SCREENING FOR MATHEMATICS DIFFICULTIES IN K–3 STUDENTS

CENTER ON INSTRUCTION

SCREENING FOR MATHEMATICS DIFFICULTIES IN K–3 STUDENTS

Russell Gersten
RG Research Group and
University of Oregon

Benjamin S. Clarke
RG Research Group and
Pacific Institutes for Research

Nancy C. Jordan
University of Delaware

2007

CENTER ON INSTRUCTION

This publication was created for the Center on Instruction by RG Research Group.

The Center on Instruction is operated by RMC Research Corporation in partnership with the Florida Center for Reading Research at Florida State University; RG Research Group; the Texas Institute for Measurement, Evaluation, and Statistics at the University of Houston; and the Vaughn Gross Center for Reading and Language Arts at the University of Texas at Austin.

The contents of this document were developed under cooperative agreement S283B050034 with the U.S. Department of Education. However, these contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

This research was supported in part by the National Center for Learning Disabilities.

Preferred citation:
Gersten, R., Clarke, B. S., Jordan, N. C. (2007). *Screening for mathematics difficulties in K-3 students*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

TABLE OF CONTENTS

INTRODUCTION

A major advance in the field of reading over the past 15 years has been the development and validation of screening measures that can detect with reasonable accuracy those kindergartners and first graders likely to experience difficulty in learning to read. Increasingly, these students now receive additional instructional support during the critical early years of schooling. This is especially important since we know that most students who are weak readers at the end of first grade remain struggling readers throughout the elementary grades (Juel, 1988).

We believe the field of mathematics is ready for a similar leap forward in identifying and providing intensive support for young students who are experiencing difficulties in learning the fundamentals of mathematics. The field now possesses adequate information on valid early screening measures that could be used by teachers and school personnel to identify these students. While currently we cannot assert that the same students who do poorly in mathematics at the end of first grade are likely to remain weak in mathematics for the rest of their academic careers (Hanich & Jordan, 2001), it makes sense to help those students who are still struggling with basic concepts and procedures in the early grades (Griffin, Case & Siegler, 1994).

There is some evidence that numerical concepts children acquire in early childhood lay the foundation for later acquisition of advanced mathematical concepts (Ginsburg & Allardice, 1984; Griffin et al., 1994), and that success or failure in acquiring early numerical concepts influences the interest and confidence students bring to new mathematical tasks and may fundamentally alter a student's success in mathematics throughout the elementary grades (Jordan, 1995). Thus, over the past 15 years, researchers have tried to assess the most salient aspects of a child's understanding of basic numerical relationships and operations and develop potential screening measures.

Key Variables

Screening tools to identify students at risk for later mathematics difficulties must address a number of critical variables, including predictive validity and content selection. For example, in designing early identification measures in mathematics, a critical variable is the extent to which performance on those measures relates to later mathematics performance. A student's score on a kindergarten screening measure would need to predict difficulty in mathematics at the end of first grade, second grade, and so on. Assessments that show evidence of predictive validity can aid in instructional decision-making. If evidence indicating a score below a certain threshold on a kindergarten or beginning of first-grade measure of mathematics predicts later failure, schools and teachers can use that information to allocate resources for instructional or intervention services to those students early on in their regular classroom setting. The belief is that early intervention—which might simply entail small group instruction that provides additional practice, explanation, and/or feedback—might be sufficient for young students who are behind their peers in acquiring critical foundational skills.

The design of screening instruments also needs to be guided by information from the fields of developmental and cognitive psychology on how children develop an emerging understanding of mathematics. Further, when developing assessment tools, mathematics educators' expertise should be tapped; this would result in screening tools that integrate knowledge bases in both fields. In the remainder of this report, we describe the aspects of numerical proficiency that seem to emerge consistently as most important to assess in young students and specify areas that seem most fruitful to assess in early screening batteries.

THE ROLE OF NUMBER SENSE IN MATHEMATICS DEVELOPMENT

We begin with a discussion of number sense, since this concept permeates the research on early development of numerical proficiency. Kalchman, Moss, and Case (2001) defined number sense:

> *The characteristics of good number sense include: a) fluency in estimating and judging magnitude, b) ability to recognize unreasonable results, c) flexibility when mentally computing, [and] d) ability to move among different representations and to use the most appropriate representation. (p. 2).*

However, as Case (1998) noted, "number sense is difficult to define but easy to recognize" (p. 1). In fact, precise definitions of number sense remain controversial and elusive. Berch (2005) captured the complexities of articulating a working definition of number sense in his article, appropriately titled *Making Sense of Number Sense: Implications for Children with Mathematical Disabilities*:

> *Possessing number sense ostensibly permits one to achieve everything from understanding the meaning of numbers to developing strategies for solving complex math problems; from making simple magnitude comparisons to inventing procedures for conducting numerical operations; and from recognizing gross numerical errors to using quantitative methods for communicating, processing, and interpreting information (p. 334).*

