

Predicting First Graders' Development of Calculation Versus Word-Problem Performance: The Role of Dynamic Assessment

Pamela M. Seethaler, Lynn S. Fuchs, Douglas Fuchs, and Donald L. Compton
Vanderbilt University

The purpose of this study was to assess the value of dynamic assessment (DA; degree of scaffolding required to learn unfamiliar mathematics content) for predicting 1st-grade calculations (CAs) and word problems (WPs) development, while controlling for the role of traditional assessments. Among 184 1st graders, predictors (DA, Quantity Discrimination, Test of Mathematics Ability, language, and reasoning) were assessed near the start of 1st grade. CA and WP were assessed near the end of 1st grade. Planned regression and commonality analyses indicated that for forecasting CA development, Quantity Discrimination, which accounted for 8.84% of explained variance, was the single most powerful predictor, followed by Test of Mathematics Ability and DA; language and reasoning were not uniquely predictive. By contrast, for predicting WP development, DA was the single most powerful predictor, which accounted for 12.01% of explained variance, with Test of Mathematics Ability, Quantity Discrimination, and language also uniquely predictive. Results suggest that different constellations of cognitive resources are required for CA vs. WP development and that DA may be useful in predicting 1st-grade mathematics development, especially WP.

Keywords: mathematics, dynamic assessment, kindergarten, calculations, word problems

Dynamic assessment (DA) involves structuring a learning task, providing feedback or instruction to help the examinee learn the task, and indexing responsiveness to the assisted learning experience as a measure of the examinee's capacity to profit from future instruction. Beginning with Vygotsky's (e.g., 1934/1962) proposal more than 75 years ago, discussions have centered on whether DA might serve as an alternative to the conventional assessment paradigm, in which examinees respond without assistance. The concern is that such static assessments reveal only two states, unaided success or failure (Sternberg, 1996; Tzuriel & Haywood, 1992), which masks distinctions among children who cannot perform a task independently but can succeed with varying levels of assistance.

The literature on DA is diffuse. Studies vary with respect to how DAs are structured. In terms of scoring, DAs may quantify responsiveness to the assisted learning experience as improvement from unassisted pretest to unassisted posttest (e.g., Ferrara, Brown, &

Campione, 1986) or the amount of scaffolding required during the assisted learning experience to reach criterion performance (e.g., Murray, Smith, & Murray, 2000; Spector, 1992). Interaction style is another dimension along which DAs vary. With standardized DAs (e.g., Ferrara et al., 1986), testers rely on a fixed series of prompts; other DAs (e.g., Tzuriel & Feuerstein, 1992) are individualized, with testers addressing the specific obstacles examinees reveal. Yet another dimension along which DAs differ is the nature of the tasks used for the assisted learning experience, which may focus on domain-general cognitive abilities (e.g., Budoff, 1967; Feuerstein, 1979) or on cognitive abilities presumed to underlie the academic domain to be predicted (e.g., Swanson & Howard, 2005) or on domain-specific abilities such as reading or mathematics tasks (e.g., Bransford, Delclos, Vye, Burns, & Hasselbring, 1987; Campione, 1989; Campione & Brown, 1987; Spector, 1992).

The DA literature also varies in terms of research questions and methodological features. Researchers who index pretest–posttest improvement typically investigate whether the DA score distinguishes between individuals with and without a preestablished diagnosis associated with poor learning (e.g., Tzuriel & Feuerstein, 1992). By contrast, researchers who index degree of scaffolding typically examine the value of that score in predicting a learning outcome external to the DA (e.g., Spector, 1992). This second type of study can be further categorized in terms of whether static, competing predictors of outcome are considered and whether the external learning outcome is assessed concurrently with the DA or at a future time. Studies that control for competing predictors or measure the external learning outcome at a later time impose a more stringent test of DA's value. For these reasons, in the present study, we examined the contribution of DA in forecasting future external learning while considering the contribution of competing predictors. We were interested in DA's contribution in predicting

This article was published Online First August 22, 2011.

Pamela M. Seethaler, Lynn S. Fuchs, Douglas Fuchs, and Donald L. Compton, Department of Special Education, Vanderbilt University.

This research was supported by U.S. Department of Education Award Number R324A090039 and by Eunice Kennedy Shriver National Institute of Child Health & Human Development Award Number R01HD059179 and Core Grant Number HD15052 awarded to Vanderbilt University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the U.S. Department of Education or the Eunice Kennedy Shriver National Institute of Child Health & Human Development or the National Institutes of Health.

Correspondence concerning this article should be addressed to Pamela M. Seethaler, 228 Peabody, Vanderbilt University, Nashville, TN 37203. E-mail: pamela.m.seethaler@vanderbilt.edu

two transparently different forms of mathematics development: calculations (CAs) and word problems (WPs).

Prior DA Studies Predicting Mathematics Learning External to the DA While Considering Competing Predictors

To establish the context and rationale for the present investigation, we describe prior studies that have explored DA's contribution in predicting learning external to the DA while controlling for competing predictors. We considered investigations that predicted concurrent or future outcomes as well as DAs of varying structure and design, while limiting our search to studies that focused on mathematics. This netted three relevant investigations. Speece, Cooper, and Kibler (1990) measured first-grade students on a DA task associated with overall cognitive ability: solving matrices. Using a standardized style of interaction, they indexed the number of prompts students required during the assisted learning experience. This score accounted for unique variance, beyond verbal IQ, pre-DA matrices performance, and language ability, in explaining individual differences on the Wide-Range Achievement Test-Arithmetic subtest (WRAT; Wilkinson, 1993), with 2% of the variance in WRAT unique to the DA.

Swanson and Howard (2005) extended Speece et al.'s (1990) study by centering their DA on cognitive resources more specifically presumed to underlie reading or math performance: phonological working memory (i.e., rhyming tasks that required recall of acoustically similar words) and semantic working memory (i.e., digit/sentence tasks that required recall of numerical information embedded in short sentences). The interaction style was individualized, with DA testers choosing among four standardized hints to select the least obvious hint that aligned best with the student's errors. Three DA scores were generated: gain score (highest score obtained with assistance); maintenance score (stability of the highest level obtained after assistance was removed); and probe score (number of hints to achieve highest level). DA scores for phonological working memory were combined into a factor score, as was done for DA semantic working memory; both were used to predict concurrent WRAT performance. Among students averaging 10–12 years of age, static measures of verbal IQ and pre-DA phonological working memory as well as the semantic DA score uniquely accounted for individual differences in WRAT performance; the variance uniquely attributable to the semantic DA factor was 25%.

