

Math Problem-Solving and Cognition among Emerging Bilingual Children at Risk

And Not at Risk for Math Difficulties

H. Lee Swanson

Jennifer Kong

Stefania D. Petcu

University of New Mexico

H. Lee Swanson, Ph.D., Educational Psychology, University of California-Riverside (Lee.Swanson@ucr.edu) and University of New Mexico (HLswanson@unm.edu). Jennifer Kong, Ph.D., Special Education, University of California-Riverside and post-doctoral University of New Mexico (JenniferKong@unm.edu) and Stefania D. Petcu, Ph.D., Office of Research and Community Engagement, University of New Mexico (spetcu@unm.edu) .

This research was support by an NSF grant, Division of Research on Learning (award number 1660828) Sciences awarded to the first author. This study does not necessarily reflect the views of NSF, or the participating school districts. Special appreciation is given to Monica Ascencio Pimentel, Mario Del Angel Guevara, Lorena Luevano, Julianna Massa, Catherine Riskie, Yuliana Kenfield, Tara Curtis, Christy Yan, Karen Flores, Sindy Zambrano, Karen Cruz, and Bernadette Hall-Cuaron and School District Liaisons and Consultant: Rosie Gonzales and Erin Bostick Mason for data collection and/or analysis.

Correspondence concerning this article should be addressed to H. Lee Swanson, Educational Psychology, College of Education, University of New Mexico, and Albuquerque, NM 87131.

E-mail:HLswanson@unm.edu or Lee.Swanson@ucr.edu

Date revised submitted: September 24, 2019

Abstract

Cognitive processes that underlie individual differences in mathematical problem-solution accuracy in elementary emerging bilingual children (English Learners) at risk and not at risk for math problem-solving difficulties (MD) were examined. A battery of tests was administered in both English and Spanish that assessed problem-solving, achievement, and cognitive processing in children in first ($N = 155$ / MD $N=23$), second ($N = 129$ /MD $N=44$) and third grades ($N = 110$ / MD $N=39$). The results were that (a) the executive component of working memory (WM) predicted MD status independent of measures of fluid intelligence, reading, calculation, knowledge of algorithms, processing speed, short-term memory, and inhibition, (b) low performance on Spanish measures of numeracy and executive component of WM were major predictors of the odds of being classified as MD and (c) bilingual proficiency primarily moderated English rather than Spanish measures of cognition. The results support the notion that the executive system of WM is an important predictor of emerging bilingual children's math problem-solving difficulties.

Key words: Executive processing, working memory, math word problems, English learners, cross-sectional

Math Problem-solving and Cognition among Emerging Bilingual Children at Risk
And Not at Risk for Math Difficulties

Although current categories of learning disabilities include specific disabilities in calculation and mathematical problem-solving [see IDEA reauthorization, 2004, Sec. 300.8(c) (10)], the majority of the research on math disabilities (MD) focuses on cognitive processes related to calculation (Geary, 2013; Murphy et al., 2007; Peng et al., 2018; Swanson & Jerman, 2006). This focus is limiting because mathematical problem-solving constitutes one of the most important mediums through which children can potentially learn to select and apply strategies necessary for coping with everyday problems. In addition, according to the National Mathematics Advisory Panel (2008), and to PISA (Programme for International Student Assessment; OCED 2012), U.S. children show substantial weaknesses when asked to solve math word problems relative to other achievement domains and in comparison to other industrialized countries, which suggests a need to understand the cognitive mechanisms and processes that underlie math word problem-solving. Difficulties in mathematical problem-solving are especially compounded in EL children learning English as a second language. Spanish-speaking ELs makes up a large percentage (73.1%) of the EL population in the U.S., and represent a substantial number of students who do not demonstrate proficiency in math (National Assessment of Educational Progress, 2011; 2017; National Center for Education Statistics, 2011). One of the key areas found particularly difficult for EL children whose first language is Spanish is solving word problems (e.g., Macizo, Herra, Roman, & Marin, 2011; Martiniello, 2008; 2009; Ockey, 2007).