He compiled 30 possible components of number sense, based on research from cognitive psychology, developmental psychology, and educational research. One recurrent component in all operational definitions of number sense is magnitude comparison ability (i.e., the ability to discern quickly which number is the greatest in a set, and to be able to weigh relative differences in magnitude efficiently, e.g., to know that 11 is a bit bigger than 9, but 18 is a lot bigger than 9). Another frequently cited component of number sense is the ability to decompose numbers to solve a problem. For example, students with good number sense can solve 54 + 48 by first decomposing 48 to 4 tens and 8 ones. They can then add the 4 tens to 54 (64, 74, 84, 94) and then add 8 ones to 94. (NRC, 2001). (For a full list of possible components of number sense, see Berch (2005).)

Okamato and Case more formally, but also more forcefully, describe number sense as "the presence of powerful organizing schemata that we refer to as *central conceptual structures*. (p. 2)." They discuss how these conceptual structures are typically sets of mental number lines and how developing these mental number lines is critical to both proficiency in mathematical procedures and understanding of mathematical concepts. Both Berch (2005) and Griffin et al. (1994) also note that people with good number sense seem to develop a mental number line on which they can represent and manipulate numerical quantities. The development of a mental number line would facilitate the solving of a variety of mathematical problems.

Griffin et al. (1994) note that number sense is developed in large part from both formal and informal instruction by parents, siblings, or teachers, although genetic aspects are also clearly involved (Geary, 2004; Petrill, 2006).

Assessing Individual Components of Number Sense for Early Screening and Identification

Researchers in this field (e.g., Geary, 2004; Jordan, Kaplan, Olah, & Locuniak, 2006; Clarke & Shinn, 2004; Fuchs et al., 2006; Bryant, Bryant & Gersten, 2006 AERA presentation) have focused on a broad array of discrete aspects of numerical proficiency deemed critical for future success in mathematics. In some ways, this approach is similar to the approach used by reading researchers for assessing beginning reading using separate tests for discrete skills such as letter naming fluency, initial sound identification, phoneme segmentation, and the reading of short pseudowords. Most of these measures are fairly easy to administer and can be completed in a few minutes. Because the skill measures most researchers use in beginning mathematics are more focused, they are faster to administer and can more easily be used school- or district-wide with large numbers of students. Such measures could be used to quickly identify students whose mathematics achievement is either on track or at risk in one or more critical areas and prompt additional support. However, as with any screening measure, these brief measures cannot provide a full diagnostic profile.

In the next sections we describe the proficiencies deemed critical for future success in mathematics and provide information on some measures of these proficiencies. Table 2 lists the key proficiencies and results from selected measures of these proficiencies.

SELECTED COMPONENTS OF DEVELOPING NUMERICAL PROFICIENCY

Magnitude Comparison

As children develop a more sophisticated understanding of number and quantity, they are able to make increasingly complex judgments about magnitude. For example, one preschooler may know that 9 is bigger than 3, while another will know that 9 is 6 greater than 3. Riley, Greeno & Heller (1983, cited in NRC, 2001) found that, given a hypothetical scenario with a picture of five birds and one worm, most preschoolers can answer questions such as, "Suppose the birds all race over and each one tries to get a worm. Will every bird get a worm?" Their answers demonstrate a gross magnitude judgment that there are more birds than worms. But given a specific question about magnitude, for example "How many birds won't get a worm?" (p. 169), most preschoolers could not answer correctly.

The ability to make these finite types of magnitude comparisons is a critical underpinning of the ability to calculate. As the reader will note, many items in the Number Knowledge Test (Okamato & Case, 1996) involve magnitude comparison. In their view, magnitude comparison is at the heart of number sense.

Almost all researchers who develop assessments for early screening use some measure of magnitude comparison. This task requires some ability at mental calculations and also entails an understanding of place value. Using magnitude comparison in screening illustrates that screening instruments are not by nature designed to be comprehensive, and that a good screening instrument will be related to other critical aspects of performance. Thus, while we may not measure mental calculation and place value directly, a measure of magnitude comparison serves as an indicator of likely performance in those areas. Although magnitude comparison is rarely taught in traditional texts, research by Griffin, Case, & Siegler (1994) suggests that it is often taught (informally but explicitly) in middle-income homes and rarely taught in low-income homes. They found that high SES students entering kindergarten answered the magnitude comparison problems correctly 96% of the time, compared with low SES children, who answered correctly only 18% of the time.