Therefore, Swanson and Howard (2005) found stronger support for concurrent relations with calculations outcomes for a DA centered on cognitive abilities presumed to underlie mathematics performance than did Speece et al. (1990), whose DA addressed a task associated with more general cognitive ability. However, neither study assessed mathematics development at a future time or used a DA that involved a domain-specific, mathematics task. We identified only one such study. L. S. Fuchs, Fuchs, Compton, et al. (2008) developed a domain-specific DA designed to be novel to the third-grade participants by focusing on early algebraic cognition tasks. In the fall, students were assessed on cognitive resources associated with word-problem performance, initial CA and WP performance, as well as DA. On the basis of random assignment, students received 16 weeks of validated WP instruction or conventional WP instruction. Near the end of the school

year, students were assessed on WP measures proximal and distal to instruction. Structural equation measurement models showed that DA measured a distinct dimension of pretreatment ability. Structural equation modeling showed that the nature of instruction (validated vs. conventional) was sufficient to account for WP development proximal to instruction; yet, language, pretreatment math performance, and DA were uniquely predictive in forecasting WP development more distal to instruction.

In the present study, we extended L. S. Fuchs, Fuchs, Compton et al. (2008) in three ways. First, we examined the role of DA in predicting future mathematics learning using a different domain-specific DA: solving four types of nonstandard expressions. We selected this domain because (a) we could assume it was unfamiliar and sufficiently difficult that most first graders would not be able to solve nonstandard expressions without assistance, but could learn the content with varying amounts of support; (b) we could assume that beginning first graders would have the prerequisite skills to support the assisted learning experience—representations of, Arabic numeral names for, and counting skills associated with small quantities (1–10); (c) we could delineate strategies for solving the nonstandard expressions, which we used to construct clear explanations within a graduated sequence of prompts; and (d) via pilot work, we had established that the DA's four types of nonstandard expressions were increasingly difficult, with later types building on earlier types, such that transfer across the four DA equation types might facilitate higher DA scores.

Beyond focusing on a different DA, a second and more important extension to L. S. Fuchs, Fuchs, Compton, et al. (2008) was that, in the present study, we assessed the utility of DA at the start of first grade, when forecasting learning has proved especially challenging (e.g., Compton et al., 2010; Johnson, Jenkins, Petscher, & Catts, 2009). This is due to difficulty in distinguishing between two types of students who score poorly on static measures: those with poor learning potential who require special intervention versus those whose low score is due to limited prior learning experience but who have good potential to learn in response to generally strong classroom instruction. DA, which indexes how much instructional scaffolding is required to produce learning, may be more useful than static measures for making such distinctions.

The final and most important extension to L. S. Fuchs, Fuchs, Compton, et al. (2008) was that the present study focused on DA's value as a predictor of learning as a function of type of mathematics development: CA versus WP performance. These two forms of mathematics development are transparently different. Whereas CA problems are set up for solution, WPs require students to use linguistic information to construct a problem model: identifying missing information, constructing a number sentence, and setting up a CA problem for solution. Beyond the transparent differences between CA and WP, prior work suggests that the cognitive characteristics underlying development in CA versus WP differ (e.g., L. S. Fuchs, Fuchs, Stuebing, et al., 2008; L. S. Fuchs, Geary, Compton, Fuchs, Hamlett, Seethaler, et al., 2010; Swanson & Beebe-Frankenberger, 2004). For example, processing speed (L. S. Fuchs, Fuchs, Stuebing, et al.) and working memory (Bull & Johnston, 1997) seem to contribute to CA development, whereas WP skill appears to be uniquely predicted by concept formation, nonverbal reasoning, sight-word proficiency, language, and reading (L. S. Fuchs et al., 2006; L. S. Fuchs, Geary, Compton,

Fuchs, Hamlett, Seethaler, et al., 2010; Swanson, 2006). Moreover, skill with WP is significantly linked to CA skill, as L. S. Fuchs et al. (2006) showed with path analysis of arithmetic, arithmetic computation, and arithmetic WP performance of third-grade students, making CA skill necessary but not sufficient for solving WPs.

Although these transparent differences create different demands on students, we identified no prior studies that examined DA's value for these (or other) contrasting subdomains of mathematics performance. We hypothesized that DA's predictive value might differ on the basis of these transparent differences and differences in cognitive correlates. On one hand, the present study's DA, which focuses on balancing equations, may reflect conceptual understanding of arithmetic or the equal sign, which may be more central for WP than CA, because WP (but not CA) development seems to be linked with concept formation and reasoning. The present study's DA incorporates increasingly explicit, conceptually based worked examples, which may draw on the same cognitive resources as does solving WP. On the other hand, the DA includes strategies for deriving answers to CA problems, while avoiding any narrative WP context and instruction; in this way, the DA may better reflect capacity for CA than WP development.

Competing Predictors

In considering the value of DA for predicting CA versus WP development, we were interested in controlling variance associated with predictors that represent traditional (static) domain-specific numerical competencies and domain-general cognitive resources.

Domain-Specific Numerical Competencies

Okamoto (2000, cited in Kalchman, Moss, & Case, 2001) identified the ability to discriminate between quantities as a distinct dimension of kindergarteners' mathematics performance, requiring children not only to distinguish between numerosities but also to move across representational systems. Kindergarteners differ in their ability to discriminate between quantities (i.e., Which number is bigger, 4 or 6?), even when controlling for counting and simple computation (Griffin, Case, & Siegler, 1994). Quantity Discrimination (QD; Chard et al., 2005) is a measure of the speed and accuracy with which children distinguish between and map Arabic numerals onto small numerosities (i.e., students quickly identify the larger quantity in pairs of Arabic numerals ranging from 1 to 10).

In the present study, we included QD because evidence indicates that, at the beginning of first grade, it is a strong predictor of subsequent mathematics achievement (e.g., Chard et al., 2005; Clarke & Shinn, 2004; Lembke & Foegen, 2005). We had three additional reasons for including QD. First, it is commonly used in schools for screening risk for poor mathematics development at the start of first grade. Second, in choosing QD, we opted against measures that require operational manipulations of small quantities as in the Number Sets Test (Geary, Bailey, & Hoard, 2009) or Curriculum-Based Measurement-Computation (L. S. Fuchs et al., 2007), because we sought a measure of early numerical competency that did not require the operations required for our CA and WP outcomes. Finally, by fixing on small quantities (rather than larger magnitudes, as in Number Line Estimation; Siegler &

Booth, 2004), we distinguished variance associated with knowledge of small quantities, which is transparently prerequisite to and should support performance on the DA, from variance associated with what the DA was designed to index: scaffolding required to learn novel mathematics content. Given that QD predicts CA outcomes (Chard et al., 2005; Lembke & Foegen, 2005) and that CA is required for WPs, we anticipated that QD would explain variance in both types of outcomes.

At the same time, because QD is a speeded assessment that focuses on a limited conceptualization of early numerical competency, we also included a power test assessing a broader set of early mathematical competencies, including informal and formal knowledge: the Test of Early Mathematics Ability-3 (TEMA; Ginsberg & Baroody, 2003). In terms of informal knowledge constructs, TEMA assesses numbering (e.g., counting by 1, by 10s, or from a number; identifying number before or after), number comparison (e.g., identifying smaller/larger quantities from collections of items or from Arabic numerals; selecting the Arabic numeral closer to a given numeral), CAs (e.g., solving mental, nonverbal addition problems with sums to 12; demonstrating addition of one or more objects), and understanding of cardinality (e.g., shown a collection of printed stars, counting and saying how many) and equal partitioning (e.g., given a set of tokens, showing how to share the "cookies" fairly between two sisters). The types of formal knowledge assessed are numeral literacy (reading/writing numerals from 1 to 4 digits), number facts (speeded in answering addition or multiplication facts), CAs (performing written or mental calculations of two-digit numerals), and understanding the additive commutativity principle (e.g., $9 + 7$ is the same as $7 + 9$) and base ten (e.g., identifying how many \$10 bills equal one \$100 bill).

