reading Assessment and Classification Evaluation (GRADE). Pearson Education Inc.

References

Graesser, A. C., Leon, J. A., & Otero, J. (2002). Introduction to the psychology of science text comprehension. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), The Psychology of Science Text Comprehension (pp. 1-15). Mahwah, NJ: Erlbaum.

Graesser, A. C., McNamara, D. S., & Louwerse, M. M (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A.P. Sweet and C.E. Snow (Eds.), Rethinking reading comprehension (pp. 82-98). New York: Guilford Publications.

Graesser, A. C., McNamara, D. S., Lowerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202

Graesser, A.C., Olde, B. A., & Klettke, B. (2002). How does the mind construct and represent stories? In M. Green, J. Strange, and T. Brock (Eds.), Narrative Impact: Social and Cognitive Foundations. Mahwah: NJ: Erlbaum.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395.

Grzybowski, M., & Younger, J. G. (1997). Statistical methodology: III. Receiver operating characteristic (ROC) curves. *Academic Emergency Medicine*, *4*, 818-826.

Griffin, S. (2008). Early intervention for children at risk of developing mathematical learning difficulties. In D. Berch & M. Mazzocco (Eds.) *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities.* (pp. 373 - 395) Baltimore, MD: Paul H. Brookes.

Guyer, R., & Thompson, N.A. (2012). *User's Manual for Xcalibre item response theory calibration software, version 4.1.6*. St. Paul MN: Assessment Systems Corporation. Available from http://www.assess.com/

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied measurement in education*, *15*, 309-333.

Haladyna, T.M. (2007). Roles and importance of validity studies in test development (pp. 739-760). In S.M. Downing and T.M. Haladyna (Eds.) *Handbook of test development.* Mahwah, NJ: Lawrence Erlbaum Associates.

Hardy, M., Stennett, R., & Smythe, P. (1974). Development of auditory and visual language concepts and relationship to instructional strategies in Kindergarten. *Elementary English Journal*, 51, 525-532.

Harlin, R., & Lipa, S. (1990). Emergent literacy: A comparison of formal and informal assessment methods. *Reading Horizons,* 20, 209-223.

Henard, D.H. (2000). Item response theory. In L. Grimm & P. Yarnold (Eds.), Reading and understanding more multivariate statistics (pp. 67-97). Washington, DC: American Psychological Association.

Hiebert, E. H., & Taylor, B.M. (2000). Beginning reading instruction: Research on early interventions. In M. J. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.). *Handbook of* reading *research, Volume III* (pp. 455-482) Mahwah, NJ: Lawrence Earlbaum.

Hiebert, E. H., & Fisher, C.W. (2007). The critical word factor in texts for beginning readers. Journal of *Educational Research*. 101(1), 3-11.

Hintze, J. M., & Christ, T. J. (2004). An examination of variability as a function of passage variance in CBM progress monitoring. *School Psychology Review*, 33, 204-217.

Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the classification accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review*, *34*, 372-386.

Hoffman, K. I. (1993). The USMLE, the NBME subject examinations, and assessment of individual academic achievement. *Academic Medicine*, *68*(10), 740-7.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). The ABC"s of CBM: A practical guide to Curriculum-Based measurement. New York, NY: The Guilford Press.

Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. STRUCTURAL EQUATION MODELING. 6, 1-55.

Jenkins, D. (2001). Impact of the implementation of the teaching/learning cycle on teacher decision-making and emergent readers. *Reading Psychology*, 22(4), 267-288.

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a Response-to-Intervention (RTI) framework. *School Psychology Review*. 36(4).

Jenkins, J. R., Zumeta, R., Dupree, O., & Johnson, K. (2005). Measuring gains in reading ability with passage reading fluency. *Learning Disabilities Research & Practice*, 20(4), 245-253.

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review*, *36*, 582-600.

Jenkins, J. R., & Jewell, M. (1992). An examination of the concurrent validity of the Basic Academic Skills Samples (BASS). *Assessment for Effective Intervention*, *17*(4), 273-288.