Strategic Counting

The ability to understand how to count efficiently and to use counting to solve problems is fundamental to developing mathematical understanding and proficiency (Siegler & Robinson, 1982). Geary (2004) notes that using inefficient counting strategies is a key indicator of which young students are likely to have difficulty learning mathematics. Researchers typically differentiate between knowledge of counting principles and skill in counting. An example of rudimentary counting principle is the realization that "changing the order of counting, or the perceptual appearance of an array, will not affect the quantity, whereas addition and subtraction of an object will affect the quantity," (Dowker, 2005, p.85). A second example is the knowledge that, given a group of 5 objects and a group of 3 objects, in order to know how many objects you have all together, you can "count on" from 5 (i.e., count 6, 7, 8). This approach is certainly more efficient than one used by very young children: They tend to count out 3 objects, then 5 objects, and then put them together and begin counting all over again from 1.

In most cases, competence in counting is strongly related to knowledge of counting principles (Dowker, 2005). Siegler (1987, 1988) studied the evolution of the min strategy in young children in depth. For example, once a child possesses the min strategy, if asked "what is 9 more than 2," she will automatically know that it is much more efficient to reverse the problem to 2 more than 9, and simply "count on" from 9. Of course, grasping the min principle demonstrates a grasp of the commutative principle. Students with math difficulties or disabilities (MD) almost invariably use more immature and inefficient counting strategies to solve problems.

Although sequence counting (i.e., reciting the counting words without reference to objects) is an important skill for students to master in preschool, using strategic counting to solve mathematical problems is the more critical mathematical skill. For that reason, most researchers attempt to include a measure of strategic counting in their assessment batteries.

Geary (1990) examined the counting strategy use of first graders with MD compared with their peers. Although he found that both groups used similar strategies for solving problems, students with MD were three to four times more likely to make procedural errors. For example, when students with MD counted on their fingers they were incorrect half of the time, and when they used verbal counting strategies they were incorrect one third of the time. Thus some researchers also assess counting skill and accuracy, although it seems that the ability to count strategically and effectively is more likely foundational to future success in arithmetic. As students become more efficient in applying effective and efficient counting strategies to solve basic arithmetic combinations their conceptual understanding of important mathematical principles (e.g., commutativity and the associative law) is reinforced and strengthened.

Retrieval of Basic Arithmetic Facts

Early theoretical research on mathematics difficulties focused on *correlates* of students identified as having a mathematics learning disability. One consistent finding (Goldman, Pellegrino, & Mertz, 1988; Hasselbring, Sherwood, Bransford, Fleenor, Griffith, & Goin, 1988) was that students who struggled with mathematics in the elementary grades were unable to automatically retrieve addition and subtraction number combinations. More recently, Geary (2004) found that children with difficulties in mathematics typically fail to make the transformation from solving problems by counting on their fingers (or with objects) to solving problems in their heads without needing to use their fingers.

The research seems to indicate that, although students with MD often make good strides in terms of facility with algorithms, procedures, and simple word problems when provided with classroom instruction, deficits remain in their retrieval of basic combinations (Geary, 2004; Hanich & Jordan, 2001). These deficiencies suggest underlying problems with what Geary calls *semantic memory* (i.e., the ability to store and retrieve abstract information efficiently). This ability appears to be critical for students to succeed in mathematics and, ultimately, to understand mathematics.

Word Problems

Adding It Up, the report on mathematics developed by the National Research Panel (2001), concluded that although adults often think that children have a hard time solving word problems, children in fact find them easier than either simple number sentences or simple equations. Jordan, Levine, & Huttenlocher (1994) found that before formal instruction, young children solve simple word problems involving addition and subtraction more readily than they do number combinations, which do not refer to objects or provide any context. Word problems have only recently been added to early screening batteries.

Numeral Recognition: Learning to Link Numerals with Names

Numeral recognition is notoriously difficult in English compared with other languages; some researchers have suggested it may be one of many factors impeding how quickly American students learn mathematics.

While numeral recognition is not a mathematics skill in and of itself, it serves as a gateway skill to formal mathematics. In that regard, it is analogous to a child's ability to recognize letters as a means of accessing the written code. Just as letter-naming accuracy and speed are good predictors of a child's ability to benefit from typical reading instruction, numeral recognition may be a similar predictor in early screening of students for possible difficulties in mathematics. Thus while numeral recognition may not be a critical focus of mathematical instruction, it may serve as an indicator of risk for later failure in mathematics. Children begin to learn about the written symbol system for numerals before they enter school; thus a screening measure assessing numeral recognition could be a valuable tool to identify at-risk students as they enter kindergarten.

The numbers that children are most often first exposed to are descriptive, such as a home address or telephone number. This is in direct contrast to the use of numbers in formal school settings, where the cardinality of numbers and their use in abstract computations is emphasized. For example, figuring out how to solve a simple addition problem on a worksheet depends on a student's recognizing the number symbols and then using other facets of his mathematical understanding, including the concepts of cardinality, magnitude comparison, and counting.

EMPIRICAL STUDIES OF DISCRETE SKILL MEASURES: A BRIEF OVERVIEW OF PREDICTIVE VALIDITY

The following section highlights the major data sources used to prepare this report. Table 2 provides key details about each relevant study. This section is intended to offer context to the table and our recommendations in the final section.