In previous predictive validity research with kindergarten or first-grade samples, TEMA has been used as an outcome measure, with coefficients ranging from .33 to .69 (Lembke & Foegen, 2005; Mazzocco & Thompson, 2005; Seethaler & Fuchs, in press; Teisl, Mazzocco, & Myers, 2001). To our knowledge, however, TEMA has not been evaluated as a predictor of outcome. In the context of the present study, in which DA is a lengthy, untimed assessment, we were interested in controlling for variance associated with another lengthy, untimed (but static) assessment of numerical competencies. Controlling in this way for the role of a more established, extended, and comprehensive (but static) measure of early mathematical competencies, which also maps more directly than DA onto the skills measured in our outcomes, created a stringent test for the predictive value of DA. Because TEMA taps multiple forms of mathematical knowledge, we hypothesized it would predict CA and WP outcomes.

Domain-General Cognitive Resources

In the present study, we also assessed the contribution of two domain-general cognitive abilities in predicting outcomes. For this purpose, we included language and reasoning as predictors of future mathematics ability, two important subtests of many traditional IQ tests, which have also been linked to WP development in first graders (e.g., L. S. Fuchs, Geary, Compton, Fuchs, Hamlett, & Bryant, 2010; L. S. Fuchs, Geary, Compton, Fuchs, Hamlett, Seethaler, et al., 2010). Language ability is important to consider given the need to process linguistic information during school

instruction. As Dehaene (1997) suggested, even infants have an informal and primary sense of number, which may be inherent. As children develop intellectually, however, they rely on symbolic and verbal comprehension of numbers, necessary for formal mathematical competence (Jordan, Glutting, & Ramineni, 2010). Jordan, Levine, and Huttenlocher (1995) documented the importance of language ability when kindergarten and first-grade language-impaired children performed significantly lower than nonimpaired peers on WPs. In addition to language, reasoning, measured by completing visually presented patterns, has been identified as a unique predictor or cognitive correlate of various aspects of mathematics development. For example, L. S. Fuchs et al. (2005) demonstrated the importance of reasoning in WP development across first grade, a finding corroborated by Agness and McLone (1987). Seethaler and Fuchs (2006) found reasoning to be a significant correlate of computational estimation skill, and reasoning again emerged as a significant predictor of fifth graders' development of computation with whole and rational numbers (Seethaler, Fuchs, Star, & Bryant, 2011). Because the relations between domain-general abilities may be stronger for WP than for CA (e.g., L. S. Fuchs, Geary, Compton, Fuchs, Hamlett, & Bryant, 2010; L. S. Fuchs, Geary, Compton, Fuchs, Hamlett, Seethaler, et al., 2010), we hypothesized that the domain-general cognitive resources would capture more variance for predicting WP than CA development.

Method

Participants. Two-hundred two participants (i.e., the maximum number of students our research funds permitted, which corresponded to power analyses indicating adequate sample size) were identified. In 61 classrooms in 17 elementary schools (14 Title I; 3 non-Title I) in a southeastern metropolitan school district, one-hundred fifteen students were excluded, who, as part of a different study, would be receiving 16 weeks of mathematics tutoring. (These students were excluded because tutoring was designed to disrupt the predictive value of their initial status on variables like the ones included in the present study.) The remaining 866 students with parental consent were administered two screening measures, the First-Grade Test of Computational Fluency and the First-Grade Test of Mathematics Concepts and Applications (L. S. Fuchs, Hamlett, & Fuchs, 1990; see the Measures section), in their general education classrooms by trained research assistants. A latent class approach (combining screening scores into a single latent factor) produced a three-class solution that specified high, average, and at-risk strata. Stratifying by classroom and strata, two-hundred two students were randomly selected (all schools and classrooms were represented). Eighteen students who moved prior to completing spring testing were comparable to the remaining students on all demographics and on all mathematics performance variables administered in the fall. The present analyses thus included the 184 students for whom fall and spring data have been collected. Of these students, 80 (43.5%) were men, and 112 (60.9%) received free or reduced lunch. In terms of ethnicity, 74 students (40.2%) were African American, 85 (46.2%) were Caucasian, 13 (7.1%) were Hispanic, eight (4.3%) were Asian, and four (2.2%) were "other." Nine students (5.5%) received special education services for a learning (1.1%), speech (3.3%), or lan-

guage (.5%) disability; six students (3.3%) were English language learners.

Screening measures to obtain a representative sample. Because the two screening measures were used to select a representative sample, items on the screeners represented a range of difficulty. Because the screeners were used for sample selections, they were not used as predictors in the study. The *First-Grade Test of Computational Fluency* (L. S. Fuchs, Hamlett, & Fuchs, 1990) is a single page of computation items representing the first-grade curriculum: nine single-digit addition items, nine single-digit subtraction items, two double-digit addition items without regrouping, three double-digit subtraction items without regrouping, and two single-digit addition items with three addends. Items are displayed in five rows of five items, and students have 2 min to write answers next to or below each item. They are instructed to first try the items that seem easier and then to go back to harder items. The score is the number of correct digits. Coefficient alpha for this sample was .87.

The *First-Grade Test of Mathematics Concepts and Applications* (L. S. Fuchs et al., 1990) comprises 25 items, displayed on three pages. Items represent the first-grade curriculum, including numeration, concepts, geometry, measurement, applied computation, money, charts, graphs, and word problems. The tester reads each item out loud, without reading key numbers or number words. As the tester reads, students follow along on a paper copy, while covering other items on the page. Before moving to the next item, the tester gives students sufficient time to respond (15 or 20 s, as dictated by standard directions, based on field data indicating adequate response time for almost all first graders). The score is the number of correctly answered items. Coefficient alpha for this sample was .85.

Predictor measures. With *QD* (Chard et al., 2005; Lembke & Foegen, 2009; Research Institute on Progress Monitoring, 2009), students have 1 min to select the larger of two numbers (ranging from 1 to 10), presented in 56 boxes across two pages (28 per page). It is individually administered. Test-retest reliability is .85-.99 (Clarke, Baker, Smolkowski, & Chard, 2008).

TEMA-3 (Ginsburg & Baroody, 2003) assesses informal and formal mathematics knowledge for children 3 years 0 months through 8 years 11 months. It comprises 72 items of increasing difficulty, with multiple trials for each item. The tester scores each trial as right or wrong by determining whether the number of correct trials warrants a point for that item. For example, some items require two of three correct trials to earn a point; other items require all trials answered correctly. Students reach a ceiling when five consecutive items do not meet criteria to earn a point; the tester ensures a basal of five consecutively correct items. Testing takes up to 45 min. Coefficient alpha for 6-year-olds is .95.