Johns, J. (1972). Children's concepts of reading and their reading achievement. *Journal of Reading Behavior*, 4, 56-57.

Johns, J. (1980). First Graders' concepts about print. *Reading Research Quarterly*, 15, 529- 549.

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*, 850-867.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of educational Psychology*, *80*, 437-447.

Juel, C. (2006). The impact of early school experience on initial reading. In D. K. Dickinson, & S. B. Neuman (Eds.). *The handbook of early literacy research* (v.2), (pp. 410-426). NY: Guilford Press.

Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly*, *35*(4), 458-492.

Kane, M. T. (2013). Validating the interpretation and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1-73.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology-Amsterdam*, 94, 67.

Kendeou, P., van den Broek, P., White, M., & Lynch, J. (2007). Preschool and early elementary comprehension: Skill development and strategy interventions. In D. S. McNamara (Ed.) *Reading comprehension strategies: Theories, interventions, and technologies,* (pp. 27-45). Mahwah, NJ: Erlbaum.

Kilgus, S.P., Chafouleeas, S.M., & Riley-Tillman, T.C. (2013). Development and initial validation of the social and academic behavior risk screener for elementary grades. *School Psychology Quarterly, 28,* 210-226.

References

Kilgus, S. P., Sims, W., von der Embse, N. P., & Riley-Tillman, T. C. (under review). Confirmation of models for interpretation and use of the Social and Academic Behavior Risk Screener (SABRS). *School Psychology Quarterly*.

Kilgus, S. P., Eklund, K., von der Embse, N. P., & Taylor, C. (2014). Classification accuracy of the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS) in elementary and middle grades. Manuscript in preparation.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up*. Washington, DC: National Academy Press.

Kim-Kang, G., & Weiss, D. J. (2007). Comparison of computerized adaptive testing and classical methods for measuring individual change. In *Proceedings of the 2007 GMAC Conference on Computerized Adaptive Testing.* Available from www. psych. umn.edu/psylabs/CATCentral.

Kim-Kang, G., & Weiss, D. J. (2008). Adaptive measurement of individual change. *Zeitschrift für Psychologie/Journal of Psychology*, *216*, 49-58.

Kingsbury, G. G., & Houser, R. L. (1999). Developing computerized adaptive tests for school children. In F. Drasgow & J. B. Olson Buchanan (Eds.), *Innovations in computerized assessment* (pp. 93-115). Mahwah, NJ: Erlbaum.

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, *30*(4), 28-37.

Kintsch, W. (1998). Comprehension: A Paradigm for Cognition. Boulder, CO: Cambridge University Press.

Kintsch, W. & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 35(5), 363-394.

Kirkcaldy, B.D., Shephard, R.J. & Siefen, R.G. (2002). The relationship between physical activity and self-image and problem behavior among adolescents. *Social Psychiatry and Psychiatric Epidemiology, 37,* 544-550.

Klare, G. R. (1974-1975). Assessing readability*. Reading Research Quarterly*, 10(1), 61-102.

Kolen, M. J., & Brennan, R. L. (2004). Test equating, scaling, and linking: Methods and practices (2nd ed.). New York: Springer-Verlag.

Kranzler, J.H., Brownell, M.T., & Miller, M.D. (1998). The construct validity of curriculum-based measurement of reading: An empirical test of a plausible rival hypothesis. *Journal of School Psychology, 36* (4), 399-415.

Kuhl, J. & Kraska, K. (1989). Self-regulation and metamotivation: Computational mechanisms, development, and assessment. In R. Kanfer, P. Ackerman, R. Cudeck (Eds.), Abilities, motivation, and methodology: The Minnesota Symposium on learning and individual differences, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 373-374.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*, 293-323.

Lee, J., Grigg, W., 8c Donahue, P. (2007). The nation's report card: Reading 2007 (NCES 2007-496). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Levy, B. A., Gong, Z., Hessels, S., Evans, M. A., & Jared, D. (2006). Understanding print: Early reading development and the contributions of home literacy experiences. *Journal of Experimental Child Psychology*, *93*(1), 63-93.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, *74*(6), 431.