Clarke (2004, 2005)

Clarke and colleagues (Clarke & Shinn, 2004; Chard, Clarke, Baker, Otterstedt, Braun, & Katz, 2005; Clark, Baker, Chard (in preparation)) used a set of individually administered timed measures each focusing on a specific component of number sense. The purpose of designing fluency measures was to create a set of feasible assessment tools for screening the entire population of a school's kindergarten or first grade students. Brief fluency measures enable schools to easily identify the kindergarten and first grade students most at risk in mathematics at the beginning of the school year and to provide interventions to prevent more serious mathematics problems in later grades.

Clarke and Shinn (2004) first tested three measures, number identification, quantity discrimination, and missing number, with first-grade students. Each measure was timed for one minute. The number identification measure required students to identify numerals between 1 and 20, the quantity discrimination measure required students to identify the bigger number from a pair of numbers between 1-20, and the missing number measure required students to identify a missing number from a sequence of 3 consecutive numbers in either the first, middle, or last position. The missing number measure functioned as a measure of strategic counting.

Predictive validity correlations for the fall to spring period ranged from .72 to .79; all were significant. Chard et al. (2005) and Chard et al. (in preparation) extended the initial work downward to a kindergarten sample and repeated it with a first grade sample. The kindergarten measures were modified to include only numbers between 1 and 10. The criterion measures in the spring were the NKT and the Stanford Achievement Test. Similar, but slightly weaker results were found with ranges of r=.46-.72 for kindergarten and r=.40-.59 for first grade.

Fuchs (2006)

Fuchs and colleagues (2006) investigated the utility of four group-administered measures from the start of first grade in predicting math disability at the end of second grade. They used both a measure of number combinations and numeral recognition/counting as well as two curriculum-based measures (CBM) that sampled skills covered in the first grade curriculum. One CBM probe sampled the concepts and applications taught in first grade (CBM-Concepts/Applications) and took 7 minutes 30 seconds to complete. The second CBM probe (CBM-Computation) took 2 minutes to administer and sampled the computation objectives for first grade. The number combination measure consisted of a one-minute addition and one-minute subtraction probe each containing 25 problems whose answers were between 0 to 12. Number identification/ counting had students fill in the last two blanks of a number sequence (e.g. 4, 5, 5, _, _). Students had two minutes to complete 8 problems. Logistic regression analyses were completed to gauge the battery's ability to correctly identify students with mathematics difficulties in arithmetic and word problems by the end of the second grade. Difficulty was defined as below the 10th percentile on a comprehensive test of mathematics. The measures linked to the curricular goals for first grade were more potent predictors, with CBM-Concepts/ Applications the better of the two. Number Identification/Counting was close to significance in predicting difficulty in arithmetic but not in word problems. The number combinations measure was not significant in predicting either arithmetic or word problem difficulties.

Mazzocco and Thompson (2005)

Mazzocco and Thompson (2005) tracked 226 students from kindergarten through third grade on several measures ranging from visual-spatial and cognitive to formal and informal mathematics achievement. Their goal was to determine the best measure or set of measures to predict which kindergarten students might be at risk for mathematics difficulty in the third grade. Running a set of regression models, the authors found that four specific items embedded in the kindergarten measures predicted later mathematic difficulty (as evidenced by standard scores of below the 10th percentile on a comprehensive measure of third grade mathematics). The four items were reading numerals, number constancy (when observing number sets below 6), magnitude judgments, and mental addition of one-digit numbers. The four-item model successfully classified 84% of third grade students, based on their kindergarten performance on the four items.

Jordan and Colleagues (2006)

Jordan, Kaplan, Olah, and Locunick (2006) developed a multi-component number sense battery (i.e., counting, number knowledge, nonverbal calculation, number combinations, and word problems). The correlation between performance on the number sense battery in the first month of kindergarten with math achievement in the last month of first grade was .67—virtually identical with the end-of-kindergarten assessment battery. The measure of simple word problem-solving served as a solid predictor.

Jordan's group has consistently studied the link between mathematics and reading disabilities, and has found that beginning reading skill (as well as overall IQ) strongly predicted later mathematics performance, and that the number sense battery added a significant proportion to the explained variance. That is, early number sense predicts later math achievement, over and above reading skill and general cognitive competencies.

VanDerHeyden and Colleagues (2001)

VanDerHeyden and colleagues (2001) created a series of one-minute, group-administered measures to assess kindergarten students' mathematical proficiency. The first measure had students count a number of circles and write the numeral corresponding to the number of circles they had counted; a modification of this measure had students count the number of circles and then circle the corresponding number from a set of choices. The last measure had students draw the number of circles that represented a numeral they were shown. Predictive validity was examined in terms of how well the measures predicted retention at the end of kindergarten. Scores predicted retention correctly in 71.4% (5/7) of cases and correctly predicted non-retention in 94.4% (17/18) of cases. (It should be noted that predicting retention was based on the three mathematics probes and three reading-readiness probes.) Concurrent validity correlations ranged from .44 to .61.