To assess language, the *Wechsler Abbreviated Scale of Intelligence (WASI) Vocabulary* test (Psychological Corporation, 1999) was used, which measures expressive vocabulary and verbal knowledge. The examiner presents words for the student to define and immediately scores the response as 0, 1, or 2 points depending on quality. For the first four items, the student sees a picture of an item to define; for the remaining items, the examiner presents the word orally. Testing is discontinued after five consecutive scores of zero. Zhu (1999) reported split-half reliability at .86-.87; the correlation with the Wechsler Intelligence Scale for Children (Wechsler, 1999) is .72.

To assess reasoning, we used *WASI Matrix Reasoning* (Psychological Corporation, 1999), which measures reasoning skill with pattern completion, classification, analogy, and serial reasoning. Students see a color picture of a matrix with one piece missing and select the correct piece to complete the picture from five choices displayed below the picture. The examiner awards 1 point for each correct answer; testing is discontinued after four out of five missed items. As per Zhu (1999), reliability is .94.

Balancing Equations Dynamic Assessment (DA; Seethaler & Fuchs, 2010a) measures the degree of scaffolding required to learn unfamiliar mathematics content, specifically, solving for missing variables in nonstandard addition and subtraction expressions. The DA comprises four types of equations of increasing difficulty. Testers present strategies of increasing explicitness to teach students to balance both sides of the equation; students progress to the level of explicitness they require to achieve mastery of that equation type, at which time they advance to the next, more difficult equation type. Beyond the reasons already discussed earlier in this article, balancing equations were chosen as the DA task because elementary school students often misinterpret the equal sign (=) as an operational rather than as a relational symbol (McNeil & Alibali, 2005; Sherman & Bisanz, 2009) and because solving equations with missing numbers is important for higher level mathematics skills and thus is valuable content for students to learn.

Four equation types constitute the DA. *Equation Type A* requires solving for a missing variable in the first or second position in equations that use 1 as an addend and for which the sum is not greater than 9 (e.g., $__ + 1 = 4$ or $8 + __ = 9$). With *Equation Type B*, students solve for a missing variable in the first or second position in equations that do not use 1 as an addend and for which the sum is no greater than 9 (e.g., $__ + 2 = 6$ or $3 + __ = 5$). For *Equation Type C*, students solve for a missing variable in the first position in subtraction equations with minuends no greater than nine (e.g., $__ - 7 = 2$). *Equation Type D* requires students to solve for a missing variable in any of four positions, with sums on both sides of the equal sign, none of which exceed 9 (e.g., $__ + 5 = 3 + 4$ or $3 + 6 = 5 + __$). The equation types are presented so that success with an earlier equation type should promote understanding of a subsequent equation type.

The administration and scoring procedures follow L. S. Fuchs, Fuchs, Compton, et al. (2008). Within each equation type, the tester begins by assessing mastery of that equation type. If mastery is demonstrated, the student advances to the next equation type. If not, instructional scaffolding begins with the least explicit level of scaffolding. Mastery testing then recurs. If mastery is achieved, the student progresses to the next equation type. If not, the next more explicit level of instructional scaffolding is presented, and mastery testing follows. In this way, a maximum of four or five (depending on equation type) increasingly explicit levels of instructional scaffolding are used. If the student fails to master an equation type after the final level of scaffolding for the equation type, the DA is terminated.

Each mastery test comprises six items representing the targeted equation type. Items repeat across alternate test forms (used for successive mastery testing within that equation type), but items are presented in different orders. Mastery test items are not used for instructional scaffolding. If a student writes nothing on a mastery test for 5 s, the tester prompts the student by asking, "Can you try

this one?" and pointing to the first item. If after 15 additional s the student still has not written anything, the tester asks, "Are you still working or are you stuck?" If student responds that he or she is stuck or if 15 additional s elapse with no observable attempt to solve the problem, the tester begins the next level of instructional scaffolding.

Each equation type includes four or five instructional scaffolding levels, each of which has two teaching items with which the examiner models and explains a problem-solving strategy. The scaffolding is scripted to ensure consistency in language and procedures. Examiners maintain student attention with frequent questions and participation. The scaffolding levels increase in instructional explicitness. Within an equation type, the first (least explicit) level only defines relevant mathematical terms (e.g., *equal means the same as*; a *plus sign means to add more*). With the second scaffolding level, the examiner uses a balance scale and 2-in. plastic teddy bear manipulatives to demonstrate balancing both sides of the equation. A 4-in. \times 2-in. equal sign (=) is printed on a white card, which is affixed to the center of the scale; student attention is drawn to note the sides of the equation as parallel to the sides of the scale, with manipulatives used to represent the amounts in the equation. The third level of scaffolding provides instruction in solving the equations while using tally marks drawn on paper to represent the numerical amounts. The next scaffolding level provides instruction in solving the equations in conjunction with an 8-in. number line printed on a piece of cardstock; students are taught to move their finger to count spaces on the number line while solving the equations. This is designed to build understanding of the inverse relation between addition and subtraction (e.g., for $___ + 2 = 6$, students count on 4 from 2 to get to 6, revealing that $6 - 2 = 4$). The final, most explicit scaffolding level increases the support for the student to successfully apply the number line strategy for building understanding of the inverse relation between addition and subtraction. Toward that end, different colored markers on the number line represent different parts of the equation. (Equation Types A and C have five levels of instruction, whereas Equation Types B and D have four because new mathematical terms are not introduced for Equation Type B or D.) Worked examples used during instructional scaffolding are not displayed during mastery testing; however, all materials necessary for applying the strategies taught during scaffolding are always displayed on the testing table (even during the initial mastery test). Students are not penalized by their choice of problem-solution strategies. An equation type is deemed mastered when at least five of the six items are answered correctly, at which time the examiner progresses to the next DA skill.

DA scores range from 0 to 22. Zero indicates a student did not master any equation type; 22 reflects a student mastering each of the four equation types on the first administration of the mastery test (i.e., without any instructional scaffolding; Equation Types A, B, C, and D are worth 6, 5, 6, and 5 points, respectively, because Equation Types A and C have five levels of instructional scaffolding, whereas Equation Types B and D have only four levels). A tester subtracts 1 point from the maximum of 22 each time a level of instructional scaffolding is required. For example, if a student demonstrates mastery on the first administration of the mastery test for Equation Types A, B, and C (without any instructional scaffolding), but requires two levels of instructional scaffolding to master Equation Type D, the examiner subtracts 2 points from the

maximum of 22, awarding a score of 20. By contrast, if a student requires three levels of instructional scaffolding to master Equation Type A, four levels of instructional scaffolding to master Equation Type B, and fails to master Equation Type C (thereby terminating the DA such that Equation Type D is not presented), the student loses 3 points for Equation Type A, 4 points for Equation Type B, 6 points for Equation Type C, and 5 points for Equation Type D, for a score of 4. Internal consistency reliability was indexed by correlating the score from each DA equation type with the DA total score, using the subset of students who had not reached a ceiling on performance prior to the administration of that DA equation type. For Equation Type A, $r = .90$; for Equation Type B, $r = .86$; for Equation Type C, $r = .82$; for Equation Type D, $r = .84$.¹

End-of-first-grade mathematics outcome measures. Both outcome measures are power tests, in which time limits are ample for all students to complete the items they are capable of answering. To assess CA performance, the Arithmetic subtest of the WRAT-3 (Wilkinson, 1993) was used, with which students have 10 min to write answers to 40 calculation items of increasing difficulty (kindergarten through Grade 12). None of the students in the present study used all 10 min. According to the manual, reliability is .94.