Lipson, M.L., Mosenthal, J. H., Mekkelsen, J., & Russ, B. (2004). Building knowledge and fashioning success one school at a time. *The Reading Teacher, 57* (6) 534-542.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly*, 13, 123-146.

Lomax, R. G., & McGee, L.M. (1987). Young children's concepts about print and reading: Toward a model of word reading acquisition. *Reading Research Quarterly,* 22, 237-256.

MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L.G. (2000). *Gates-MacGinitie Reading Tests.* 4th ed. Itasca, IL: Riverside Publishing Company.

Mandler, J. M. & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9(1), 111-151.

Markell, M. A., & Deno, S. L. (1997). Effects of increasing oral reading generalization across reading tasks. *The Journal of Special Education*, *31*(2), 233-250.

Martone, A., & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment and instruction. *Review of Educational Research, 79*(3), 1-76.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The Effects of Theoretically Different Instruction and Student Characteristics on the Skills of Struggling Readers. *Reading Research Quarterly*, 148-182.

Mazzeo D, Arens S, Germeroth C, Hein H. (2012). Stopping childhood obesity before it begins. *Phi Delta Kappan*, 93(7):10-15.

Mazzocco, M. M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability*. Learning Disabilities Research and Practice, 20,* 142 - 155.

McGregor, K. K. (2004). Developmental dependencies between lexical semantics and reading. *Handbook of language and literacy*, 302-317.

Measures of Academic Progress (MAP). Northwest Evaluation Association.

Messick, S. (1993). Validity. (In R. L. Linn (Ed), Educational measurement (2nd ed. pp. 13—104). Phoenix: American Council on Education and Oryx Press.)

Methe, S. A., Hojnoski, R., Clarke, B., Owens, B. B., Lilley, P. K., Politylo, B. C., White, K. M., & Marcotte, A. M. (2011). Innovations and future directions for early numeracy curriculum-based measurement: Commentary on the special series. *Assessment for Effective Intervention, 36,* 367 - 385.

Miller, K. (2004, October). *Developing number names: A cross cultural analysis*. Presentation at the Early Childhood Academy, University of Michigan, Ann Arbor.

Morsy, L., Kieffer, M., & Snow, C. (2010). Measure for Measure: A Critical Consumers' Guide to Reading Comprehension Assessments for Adolescents. Final Report from Carnegie Corporation of New York's Council on Advancing Adolescent Literacy. *Carnegie Corporation of New York*.

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of Kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities, 42*, 306 321.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental psychology*, *40*(5), 665.

References

Nagy, W. E. (1988). *Teaching Vocabulary to improve reading comprehension*. National Council of Teachers of English, 1111 Kenyon Rd., Urbana, IL 61801 (Stock No. 52384-015, $4.95 member, $7.50 nonmember--ISBN-0-8141-5238-4); International Reading Association, PO Box 8139, 800 Barksdale Rd., Newark, DE 19714-8139 (No. 151, $4.95 member, $7.50 nonmember--ISBN-0-87207-151-0).

National Center for Education Statistics (2013). *The Nation's Report Card: A First Look: 2013 Mathematics and Reading* (NCES 2014-451).Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Center for Education Statistics (2013). The Nation's Report Card: Trends in Academic Progress 2012 (NCES 2013 456). Institute of Education Sciences, U.S. Department of Education, Washington, D.C. National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, & Technical Subjects*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State Officers.

National Center for Response to Intervention (2010). Screening Tools Chart. U.S. Office of Special Education Programs. Retrieved from http://www.rti4success.org/sites/default/files/Screening%20Tools%20Chart.pdf

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics.* Washington, DC: Authors.

National Research Council. (2009). Mathematics learning in early childhood: Paths toward excellence and equity. Washington, DC: Author.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language & Hearing Research*, *47*(1).

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of research in reading*, *27*(4), 342-356.