SUMMARY

Findings across these studies sketch an emerging picture of the critical aspects of measuring early numerical proficiency. First, many of the measures assess different discrete skills, with varying degrees of success. The fact that screening for different components of number sense can produce acceptable results further reinforces the understanding that numerical proficiency, even at the kindergarten and first grade level, is multi-faceted. Second, strategic counting and magnitude comparison emerged as two key variables among all the studies. Strategic counting and magnitude comparison provide students a foundation for more advanced mathematical thinking that they can use to solve problems (e.g. counting on from the larger addend to solve an addition problem). Lastly, the work of Fuchs et al. (in press) suggests that a broader measure of mathematics achievement reflecting the first grade curriculum may hold promise for predicting outcomes at later grades. The ability of different measures to predict both within and across grades will require additional research.

USE OF THE NUMBER KNOWLEDGE TEST (NKT) FOR EARLY SCREENING AND DIAGNOSTIC ASSESSMENT

The Number Knowledge Test (Okamato & Case, 1996), an individually administered measure (which takes about 10-15 minutes), attempts to assess students' procedural and conceptual knowledge related to whole numbers. The test examines students' understanding of magnitude, their counting ability, and their competence with basic arithmetic operations.

As the name implies, the NKT focuses exclusively on the domain of number, but unlike measures which assess discrete skills and abilities in numerical proficiency, the NKT assesses multiple facets of a student's numerical proficiency, including the application of number to basic arithmetic concepts and operations. The NKT contains four levels of increasing difficulty, each providing a deeper analysis. For example, the NKT includes problems to assess a child's ability to make magnitude comparisons; they increase in complexity as the child advances from a lower to a higher level. The magnitude comparison questions explore a child's understanding of magnitude, the word "bigger," and whether a child understands that traditional counting goes from smaller to larger numbers. Table 1 presents sample items from the Test of Number Knowledge.

The goal of this assessment is to inform kindergarten and first grade teachers of any gaps in students' knowledge of fundamental concepts involving quantity and number that are essential for learning school mathematics. It is also intended to help teachers differentiate instruction and in some cases offer more intensive intervention to students who lack the foundational knowledge to understand and master the primary grade curriculum. Our preliminary research with this measure suggests that its results can demonstrate to kindergarten and first grade teachers that some of their students enter school without even basic knowledge of counting and quantitative relationships. Failure to teach students these concepts and procedures in the early grades would tend to leave students with, at best, a superficial grasp of arithmetic, and, at worst, consistent failure.

Research Supporting Use of the Number Knowledge Test

We have explored the predictive validity of the Number Knowledge Test as an early screening and diagnostic measure. When Baker, Gersten, Flojo, Katz, Chard, & Clarke, 2002; Gersten, Jordan & Flojo (2005) administered it in kindergarten to predict subsequent performance a year later, the NKT demonstrated significant predictive validity correlations of .73 to the SAT-9 (Harcourt Educational Measurement, 2001) Total Mathematics score administered to students one year later, at the end of first grade. The NKT strongly predicted performance on both the Procedures subtest ($r=.64$), and the Problem Solving subtest ($r=.69$).

We also explored whether any other measure might add significantly more predictive validity. Because of a limited sample size, only one other measure emerged, a digit span measure (Geary, 2003), which is similar to the item in a typical IQ test. The digit span measure is hypothesized to function as a measure of working memory. Many researchers (Geary & Brown, 1991; Siegel & Ryan, 1988; Swanson & Beebe-Frankenberger, 2004) have identified working memory as related to mathematics disability. The numerical digit span measure did, in fact, add significantly to the battery's predictive power. The two measures accounted for almost 55% of the variance on the Spring of first grade SAT-9 total math score administered one year later ($R=.74$; $F(2,64) = 38.50$, $p<.01$).

Finally, we used item response theory (IRT) to establish the internal consistency reliability of the NKT and to examine the extent to which the four levels of the NKT developed by Okamato and Case (1996) were empirically supported. The analyses indicate reliability of 0.94. One of the most interesting findings of the IRT analyses (Baker, Yovanoff, & Gersten, 2002) was that the measure did not provide adequate numbers of items for kindergartners and first graders with low proficiency. This lack of lower-level items could complicate screening for at-risk status if large numbers of students would have scores of zero or one. This can be easily addressed by adding more items at the lower levels.