The purpose of this study was to explore those cognitive processes that underlie math problem-solving difficulties in EL or emerging bilingual children.¹ Clearly, not all emerging

MATH PROBLEM-SOLVING AND COGNITION

bilingual children experience math problem-solving difficulties and therefore it is important to determine some of the cognitive strengths of these children compared to their lower-performing peers. Although some of the difficulties in math experienced by emerging bilingual children with Spanish as a first language have been partially attributed to cross language transfer, oral language, linguistic complexity, and reading skill (e.g., Farnia & Geva, 2011; Han, 2012; Macizo et al., 2011; Martiniello, 2008; 2009; Ockey, 2007; Vukovic & Lesaux, 2013), other processes besides language may play a critical role in such children's math difficulties (MD). Therefore, it is important to determine some of the cognitive measures that predict success on problem-solving measures for Spanish speakers so intervention programs can be developed and tested.

One process of particular importance in this study is the executive component of working memory (WM). This is because executive processing is viewed as a cognitive asset in bilingual children (e.g., Bialystok, 2011; Bialystok & Feng, 2009) and because the executive component of WM has been found to be a major cognitive component that underlies math problem-solving difficulties in monolingual children (e.g., David, 2012; Geary, 2012; Swanson & Beebe-Frankenberger, 2004). Prior to identifying some of these processes related to WM that may underlie MD, a review of those measures commonly attributed to math difficulties is considered. Some of the measures given priority are drawn from recent meta-analyses and syntheses of literature that highlight academic, language and cognitive performance differences comparing monolingual children with and without MD (e.g., Han, 2012; Peng, Namkung, Barnes, & Sun, 2016; Swanson & Jerman, 2006).

Vocabulary and Reading Proficiency

Performance on measures of vocabulary and reading has been clearly related to word problem-solving accuracy (e.g., Bjork & Bowyer-Crane, 2013; Farnia & Geva, 2011; Fuchs et

MATH PROBLEM-SOLVING AND COGNITION

al. 2006, 2016; Gorman, 2012; Harvey & Miller, 2017; Lee, Ng, Ng, & Lim, 2004; Swanson, Jerman, & Zheng, 2008). Several studies show that L1 performance in vocabulary (e.g., Farnia & Geva, 2011) and reading (e.g. Lonigan, Allan, Goodrich & Farrington & Philips, 2017; Swanson, Orosco, & Lussier, 2012) is related to reading comprehension, which in turn is highly correlated with math problem-solving accuracy (e.g., Fuchs, Fuchs, Compton, Hamlett & Wang, 2015). For example, previous studies have identified the instrumental role of language in mathematical understanding, reasoning, problem-solving and expression of solutions (Bjork & Bowyer-Crane, 2013; Vukovic & Lesaux, 2013). Likewise, there is a strong association between problem-solving and reading. Several studies find that MD co-occurs with reading disabilities more frequently than expected by chance (e.g., Branum-Martin, Fletcher, & Stuebing, 2013; Landerl & Moll, 2010).

Number Estimation and Magnitude Judgement

Children's mental representations of numbers also play a significant role in math problem-solving performance in children with MD. A core competency in math performance is children's ability to mentally map Arabic numerals onto representations of small or large quantities (e.g., Booth & Siegler, 2008; Kolkman, Kroesbergen, & Leseman, 2014). For example, magnitude judgment and number line estimation have been linked to math difficulties (e.g., Fuchs et al. 2012; Geary, 2011; Simmons et al. 2012). Various forms of mental number representations that may underlie math difficulties are judgments of estimation (e.g., Fuchs et al., 2012) and number magnitude (e.g., which is larger: 8 or 5; e.g., Rousselle & Noël, 2007). For example, a number-line estimation task requires children to estimate the position of target numbers on a line within numerals at endpoints (e.g., 0 and 100). Several studies (e.g., Fuchs et al., 2012; Geary, 2011; Rousselle & Noël, 2007) have found that accuracy in making placement on a mathematical

MATH PROBLEM-SOLVING AND COGNITION

number line was uniquely predictive of math skills. Other specific forms of knowledge that may play an important role in math difficulties include recognizing propositions within word problems. These specific propositions within word problems are related to accessing numerical, relational, question the problem is asking, and extraneous information, as well as accessing the appropriate operations and algorithms for the solution (e.g., Mayer & Hegarty, 1996).