National Mathematics Advisory Panel [NMAP], (2008) Foundations for Success: The Final Report of the National Mathematics Advisory Panel (No. ED04CO0015/0006): U.S. Department of Education: Washington, D.C.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000a). Report of the National Reading Panel: Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: reports of the subgroups. National Institute of Child Health and Human Development, National Institutes of Health.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000b). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. National Institute of Child Health and Human Development, National Institutes of Health.

NCTM. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council for Teachers of Mathematics.

Neuenschwander, R., Rothlisberger, M, Cimeli, P., Roebers, C.M. (2012). How do different aspects of self-regulation predict successful adaptation to school? Journal of Experimental Child Psychology, 113, 353-371.

Neuman, S.B. & Roskos, K. (2005). The state of pre-Kindergarten standards. *Early Childhood Research Quarterly, 20,* 125-145.

Nichols, W. D., Rickelman, R.J., & Rupley, W. H. (2004). Examining phonemic awareness and concepts of print patterns of Kindergarten students. *Reading Research and Instruction*. 43, 56-82.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse processes*, *22*(3), 247-288.

Northwest Evaluation Association (NWEA) (2011). *Measures of Academic Progress (MAP).* Portland, OR.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York: McGraw-Hill.

Nydick, S. W., & Weiss, D. J. (2009). A hybrid simulation procedure for the development of CATs. In Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing. Retrieved from www. psych.umn.edu/psylabs/CATCentral.

Oakhill, J., & Cain, K. (2007). Introduction to comprehension development. In K. Cain & J. Oakhill (Eds.), *Children's comprehension problems in oral and written language: A cognitive perspective* (pp. 3-40). New York: Guilford Press.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and cognitive processes*, *18*, 443-468.

Okamoto, T. (1996, July). On relationships between statistical zero-knowledge proofs. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (pp. 649-658). ACM.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*(2), 184-202.

Perfetti, C. A. (1994). Psycholinguistics and reading ability. In M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 849-894). San Diego, CA: Academic Press.

Perfetti, C. A. (1992). The representation problem in reading acquisition.

Porges, S. (2003). The polyvagal theory: Phylogenetic contributions to social behavior. Physiology and Behavior, 79, 503-513.

Phillips, B. M., & Torgesen, J. K. (2006). Phonemic awareness and reading: Beyond the growth of initial reading accuracy. *Handbook of early literacy research*, *2*, 101-112.

Pratt, K., Martin, M., White, M.J., & Christ, T.J. (2010). Development of a FAIP-R First Grade Probe Set (Technical Report No. 3). Minneapolis, MN: University of Minnesota, Department of Educational Psychology.

Purpura, D. J., & Lonigan, C. J. (2013). Informal numeracy skills: The structure and relations among numbering, relations, and arithmetic operations in preschool. *American Educational Research Journal, 50*, 178-209.

RAND Reading Study Group (2002). Reading for understanding: Toward an R&D program in reading comprehension. Santa Monica, CA: RAND.

Raver, C. C. & Knitzer, J. (2002). Ready to enter: What research tells policymakers about strategies to promote social and emotional school readiness among three- and four-year-old children. Promoting the Emotional Well-being of Children and Families. Policy Paper No. 3. Columbia, New York: National Center for Children in Poverty.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. Jr. (2012). Psychology of reading (2nd ed.). New York: Psychology Press.

Raudenbush, S. W. and Anthony S. Bryk (2002). Hierarchical linear models: Applications and data analysis methods, Second Edition. Newbury Park, CA: Sage

Reschly, A. L., Busch, T.W., Betts, J., Deno, S. L., and Long, J.D. (2009). Curriculum-Based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, 47, 427-469.

References

Ricketts, J., Nation, K., & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, *11*(3), 235-257.

Riddle Buly, M., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading tests. *Educational Evaluation and Policy Analysis, 24*, 219-239.

Renaissance Learning. (1998b). *STAR Math.* Wisconsin Rapids, WI: Renaissance Learning.

Rubin, D. C. (1995). Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out-rhymes. New York: Oxford University Press.

Rydell, A., Berlin, L., & Bohlin, G. (2003). Emotionality, emotion regulation, and adaptation among 5- to 8-year-old children. Emotion, 3, 30-47.