To assess WP performance, Story Problems (Jordan & Hanich, 2000) was used, which comprises 14 single-step addition and subtraction WPs of the types most often encountered in the primary grades: combine, compare, and change. The tester reads each item aloud and provides one additional reading if requested to do so. Students have 30 s to answer on their paper copy of the test, which they have available throughout testing so they can read along while the tester reads or refer back to problems as they derive solutions. Then, the tester reads the next problem. Students completed all work within the 30-s time limit for each problem. Coefficient alpha on this sample was .90.

Procedure. In late September and October, the WASI Vocabulary, WASI Matrix Reasoning, QD, and TEMA were administered. In October and November, the DA was administered. Competing predictors were administered before the DA so that the teaching in the DA would not influence estimates of performance on other measures. The goal was to complete the two sessions for each student within 1 school week. This happened for a majority of students, but sometimes took an additional week (if the student was absent or otherwise unavailable for testing in the targeted time frame). In May, the CA and WP tests were administered. Tests were administered individually by graduate students who demonstrated 100% accuracy during practice administration of the measures. All testing sessions were audiotaped, and 16% of the sessions, distributed equally across testers, were randomly sampled to assess fidelity of administration. Scoring agreement was 99.4%. All data were independently entered into two databases, which were compared for discrepancies and resolved against the original protocols.

Data Analysis and Results

Table 1 provides means, standard deviations (*SDs*), and correlations among the predictors and outcomes. We provide raw scores as well as standard scores when applicable. In Table 2, we show results of regression analyses, into which all predictors were en-

tered simultaneously. In explaining CA development, the combination of QD, TEMA, language, reasoning, and DA accounted for 57.7% of the variance, $F(5, 178) = 48.63$, $p < .001$. Four of the five predictors made a unique contribution in explaining individual differences in CA development. QD was the strongest contributor, with TEMA a close second, followed by DA; language and reasoning were not uniquely predictive. By contrast, in explaining WP development, the combination of the five predictors accounted for a larger percentage (71.6%) of variance, $F(5, 178) = 89.67$, $p < .001$. The predictors that made a unique contribution in explaining individual differences in WP development also differed from those involved in CA. For WP, DA was the strongest predictor, followed by language, TEMA, and QD; reasoning was not uniquely predictive.

To supplement these regressions, we conducted a complete commonality analysis (Beaton, 1973; Capraro & Capraro, 2001; Newton & Spurell, 1967; Nimon, Lewis, Kane, & Haynes, 2008) specifying the unique and shared variance associated with each predictor and each combination of predictors for CA development and for WP development. See Table 3, in which we list the predictors and all possible combinations thereof in the first column. In the second and fourth columns, we show coefficients expressing the *proportion of total variance explained* by the predictor(s). In the third and fifth columns, we translated the proportions of total variance to *percentages of explained variance*. To derive *percentages of explained variance*, we took the coefficient expressing the proportion of total variance explained by a given predictor and divided that coefficient by the total amount of explained variance across predictors, then multiplied by 100. For example, in explaining individual differences in WP development, the coefficient expressing the proportion of total variance explained by DA is .086, and the coefficient denoting the total proportion of variance explained by all the predictors (individually and in combination) is .716. To derive the *percentage of explained variance* accounted by DA in WP development, we divided .086 by .716, which equals .1201, and then multiplied by 100: 12.01%. In the discussion that follows, we rely on percentage of explained variance to facilitate comparisons across CA and WP.

Discussion

The purpose of this study was to assess the value of DA in predicting two transparently different forms of mathematics development: CA and WP. In evaluating the role of DA, we controlled for traditional assessments, in which examinees respond without assistance. In this way, we contrasted the assessment of what students already know (static tests) against the assessment of students' capacity to learn (DA). To create a stringent test of DA's value in forecasting development, we included different types of static assessments in our model.

One of those static measures, QD, assesses the speed and accuracy with which children distinguish between and map Arabic numerals onto small numerosities. Including the QD predictor permitted us to distinguish variance associated with knowledge of small quantities, which is a transparent prerequisite to and should support DA performance, from variance associated with what the

¹ Contact Pamela M. Seethaler for more information on the DA.

Table 1
Means, Standard Deviations, and Correlations^a Among Predictor and Outcome Measures ($n = 184$)

Measure	Raw score		Standard score ^b		L	R	QD	T	DA	CA	WP
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)							
Predictors											
Language (L)	19.20	(6.51)	43.72	(10.23)	—						
Reasoning (R)	9.95	(5.78)	50.30	(9.88)	.42	—					
Quantity Discrimination (QD)	31.58	(10.13)			.39	.40	—				
TEMA (T)	39.22	(9.96)	100.21	(14.08)	.50	.54	.64	—			
Dynamic Assessment (DA)	9.20	(8.08)			.48	.62	.53	.68	—		
Outcomes											
Calculations (CA)	18.59	(3.45)	100.70	(15.84)	.44	.49	.64	.69	.63	—	
Word Problems (WP)	7.01	(4.40)			.59	.58	.60	.71	.78	.71	—

Note. Language is Wechsler Abbreviated Scale of Intelligence (WASI) Vocabulary; Reasoning is WASI Matrix Reasoning. TEMA = Test of Early Mathematics Ability, third edition.

^a All correlations significant at $p < .01$. ^b Standard scores for WASI Vocabulary and WASI Matrix Reasoning are *T* scores ($M = 50$; $SD = 10$); for Calculations (Wide-Range Achievement Test-Arithmetic), the mean is 100 ($SD = 15$).

DA was designed to index: ability to learn mathematics. So it is noteworthy that QD, which takes only 1 min to administer and involves none of the addition or subtraction demands of the CA outcome, was the strongest single predictor of CA development across first grade, uniquely accounting for 8.84% of explained variance. The power of QD in forecasting individual differences in CA development underscores children's appreciation of magnitudes as foundational to formal mathematics learning (e.g., Berch, 2005; Dehaene, 1997; Okamoto & Case, 1996), while illustrating the importance of prerequisite knowledge as a condition of future learning.

QD's predictive role is especially noteworthy given that we included in our model another, more comprehensive and lengthy static index of incoming mathematics performance, TEMA, which in part also assesses understanding of small magnitudes. Another component of the TEMA battery is incoming CA skill, which creates better alignment than QD with the CA outcome. So it is not

surprising that TEMA also accounted for a sizable percentage of explained variance in CA development (6.41%). Even so, it is also impressive that with the two static domain-specific assessments already uniquely accounting for 15.25% of explained variance, DA made a uniquely significant, albeit smaller, contribution to predicting CA outcome, accounting for 2.77% of explained variance ($\beta = .20$). At the same time, as revealed in the commonality analysis, these three domain-specific predictors also shared a substantial amount of additional variance in predicting CA. So the bulk of explained variance was attributable to domain-specific predictors, with the domain-general language and reasoning variables, traditionally incorporated in intelligence tests to predict school learning, failing to achieve statistical significance (β s = .05 and .06).