Short-Term Memory Storage and Working Memory

In addition to language, reading and domain-specific processes, recent studies have linked short-term memory (STM) and working memory (WM) to emerging bilingual children at risk for MD (e.g., Swanson, Kong, & Petcu, 2018).² Although both STM and WM involve transient memory, their relationships to math performance may be different (e.g., Alloway, Gathercole, Willis, & Adams, 2004). Theoretically, STM involves passive storage, in which the learner mentally rehearses verbal and/or visual-spatial information (Baddeley, 2007). Although WM or complex span tasks share the same processes (e.g., rehearsal, updating, controlled search) as STM or simple span tasks, STM tasks have a greater reliance on phonological processes than complex span (WM) tasks (see Unsworth & Engle, 2007, pp. 1045-1046, for a review). Several studies have shown that phonological STM is an important component of second-language vocabulary acquisition and achievement (e.g., Engel & Gathercole, 2012; Thorn & Gathercole, 2001).

The majority of studies that have compared children who vary in language and math skills assume that STM measures tap a phonological system or what Baddeley (2012) denotes as the phonological loop. The phonological loop has been referred to as STM because it involves two major components discussed in the STM literature: a speech-based phonological input store and a rehearsal process (see Baddeley, 2007, for review). Several findings converge on the view

MATH PROBLEM-SOLVING AND COGNITION

that children with relatively poor phonological memory are less successful in second language acquisition which in turn influences L2 achievement (e.g., Palladino & Ferrari, 2008).

In contrast to phonological storage, additional studies highlight the importance of the executive component of WM. That is, difficulties in executive processing (controlled attention) may also account for the high incidence of math difficulties in emerging bilingual children. Two findings underlie this assumption: (1) the executive component of WM has been associated with math difficulties (e.g., David, 2012; Martin et al., 2013; Menon, 2016; Swanson & Beebe-Frankenberger, 2004) and (2) a number of studies have shown proficiency in L1 and L2 positively influence intentional control, cognitive flexibility and overall executive processing (e.g., Bialystok, 2011). A key mechanism in this process is the inhibition of competing for language systems (Bialystok, 2011; Bialystok & Feng, 2009; Bonifacci et al., 2011). Because inhibition has been attributed to WM (e.g., Engle, 2002; Friedman, Haberstick, Willcutt, Miyake, Young, & Hewitt, 2007), as well as math difficulties (e.g., Passolunghi & Siegel, 2001; Toll, Kroesbergen, Van Luit, & Johannes, 2011; see Censabella & Noel, 2008, for review), individual differences related to WM may play an important role in emerging bilingual children problem-solving performance.

Purposes of the Study

The purpose of this study is twofold. The first purpose was to determine those aforementioned measures (e.g., reading, vocabulary, estimation) that uniquely predict problem-solving accuracy differences among emerging bilinguals who are at risk and not at risk for math problem-solving difficulties. Of interest is whether measures of WM contribute unique variance in predicting MD status when variables related to reading, vocabulary and domain-specific processes (e.g., estimation, recognition of problem-solving components) are entered into the

regression model.