In summary, the NKT appears to be a strong measure, giving teachers an overview of students' knowledge of the key foundational components of numerical proficiency with strong psychometric properties, solid predictive validity, and diagnostic utility. (Diagnostic here means information concerning the student's strengths and weaknesses, not his or her classification in a diagnostic category.) The NKT takes time to administer, precluding its use as a brief screening measure for all students in a building. We envision its use only with students whom the more brief assessments described above indicated as being at risk, to help select appropriate instructional targets for intervention.

CONCLUSION

Though the research into early mathematics assessment is in its infancy, an emerging knowledge base is permitting us to draw important conclusions that can help guide further research and practice in the field. First and foremost is the recurring finding across multiple studies that significant differences exist between students in kindergarten and first grade and, more important, that we can pinpoint those differences accurately with brief and relatively easy-to-use screening tools. While differences observed in young children may arise from exposure to mathematics before formal schooling or student performance on more formal mathematics once in school, each component of number sense measured offers a critical link to instruction and additional instructional services.

The mathematics curriculum continues to change over the years, and it is possible that certain students may initially learn math at acceptable levels, only to experience problems once content moves to a more abstract level (e.g., with the introduction of decimals, improper fractions, ratios and proportions, negative numbers). Therefore, as in the reading field (Scarborough, 2001), we will likely see some students whose mathematics performance may be acceptable in the primary grades but deteriorates in later grades (Geary, 1993).

As the field of early screening advances, and the field of instructional research in mathematics evolves, we can develop understanding with greater precision. The research we reviewed addresses early predictors of students' ability to gain competence in understanding concepts and procedures related to whole numbers. It does not necessarily help us understand which students will ultimately succeed in the early elementary grades, but struggle when more intricate and abstract topics, such as rational numbers (i.e., fractions, ratio, proportion) or geometry, enter the picture beginning in fourth and fifth grade.

All we know from extant research is that visuo-spatial awareness and ability (Geary, 1993), and facility with measurement begin to play increasingly critical roles in students' acquisition of mathematical proficiency. Geary (2004) hypothesizes that students with problems in the visuo-spatial area may experience difficulty in learning concepts in geometry and measurement that may not emerge until the middle school years. To date, there are no data to either support or refute his hypothesis.

In addition, the research we reviewed supports the importance of working memory (Swanson & Beebe-Frankenberger, 2004; Geary, 2004) in understanding mathematical proficiency at many different levels, although few have explored instructional methods for enhancing working memory in mathematics. As our understanding of mathematical development advances, so too should our design of screening instruments that reflect the complexity of mathematics. It is likely that as outcome measures become more mathematically sophisticated, following guidelines in recent documents such as *Adding it Up* (NRC, 2001), and the revised NCTM (2000) Standards and Curriculum Focal Points (NCTM, 2006), we will learn more about longer-term predictors of subsequent success.

However, we now have sound means for assessing which five- and six-year olds are likely to encounter serious difficulties in learning mathematics. Each research effort reviewed has attempted to assess some aspect of predictive validity across either one or several school years, either by examining correlations over time or, less frequently, by examining student classifications. The strength of predicting later math difficulties varies, but recent research demonstrates that to some extent earlier difficulty in mathematics may underpin struggles with later mathematical achievement.

Although success (or failure) in mathematics at one grade level does not guarantee success (or failure) in later grades, there is good reason to identify which students will likely need help in learning both basic mathematics concepts and key procedures. A small body of research (Griffin et al. 1994; Fuchs & Karns, 2001) suggests that early intervention in kindergarten can produce some lasting benefits. As more research focuses on interventions for students identified as at risk, our understanding of the relationship between deficits in foundational skills and later performance will be enriched.

We are beginning to understand more about what comprises a comprehensive battery for early assessment, but are less sure of the components of an efficient assessment battery. In part, decisions about what works best may be guided by a max-min standard. That is, how can we gain the maximum amount of information in the minimum amount of time? Brief measures of magnitude comparison and strategic counting would appear to be important elements. Measures of working memory may well add to a battery's predictive power, but working memory is less likely to be a focus of instruction and therefore less likely to be sensitive to change in the way that the first two measures would. Further, although measures of counting knowledge and number identification may not be ideal as screening measures, they do hold potential for progress monitoring (Chard et al., 2005).

An important issue for future research is to determine the advantages and disadvantages of timed measures. Our sense is that timed measures may, in many instances, be more potent screening measures than un-timed ones. It may also be true that the realities of screening of all kindergarten or first grade students require timed measures to ensure efficient data collection. The max-min principal is again relevant to any discussions of the appropriateness of assessments as screening instruments.

Lastly, while the link between assessment and instruction in early mathematics is neither fully known or articulated, the field should remain cognizant of continued efforts to develop tools compatible with the principles of responsiveness to intervention (RTI) as defined in the reauthorization of the Individuals with Disabilities Education Act (2004). That such features (Fuchs, Fuchs, & Prentice, 2004) are often present in screening tools does not automatically ensure that they will be useful in the problem-solving and formative assessment phases of RTI. Future research efforts should focus on examining the role of effective assessments tools in decision-making criteria for RTI.