With respect to the major purpose of the present study, assessing whether the contribution of these predictors differed for CA versus WP development, findings were interesting. Whereas DA was overshadowed by QD and TEMA in predicting CA, DA was the strongest single contributor to WP development, uniquely accounting for 12.01% of explained variance—nearly one and one-half times as much explained variance as QD accounted for in CA (8.84%). For DA's prediction of WP development, beta was a sizeable .45. Moreover, although the contributions of QD and TEMA were significant, betas were substantially smaller (.14 for QD; .18 for TEMA), as were the percentages of explained variance (1.54% for QD; 1.96% for TEMA)—despite that TEMA was the only predictor to explicitly assess WPs (i.e., nonverbal addition problems with sums to 12).

The second way in which findings differed for WP development concerns the role of the domain-general predictor variables. Whereas these language and reasoning predictors failed to make a significant contribution to CA development, language (but not reasoning) was uniquely predictive of WP development, with beta equal to .20 (the same as DA in predicting CA development) and with 4.05% of the explained variance in WP development uniquely attributable to language. Moreover, whereas DA shared relatively little variance with language and reasoning in predicting CA (2.27%), the corresponding figure in predicting WP development was more than 3 times larger

Table 2
Regression Models Predicting Individual Differences in First-Grade Mathematics Development

Outcome	<i>B</i>	<i>SE</i>	β	$t(5, 178)$	<i>p</i>
Calculations					
Constant	-9.58	0.84		11.47	<.001
Language	0.03	0.03	0.05	0.91	.360
Reasoning	0.04	0.04	0.06	0.98	.329
QD	0.10	0.02	0.30	4.62	<.001
TEMA	0.11	0.03	0.30	3.94	<.001
DA	0.08	0.03	0.20	2.61	.010
Word Problems					
Constant	-3.43	0.88		-3.91	<.001
Language	0.14	0.03	0.20	4.28	<.001
Reasoning	0.04	0.04	0.05	1.02	.311
QD	0.06	0.23	0.14	2.65	.009
TEMA	0.08	0.28	0.18	2.92	.004
DA	0.24	0.03	0.45	7.35	<.001

Note. Language is Wechsler Abbreviated Scale of Intelligence (WASI) Vocabulary; Reasoning is WASI Matrix Reasoning. QD = Quantity Discrimination; TEMA = Test of Early Mathematics Ability, third edition; DA = Dynamic Assessment.

Table 3
Commonality Analysis for Predicting End-of-Year Mathematics Development

Variable	Calculations development		Word-problem development	
	Coefficient: Proportion of variance explained	Percentage of explained variance	Coefficient: Proportion of variance explained	Percentage of explained variance
Unique to:				
Quantity Discrimination (QD)	.051	8.84	.011	1.54
Test of Early Mathematics Ability (TEMA)	.037	6.41	.014	1.96
Dynamic Assessment (DA)	.016	2.77	.086	12.01
Language (L)	.002	0.35	.029	4.05
Reasoning (R)	.002	0.35	.002	0.28
Common to:				
QD + TEMA	.062	10.75	.018	2.51
QD + DA	.010	1.73	.011	1.54
QD + L	.002	0.35	.003	0.42
QD + R	.000	-0.03	.000	0.00
TEMA + DA	.027	4.68	.042	5.87
TEMA + L	.005	0.87	.009	1.26
TEMA + R	.004	0.69	.002	0.28
DA + L	.002	0.35	.017	2.37
DA + R	.008	1.39	.025	3.49
L + R	.001	0.17	.002	0.32
QD + TEMA + DA	.066	11.44	.057	7.96
QD + TEMA + L	.011	1.91	.011	1.54
QD + TEMA + R	.005	0.87	.002	0.28
QD + DA + L	.002	0.35	.004	0.56
QD + DA + R	.003	0.52	.003	0.42
QD + L + R	.000	0.00	.000	0.00
TEMA + DA + L	.009	1.56	.025	3.49
TEMA + DA + R	.026	4.51	.039	5.45
TEMA + L + R	.002	0.35	.003	0.42
DA + L + R	.003	0.52	.015	2.09
QD + TEMA + DA + L	.034	5.89	.048	6.70
QD + TEMA + DA + R	.066	11.44	.059	8.24
QD + TEMA + L + R	.004	0.69	.003	0.42
QD + DA + L + R	.002	0.35	.003	0.42
TEMA + DA + L + R	.020	3.47	.047	6.56
QD + TEMA + DA + L + R	.097	16.81	.127	17.74
Total	.577	100.00	.716	100.00

Note. Language is Wechsler Abbreviated Scale of Intelligence (WASI) Vocabulary. Reasoning is WASI Matrix Reasoning.

(8.02). In these ways, DA appears to invoke the need for the same kinds of language and reasoning skills that help children profit from WP classroom instruction.

It is therefore interesting to consider that in this study's DA, students solved for missing numbers in mathematical expressions without the need to process text. In this way, the DA is more transparently aligned with CA than WP, raising questions about why DA was more predictive of WP than CA development. A possible explanation is that DA represents a measure of conceptual understanding of arithmetic or understanding of the equal sign, either of which may be required more for WP than CA. Given that DA did not require the processing of text, as in WP, it is also curious that DA shared variance with language in predicting WP development. A possible explanation for this finding is that the DA nevertheless involves language ability because its instructional scaffolding is offered via language as the examiner explains problem-solution concepts and strategies. In fact, much of the school curriculum is delivered via oral language. In mathematics, evidence suggests that language plays an important role in the acquisition of early numeracy concepts and skills (Fletcher, Lyon,

Fuchs, & Barnes, 2007; Hodent, Bryant, & Houde, 2005), whereas Seethaler et al. (2011) showed that language is a unique predictor of CA with rational numbers among fifth graders. Strength with oral language may support teachers' explanations, facilitating insights into novel concepts as required in the DA. It also suggests that the cognitive resources involved in DA may rely on similar types of mental flexibility, manipulation of symbolic associations, and maintenance of multiple representations that are reflected in oral language and reasoning abilities. This may be more true for WP than for CA, at least in part, because classroom instruction is less explicit and procedural for WP than CA.

More generally, results suggest that different types of mathematics depend on distinct aspects of mathematical cognition, as previously shown. For example, L. S. Fuchs, Fuchs, Compton, et al. (2006) documented links among skill with arithmetic, procedural calculations, and WPs, even as distinct constellations of predictors emerged for each area of mathematics performance. In a related way, Hart, Petrill, and Thompson (2009) found support for different genetic and environmental influences on students' WP versus CA performance. And, as illustrated in the present

study, for forecasting mathematics development, the relative value of predictors, including DA, differs depending on the form of mathematics to be predicted.