Samejima, F. (1994). Some critical observations of the test information function as a measure of local accuracy in ability estimation. *Psychometrika*, *59*(3), 307-329.

Samuels, S. J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly, 42*, 563-566.

Sadler, D. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119-144.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97-110). New York: Guilford Press.

Schank, Roger C., and Robert P. Abelson. 1977. Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. Hillsdale, NJ: Lawrence Erlbaum.

Schmeiser, C. B., & Welch, C. J. (2006). Test development. In R. L. Brennan (Ed.), Educational Measurement (4th ed., pp. 623-646). Westport, CT: American Council on Education and Praeger Publishers.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology*, *45*, 193-223.

Shinn, M. R. (Ed.). (1989). Curriculum-based measurement: Assessing special children. New York: Guilford Press.

Skiba, R. & Peterson, R. (2000). School discipline at a crossroads: From zero tolerance to early response. *Exceptional Children, 32*, 200-216.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. National Academies Press.

Stahl, S. A. (2001). Teaching phonics and phonological awareness. *Handbook of early literacy research*, *1*, 333-347.

Snow, C. (1991). The theoretical basis for relationships between language and literacy in development. *Journal of Research in Childhood Education*, *6*, 5-10.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. National Academies Press.

Stanovich, K. E. (1981). Attentional and automatic context effects in reading. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, N.J.: Erlbaum.

Stanovich, K.E. (1984). The interactive-compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *Remedial and Special Education*, 5, 11-19.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 360-407.

Stanovich, K. E. (1990). A call for an end to the paradigm wars in reading research. *Journal of Reading Behavior*, 22, 221-231.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*(4), 402-433.

Stein, N. L. & Glenn, C. G. (1975). An Analysis of Story Comprehension in Elementary School Children: A Test of a Schema (Report No. PS 008 544). Washington University. (ERIC Document Reproduction Service No. ED121474).

Stein, N. L. & Trabasso, T. (1982). What's in a story: An approach to comprehension and instruction. In R. Glaser (Ed.), Advances in the psychology of instruction (Vol. 2). Hillsdale, N.J.: Erlbaum.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental psychology*, *38*, 934-947.

Sugai, G., Horner, R., & Gresham, F. (2002). Behaviorally effective school environments. In Shinn, M.R., Walker, H.M.,& Stoner, G. (Eds.), Interventions for academic and behavior problems II: Preventive and remedial approaches (pp. 315-350). Bethesda, MD: National Association of School Psychologists.

Sugden, D.A. (1986). The development of proprioceptive control. Themes in Motor Development, 35, 21-39.

Swets, J. A., R. M. Dawes, & J. Monahan (2000), "Psychological Science Can Improve Classification Decisions", Psychological Science in the Public Interest 1: 1-26. Test of Silent Reading Efficiency and Comprehension (TOSREC).

Taylor, B. M., Pearson, P., Clark, K., & Walpole, S. (2000). Effective schools and accomplished teachers: Lessons about primary-grade reading instruction in low-income schools. *The Elementary School Journal*, 101(2), 121-165.

Taylor, B. M., Pearson, P., Peterson, D. S., & Rodriguez, M. C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *The Elementary School Journal*, 104(1), 3-28.

Taylor, B. M., Pearson, P., Peterson, D. S., & Rodriguez, M. C. (2005). The CIERA School Change Framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly*, 40(1), 40-69.

Thompson, S. J., Johnstone, C. J., Thurlow, M. L., & Clapper, A. T. (2004). *State literacy standards, practices, and testing: Exploring accessibility* (Technical Report 38). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Trabasso, T. & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes*, 15, 249-275.

Trabasso, T. & Stein, N. L. (1997). Narrating, representing, and remembering event sequences. In P. W. van den Broek, P. J. Bauer, & T. Bourg (Eds.), Developmental spans in event comprehension and representation: Bridging fictional and actual events (pp. 237-270). Mahwah, NJ: Erlbaum.