Despite the dearth of research in early mathematics, in recent years strides have been made to begin to build research programs that explore critical questions in early mathematics assessment and instruction. Our hope is that this paper helps synthesize the state of this research and spurs further research in the field.

REFERENCES

Baker, S., Gersten, R., Flojo, J., Katz, R., Chard, D., & Clarke, B. (2002). *Preventing mathematics difficulties in young children: Focus on effective screening of early number sense delays.* (Technical Rep. No. 0305). Eugene, OR: Pacific Institutes for Research.

Baker, S., Yovanoff, P., Chard, D., Gersten, R. (2002). *Item response theory analysis of Number Knowledge Test.* Unpublished manuscript. Eugene, OR: Pacific Institutes for Research.

Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38,* 333-339.

Bryant, D. P., Bryant, B., Gersten, R., & Kim, S. A. (2006, April). *Technical adequacy of early mathematics number and operation measures for k-2nd grade students.* Paper presented at the American Educational Research Association, San Francisco, CA.

Case, L. P., Harris, K. R., & Graham, S. (1992). Improving the mathematical problem-solving skills of students with learning disabilities: Self-regulated strategy development. *Journal of Special Education, 26,* 1-19.

Case, R. (1998). *A psychological model of number sense and its development.* Annual meeting of the American Educational Research Association, San Diego, CA.

Chard, D., Clarke, B., Baker, S. K., 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. K., & Chard, D. J. (in preparation). Developing and validating measures of number sense to identify students at-risk for mathematical disabilities.

Clarke, B., & Shinn, M. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33,* 234-248.

Dowker, A. (2005). *Individual differences in arithmetic: Implications for psychology, neuroscience, and education.* New York: Psychology Press.

Fuchs, L. S., Fuchs, D., & Prentice, K. (2004), Responsiveness to mathematical problem-solving instruction: Comparing students at risk of mathematics disability with and without risk of reading disability. *Journal of Learning Disabilities, 37,* 292-306.

Fuchs, D., & Karns, K. (2001). Enhancing kindergartners' mathematical development: Effects of peer-assisted learning strategies. *Elementary School Journal, 101,* 495-510.

Fuchs, L. S., Fuchs, D., Compton, D., Bryant, J. D., Hamlett, C. L., & Seethaler, P. M. (in press). Screening at the beginning of first grade for math disability at the end of second grade. *Exceptional Children.*

Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology, 33,* 386-404.

Geary, D. C. (1993). Mathematical disabilities: Cognition, neuropsychological and genetic components. *Psychological Bulletin, 114,* 345-362.

Geary, D. C. (2003). Learning disabilities in arithmetic: Problem solving differences and cognitive deficits. In K. Harris & S. Graham (Eds.), *Handbook of Learning Disabilities* (pp. 199-212). New York, NY: Guilford.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37,* 4-15.

Geary, D. C., & Brown, S. C. (1991). Cognitive addition: Strategy choice and speed-of- processing differences in gifted, normal, and mathematically disabled children. *Developmental Psychology, 27,* 787-797.

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

Ginsburg, H. P., & Allardice, B. S. (1984). Children's difficulties with school mathematics. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social contexts* (pp. 194-219). Cambridge, MA: Harvard University Press.

Goldman, S. R., Pellegrino, J. W., & Mertz, D. L. (1988). Extended practice of basic addition facts: Strategy changes in learning disabled students. *Cognition and Instruction, 5,* 223-265.

Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning or arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice,* (pp. 24-49). Cambridge, MA: MIT Press.

Hanich, L., & Jordan, N. (2001). Performance across different areas of mathematical cognition in children with learning disabilities. *Journal of Educational Psychology, 93,* 615-626.

Hasselbring, T., Sherwood, R., Bransford, J., Fleenor, K., Griffith, D., & Goin, L. (1988). Evaluation of a level-one instructional videodisc program. *Journal of Educational Technology Systems, 16,* 151-169.

Jordan, N. C. (1995). Clinical assessment of early mathematics disabilities: Adding up the research findings. *Learning Disabilities Research & Practice, 10,* 59-69.

Jordan, N., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77,* 153-175.

Jordan, N., Levine, S., & Huttenlocher, J. (1994). Development of calculation abilities in middle and low income children after formal instruction in school. *Journal of Applied Developmental Psychology, 15,* 223-240.

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.

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.* Mahwah, NJ: Lawrence Erlbaum.

Kaminiski, R. & Good, R. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25* (2), 215-227

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

National Research Panel (2001). Eager to learn: Educating our preschoolers. M. S. Donovan & M. S. Burns (Eds.). Washington, D. C.: National Academy Press.

Okamato, 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,* 27-59.

Petrill, S. (February 2006). *Mathematics disability: Definition, etiology, and comorbidity.* Paper presented at the 14th Annual PCRC (Pacific Coast Research Conference), Coronado, CA.