For this reason, present findings suggest that different screening processes may be required to identify risk for poor CA versus WP development to permit targeted intervention to begin early, before severe academic deficits become intractable. For forecasting CA development, a brief measure of magnitude comparison, such as QD, may provide value as a universal screener (for syntheses, see Gersten, Jordan, & Flojo, 2005; Seethaler & Fuchs, 2010b). At the same time, research (e.g., Compton et al., 2010; Johnson et al., 2009) illustrates how brief universal screening at first grade produces high rates of false positives. In this vein, we note that QD alone uniquely accounted for only 5.10% of the *total* variance in individual differences in first-grade CA development; instead, a combination of predictors was required to account for a substantial proportion of variance. Therefore, although QD may serve as an efficient universal screen for identifying risk for poor CA development, follow-up assessment for children who fail that universal screen may be needed to accurately classify risk, perhaps using measures such as DA or TEMA. However, for identifying risk for poor WP development, the more time-consuming DA along with a measure of language ability may provide a sounder basis.

In closing, we note that, in the present study, we did not consider the universe of possible predictors. Other domain-general abilities sometimes associated with CA or WP development might have been incorporated. These include working memory (e.g., Swanson & Beebe-Frankenberger, 2004), phonological processing (e.g., L. S. Fuchs et al., 2005), or processing speed (e.g., Bull & Johnston, 1997). Moreover, we did not consider a second major form of early numerical competency, approximate representations of larger quantities—typically indexed with Number Line Estimation (Siegler & Booth, 2004)—for which substantial empirical support exists (e.g., Booth & Siegler, 2006, 2008; Laski & Siegler, 2007; Siegler & Booth, 2004). With this caveat in mind, we draw two major conclusions. First, as shown in prior research in mathematics (e.g., L. S. Fuchs, Fuchs, Compton et al., 2008; Swanson & Howard, 2005) and reading (e.g., D. Fuchs, Compton, Fuchs, Bouton, & Caffrey, 2011), results underscore the potential value of DA (which provides insight into what a student is capable of learning in response to varying degrees of instructional scaffolding) over and beyond traditional, static measures (which are limited to a snapshot of what a student presently knows). Second, findings suggest that the relative value of these various types of learning potential measures differs as a function of whether CA or WP development is the predicted outcome. Future work should investigate whether similar distinctions among different forms of learning potential measures apply to other aspects of mathematics learning, while including a more comprehensive set of predictors, to gain additional insight into the nature of mathematics development and the role of DA in predicting its development.

References

- Agness, P. J., & McLone, D. G. (1987). Learning disabilities: A specific look at children with spina bifida. *Insights*, 8–9.
- Beaton, A. E. (1973). *Commonality* (ERIC Document Reproduction Service No. ED 111 829). Princeton, NJ: Educational Testing Service.
- Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical difficulties. *Journal of Learning Disabilities*, 38, 333–339. doi:10.1177/00222194050380040901
- Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, 41, 189–201. doi:10.1037/0012-1649.41.6.189
- Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, 79, 1016–1031. doi:10.1111/j.1467-8624.2008.01173.x
- Bransford, J. C., Delclos, V. R., Vye, N. J., Burns, M. S., & Hasselbring, T. S. (1987). State of the art and future directions. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 479–496). New York, NY: Guilford Press.
- Budoff, M. (1967). Learning potential among institutionalized young adult retardates. *American Journal of Mental Deficiency*, 72, 404–411.
- Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology*, 65, 1–24. doi:10.1006/jecp.1996.2358
- Campione, J. C. (1989). Assisted assessment: A taxonomy of approaches and an outline of strengths and weaknesses. *Journal of Learning Disabilities*, 22, 151–165. doi:10.1177/002221948902200303
- Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82–115). New York, NY: Guilford Press.
- Capraro, R. M., & Capraro, M. M. (2001). Commonality analysis: Understanding variance contributions to overall canonical correlation effects of attitude toward mathematics on geometry achievement. *Multiple Linear Regression Viewpoints*, 27, 16–23.
- Chard, D., Clarke, B., Baker, B., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment Issues in Special Education*, 30, 3–14.
- Clarke, B., Baker, S., Smolkowski, K., & Chard, D. J. (2008). An analysis of early numeracy curriculum-based measurement: Examining the role of growth in student outcomes. *Remedial and Special Education*, 29, 46–57. doi:10.1177/0741932507309694
- Clarke, B., & Shinn, M. R. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review*, 33, 234–248.
- Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Gilbert, J. K., Barquero, L. A., . . . Crouch, R. C. (2010). Selecting at-risk first-grade readers for early intervention: Eliminating false positives and exploring the promise of a two-stage gated screening process. *Journal of Educational Psychology*, 102, 327–340. doi:10.1037/a0018448
- Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York, NY: Oxford University Press.
- Ferrara, R. A., Brown, A. L., & Campione, J. C. (1986). Children's learning and transfer of inductive reasoning rules: Studies of proximal development. *Child Development*, 57, 1087–1099. doi:10.2307/1130433
- Feuerstein, R. (1979). *The dynamic assessment of retarded performers. The Learning Potential Assessment Device, theory, instruments, and techniques*. Baltimore, MD: University Park Press.
- Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: From identification to intervention*. New York, NY: Guilford Press.
- Fuchs, D., Compton, D. L., Fuchs, L. S., Bouton, B., & Caffrey, E. (2011). The construct and predictive validity of a dynamic assessment of young children learning to read: Implications for RTI frameworks. *Journal of Learning Disabilities*, 44, 339–347.
- Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, 97, 493–513. doi:10.1037/0022-0663.97.3.493