Tumner, W. E., Herriman, M. L., & Nesdale, A. R. (1988). Metalinguistic abilities and beginning Reading. *Reading Research Quarterly*, 23, 134-158.

Van de Walle, J. A., Karp, K. S. & Bay-Williams, J. M. (2013). Elementary and middle school mathematics: Teaching developmentally (8th ed.). New York, NY: Pearson.

van den Broek, P. & Trabasso, T. (1986). Causal networks versus goal hierarchies in summarizing text. *Discourse Processes*, 9, 1-15.

References

van den Broek, P. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.) Handbook of Psycholinguistics (pp. 539-588). New York: Academic Press.

VanDerHeyden, A. M. (2011). Technical adequacy of response to intervention decisions. *Exceptional Children*, *77*, 335-350.

VanLoy, W. J. (1996). A comparison of adaptive self-referenced testing and classical approaches to the measurement of individual change (Doctoral dissertation, University of Minnesota).

Vellutino, F. R. (2003). Individual differences as sources of variability in reading comprehension in elementary school children. In A. P. Sweet & C. E. Snow (Eds.), Rethinking Reading Comprehension (pp. 5-81). New York: Guilford Press.

Vellutino, F. R., & Scanlon, D. M. (1987). Phonological coding, phonological awareness, and reading ability: Evidence from a longitudinal and experimental study. *Merrill-Palmer Quarterly: Journal of Developmental Psychology,* 33, 321-363.

Vellutino, F. R., & Scanlon, D. M. (1991). The preeminence of phonologically based skills in learning to read. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Liberman* (pp. 237-252). Hillsdale, NJ: Erlbaum.

Vellutino, F. R., Scanlon, D. M., Small, S. G., & Tanzman, M. S. (1991). The linguistic basis of reading ability: Converting written to oral language. *Text*, *11*, 99-133.

Wayman, J., Cho, V. & Johnston, M. (2007). The Data-Informed District: A District-Wide Evaluation of Data Use in the Natrona County School District. Austin: The University of Texas. Retrieved 3/12/08 from http://edadmin.edb.utexas.edu/datause/

Weiss, D. J. (2004). Computerized Adaptive Testing for Effective and Efficient Measurement in Counseling and Education. *Measurement & Evaluation in Counseling & Development (American Counseling Association)*, *37*(2).

Weiss, D. J., & Kingsbury, G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement*, *21*(4), 361-375.

Weiss, D. J., & Kingsbury, G. (2005). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement*, *21*(4), 361-375.

Westberg, K. L. (1993). An Observational Study of Instructional and Curricular Practices Used with Gifted and Talented Students in Regular Classrooms. Research Monograph 93104.

Westberg, K. L., & Daoust, M. E. (2003). The results of the replication of the classroom practices survey replication in two states. *The National Research Center on the Gifted and Talented Newsletter*, 3-8.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child development*, *69*(3), 848-872.

Williams, K.T. (2001). Group Reading Assessment and Classification Evaluation (GRADE).Circle
Pines, MN: AGS Publishing.

Williams, K.T. (2004). Group Mathematics Assessment and Classification Evaluation (GMADE). Circle Pines, MN: AGS Publishing.

Williford, A.P., Maier, M.F., Downer, J.T., Pianta, R.C., & Howes, C. (2013). Understanding how children's engagement and teacher's interactions combine to predict school readiness. Journal of Applied Developmental Psychology, 34, 299-309.

Zeno, S.M., Ivens, S.H., Millard, R.T., Duvvuri, R. (1995). *The educator's word frequency guide.* Brewster, NY: Touchstone Applied Science Associates.

Zickar, M. J., Overton, R. C., Taylor, L. R., & Harms, H. J. (1999). The development of a computerized selection system for computer programmers in a financial services company. In F. Drasgow & J. B. Olson-Buchanan (Eds.), Innovations in computerized assessment (pp. 7-33). Mahwah, NJ: Erlbaum.

Zwaan, R. A. & Rapp, D. N. (2008). Discourse comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), Handbook of Psycholinguistics, 2nd Edition (pp. 725-764). New York: Elsevie.