Riley, M. S., Greeno, J. G., & Heller, J. H. (1983).Development of children's problem solving ability in arithmetic. In H. P.Ginsburg (Ed.), *The development of mathematical thinking* (pp.109–151). New York: Academic Press.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of Early Literacy Research* (pp. 97-110). New York, NY: Guilford.

Siegel, L., & Ryan, E. (1988). Development of grammatical-sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Developmental Psychology, 24,* 28-37.

Siegler, R. (1987). The perils of averaging data over strategies: The example of children's addition. *Journal of Experimental Psychology: General, 116,* 260-264.

Siegler, R. (1988). Individual differences in strategy choices: good students, not-so-good students, and perfectionists. *Child Development, 59,* 833-851.

Siegler, R. S., & Robinson, M. (1982). The development of numerical understandings. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in Child Development and Behavior* (pp. 241-311). New York, NY: Academic Press.

Swanson, H. L., & Beebe-Frankenberger, M. E. (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.

VenDerHeyden, A. M., Witt, J. C., Naquin, G. & Noell, G. (2001). The reliability and validity of curriculum-based measurement readiness probes for kindergarten students. *School Psychology Review, 30*, 363-382.

Table 1. Example Items from the Number Knowledge Test

Number Knowledge Test	Example Items
Level 0	1. Here are some circles and triangles. Count just the triangles and tell me how many there are.
Level 1	1. If you had 4 chocolates and someone gave you 3 more, how many chocolates would you have? 2. Which is bigger: 5 or 4?
Level 2	1. Which is bigger, 19 or 21? 2. What number comes 4 numbers before 17?
Level 3	1. What number comes 9 numbers after 999? 2. Which difference is smaller: the difference between 48 and 36 or the difference between 84 and 73?

Table 2. Summary of Validity Evidence for Select Early Screening Measures

Construct	Author	Measure	Description			
			N	Time Frame	Criterion	Notes
Strategic Counting	Clarke	Missing Number	52	Fall to Spring 1st	WJ-AP M-CBM	Timed one minute measure r=.67-.72
	Clarke	Missing Number	95	Winter to Spring K	SE-SAT 2 NKT	Timed one minute measure r=.46-.63
	Clarke	Missing Number	123	Fall to Spring 1st	SAT-10 NKT	Timed one minute measure r=.44-.59
	Jordan	Counting Knowledge	277	Fall K to Spring 1 Spring K to Spring 1	WJ-3 WJ-3	r=.36 r=.35
	Geary	Strategic Counting (Accuracy and Retrieval)	42	K Concurrent	WRAT	r=.71
Magnitude Comparison	Clarke	Quantity Discrimi-nation	52	Fall to Spring 1st	WJ-AP M-CBM	Timed one minute measure r=.70-.79
	Clarke		95	Winter to Spring K	SE-SAT 2 NKT	Timed one minute measure r=.64-.72
	Clarke		123	Winter to Spring 1st	SAT-10 NKT	Timed one minute measure r=.40-.59
	Jordan	Number knowledge	277	Fall K to Spring 1 Spring K to Spring 1	WJ-3 WJ-3	r=.54 r=.59
Word Problems	Jordan		277	Fall K to Spring 1 Spring K to Spring 1	WJ-3 WJ-3	r=.47 r=.62

(continued)

Table 2. Summary of Validity Evidence for Select Early Screening Measures (cont.)

Construct	Author	Measure	Description			
			N	Time Frame	Criterion	Notes
Arithmetic Combinations	Fuchs	CBM Computation	170	Fall 1 to Spring 2	WRAT	r=.34
	Jordan	Nonverbal calculation	277	Fall K to Spring 1 Fall K to Spring 1 Spring K to Spring 1	WJ-3 WJ-3 WJ-3	r=.52 r =.5 r=.58
	Jordan	Number Combinations	277	Fall K to Spring 1 Spring K to Spring 1	WJ-3 WJ-3	r=.58 r=.64
Numeral Recognition	Clarke	Number Identification	52	Fall to Spring 1	WJ-AP M-CBM	Timed one minute measure r=.67-.72
	Clarke		95	Winter to Spring K	SE-SAT 2 NKT	Timed one minute measure r=.62-.65
	Clarke		123	Winter to Spring 1	SAT-10 NKT	Timed one minute measure r=.40-.55
Compre-hensive	Jordan		277	Fall K to Spring 1 Spring K to Spring 1	WJ-3 WJ-3	r=.70 r=.73
	Mazzacco	Items from screening battery	226	K to Grade 3		ROC value .88 (items were reading numerals, number constancy, magnitude judgments, mental addition)
	Gersten et al.	Number Knowledge Test	65	Spring K to Spring 1st	SAT-9	r=.72

CENTER ON
INSTRUCTION