- Fuchs, L. S., Fuchs, D., Compton, D. L., Bryant, J. D., Hamlett, C. L., & Seethaler, P. M. (2007). Mathematics screening and progress monitoring at first grade: Implications for responsiveness-to-intervention. *Exceptional Children, 73*, 311–330.
- Fuchs, L. S., Fuchs, D., Compton, D. L., Hollenbeck, K. N., Craddock, C. F., & Hamlett, C. L. (2008). Dynamic assessment of algebraic learning in predicting third graders' development of mathematical problem solving. *Journal of Educational Psychology, 100*, 829–850. doi:10.1037/a0012657
- Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., . . . Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*, 29–43. doi:10.1037/0022-0663.98.1.29
- Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*, 30–47. doi:10.1037/0022-0663.100.1.30
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*, 1520–1533. doi:10.1111/j.1467-8624.2010.01489.x
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., . . . Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology, 46*, 1731–1746. doi:10.1037/a0020662
- Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1990). *First-Grade Test of Computational Fluency, First-Grade Test of Concepts and Applications*. Available from L. S. Fuchs, 328 Peabody, Vanderbilt University, Nashville, TN 37203.
- Geary, D. C., Bailey, D. H., & Hoard, M. K. (2009). Predicting mathematical achievement and mathematical learning disability with a simple screening tool: The number sets test. *Journal of Psychoeducational Assessment, 27*, 265–279. doi:10.1177/0734282908330592
- Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*, 293–304. doi:10.1177/00222194050380040301
- Ginsburg, H., & Baroody, A. (2003). *Test of Early Mathematics Ability* (3rd ed.). Austin, TX: Pro-Ed.
- Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 24–29). Cambridge, MA: MIT Press.
- Hart, S. A., Petrill, S. A., Thompson, L. A., & Plomin, R. (2009). The ABCs of math: A genetic analysis of mathematics and its links with reading ability and general cognitive ability. *Journal of Educational Psychology, 101*, 388–402. doi:10.1037/a0015115
- Hodent, C., Bryant, P., & Houde, O. (2005). Language-specific effects on number computation in toddlers. *Developmental Science, 8*, 420–423. doi:10.1111/j.1467-7687.2005.00430.x
- Johnson, E. S., Jenkins, J. R., Petscher, Y., & Catts, H. W. (2009). How can we improve the accuracy of screening instruments? *Learning Disabilities Research and Practice, 24*, 174–185. doi:10.1111/j.1540-5826.2009.00291.x
- Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences, 20*, 82–88. doi:10.1016/j.lindif.2009.07.004
- Jordan, N. C., & Hanich, L. B. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities, 33*, 567–578. doi:10.1177/002221940003300605
- Jordan, N. C., Levine, S. C., & Huttenlocher, J. (1995). Calculation abilities in young children with different patterns of cognitive functioning. *Journal of Learning Disabilities, 28*, 53–64. doi:10.1177/002221949502800109
- Kalchman, M., Moss, J., & Case, R. (2001). Psychological models for the development of mathematical understanding: Rational numbers and functions. In S. Carver & D. Klahr (Eds.), *Cognition and instruction* (pp. 1–38). Mahwah, NJ: Erlbaum.
- Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 78*, 1723–1743. doi:10.1111/j.1467-8624.2007.01087.x
- Lembke, E., & Foegen, A. (2005, February). *Monitoring student progress in early math*. Paper presented at the Pacific Coast Research Conference, San Diego, CA.
- Lembke, E., & Foegen, A. (2009). Identifying early numeracy indicators for kindergarten and first-grade students. *Learning Disabilities Research & Practice, 24*, 12–20. doi:10.1111/j.1540-5826.2008.01273.x
- Mazzocco, M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice, 20*, 142–155. doi:10.1111/j.1540-5826.2005.00129.x
- McNeil, N. M., & Alibali, M. W. (2005). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition & Development, 6*, 285–306. doi:10.1207/s15327647jcd0602_6
- Murray, B. A., Smith, K. A., & Murray, G. G. (2000). The test of phoneme identities: Predicting alphabetic insight in pre-alphabetic readers. *Journal of Literacy Research, 32*, 421–477. doi:10.1080/10862960009548087
- Newton, R. G., & Spurrell, D. J. (1967). Examples of the use of elements for classifying regression analysis. *Applied Statistics, 16*, 165–172. doi:10.2307/2985778
- Nimon, K., Lewis, M., Kane, R., & Haynes, R. M. (2008). An R package to compute commonality coefficients in the multiple regression case: An introduction to the package and a practical example. *Behavior Research Methods, 40*, 457–466. doi:10.3758/BRM.40.2.457
- Okamoto, Y., & Case, R. (1996). Exploring the microstructure of children's central conceptual structures in the domain of number. *Monographs of the Society for Research in Child Development, 61*, 27–58.
- Psychological Corporation. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Harcourt Brace & Company.
- Research Institute on Progress Monitoring. (2009). *Early numeracy indicators (Number Identification, Quantity Discrimination, Missing Number, Mixed Numeracy)*. Retrieved from <http://www.progressmonitoring.org/pdf/TREarlymath6.pdf>
- Seethaler, P. M., & Fuchs, L. S. (2006). The cognitive correlates of computational estimation skill among third-grade students. *Learning Disabilities Research and Practice, 21*, 233–243. doi:10.1111/j.1540-5826.2006.00220.x
- Seethaler, P. M., & Fuchs, L. S. (2010a). *Balancing Equations Dynamic Assessment*. Available from P. M. Seethaler, 238 Peabody, Nashville, TN 37203.
- Seethaler, P. M., & Fuchs, L. S. (2010b). The predictive utility of kindergarten screening for math difficulty. *Exceptional Children, 77*, 37–59.
- Seethaler, P. M., & Fuchs, L. S. (in press). Using curriculum-based measurement to monitor kindergarteners' mathematics development. *Assessment for Effective Intervention*. doi:10.1177/1534508411413566
- Seethaler, P. M., Fuchs, L. S., Star, J. R., & Bryant, J. (2011). *The cognitive predictors of computational skill with whole versus rational numbers: An exploratory study*. Manuscript submitted for publication.
- Sherman, J., & Bisanz, J. (2009). Equivalence in symbolic and nonsymbolic contexts: Benefits of solving problems with manipulatives. *Journal of Educational Psychology, 101*, 88–100. doi:10.1037/a0013156
- Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*, 428–444. doi:10.1111/j.1467-8624.2004.00684.x
- Spector, J. E. (1992). Predicting progress in beginning reading: Dynamic

- assessment of phonemic awareness. *Journal of Educational Psychology*, 84, 353–363. doi:10.1037/0022-0663.84.3.353
- Speece, D. L., Cooper, D. H., & Kibler, J. M. (1990). Dynamic assessment, individual differences, and academic achievement. *Learning and Individual Differences*, 2, 113–127. doi:10.1016/1041-6080(90)90019-D
- Sternberg, R. J. (1996). *Successful intelligence*. New York, NY: Simon & Schuster.
- Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psychology*, 98, 265–281. doi:10.1037/0022-0663.98.2.265
- Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, 96, 471–491. doi:10.1037/0022-0663.96.3.471
- Swanson, H. L., & Howard, C. B. (2005). Children with reading disability: Does dynamic assessment help in the classification? *Learning Disability Quarterly*, 28, 17–34. doi:10.2307/4126971
- Teisl, J. T., Mazzocco, M. M., & Myers, G. F. (2001). The utility of kindergarten teacher ratings for predicting low academic achievement in first grade. *Journal of Learning Disabilities*, 34, 286–293. doi:10.1177/002221940103400308
- Tzuriel, D., & Feuerstein, R. (1992). Dynamic group testing for prescriptive teaching: Differential effects of treatment. In H. C. Haywood & D. Tzuriel (Eds.), *Interactive assessment* (pp. 187–206). New York, NY: Springer-Verlag.
- Tzuriel, D., & Haywood, H. C. (1992). The development of interactive-dynamic approaches for assessment of learning potential. In H. C. Haywood & D. Tzuriel (Eds.), *Interactive assessment* (pp. 3–37). New York, NY: Springer-Verlag.
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press. (Original work published in 1934) doi:10.1037/11193-000
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Psychological Corporation.
- Wilkinson, G. S. (1993). *Wide Range Achievement Test 3*. Wilmington, DE: Wide Range.
- Zhu, J. (1999). *WASI manual*. San Antonio, TX: Psychological Corporation.

Received January 31, 2011
Revision received July 6, 2011
Accepted July 6, 2011 ■