The second purpose was to determine if bilingual proficiency predicts risk for problem-solving difficulties. For lack of a better term, bilingual proficiency was defined in this study as the combination of both English and Spanish vocabulary skills (further discussed in the methods section). Although bilingual, emerging bilingual children vary in their mastery level of each language which in turn plays a role toward identifying the cognitive processes that influence achievement (math) (Cummins, 1979; Kempert et al, 2011; Lonigan et al., 2017; Rosselli, Aridila, Lalwani & Velez-Urbe, 2016). However, the majority of these studies have not specified the role that bilingual proficiency plays in the probability of being identified at risk for MD. Therefore, an objective of this study was to investigate the influence of bilingual proficiency on the relationship between cognition (e.g., STM, WM) and difficulties in math problem-solving in bilingual children within both language systems (i.e., Spanish and English). In summary, the two questions that directed this study were:

1. Do measures of WM uniquely predict MD status when measures of vocabulary, reading and domain-specific measures are entered into the regression model?
2. Does bilingual proficiency influence the role of cognitive measures in predictions of MD status?

Method

Participants

Three hundred and ninety-four ($N=394$) students in grades 1 ($n=155$), 2 ($n=129$) and 3 ($n=110$) from two large school districts in the southwest United States participated in this study. Children were drawn from the first year of a federally funded longitudinal study focused on the

MATH PROBLEM-SOLVING AND COGNITION

cognitive development of English language learners (Swanson, Kong, & Petcu, 2019). The children were designated as EL or emerging bilingual by their school and were selected from 30 classrooms. These children were selected from urban schools with a high poverty representation (over 98 percent of the children participated in a full or reduced Federal lunch program) as well as a high Hispanic representation (> 95 %). The final sample included 192 boys and 202 girls who returned signed consent forms. School records indicated children's primary home spoken language was Spanish (80%). All children were selected from dual language classrooms in which math instruction was in both English and Spanish. The math programs in each school district placed a heavy emphasis on math facts and word identification. No significant differences in gender representation emerged across the grades, $\chi^2(df=2, N=394)=2.88, p=.23$.

Math difficulties. Performance at or below the 25th percentile on a normed referenced math measure is commonly used to designate children at risk for MD, and therefore, it is useful to use cut-off scores as used by other researchers as well as what is practiced in the schools (Fuchs et al., 2005; Geary, 2011; Geary, Hoard, Nugent, 2012). This 25th percentile cut-off point has been shown to yield a discrete subgroup of children at risk for MD (Swanson, Olide, & Kong, 2018). The present study administered four (2 English and 2 Spanish) normed referenced measures to tap math problem-solving performance. Those participants with a median standard score at or below a standard score of 90 (25th percentile) across English and Spanish problem-solving measures were regarded “as at risk for MD” ($N=108$) and those above a median of 90 ($N=286$) were considered as “not at risk” for MD. No group differences emerged related to gender, $\chi^2(df=1, N=394)=2.88, p=.09$, but significant group differences emerged related to grade level, $\chi^2(df=2, N=394)=16.40, p=.0003$. The incidence of children at risk for MD was

MATH PROBLEM-SOLVING AND COGNITION

more frequent in the older grades than younger grades, and therefore grade-level was a covariate in the subsequent analyses.

The subsequent analyses also determined the role of bilingual proficiency in predictions of MD status. As shown in Table 1, our sampling of emerging bilingual children yielded clear variations in English and Spanish on normed referenced vocabulary measures suggesting potential variations in the balancing of the two language systems. The literature varies in methodology for operationalizing variations in bilingualism (e.g., balanced vs. unbalanced bilinguals; Peña, Bedore, & Kester, 2016; Rosselli, Lalwani, & Vélez-Urbe, 2016), with some studies suggesting a focus on proficiency within each language system that in turn yields a total score and/or focusing on conceptual proficiency (tests that allow for responses in either language that yields the highest score in the preferred language, see Peña et al., 2016; for a comprehensive review). For the present study, similar to studies measuring balanced and unbalanced bilinguals (e.g., Rosselli et al., 2016), bilingual proficiency was measured as the sum of the four norm-referenced language measures (English and Spanish receptive and expressive language). Thus, we used a continuous variable to reflect overall language proficiency.

Measures

The study included group and individual administrations of a battery of tests. The series of tests were counterbalanced into one of four presentation orders. No Spanish and English versions of the same test (except for the Expressive One-word Picture Vocabulary Test, Spanish-Bilingual Edition; Brownell, 2001) were presented simultaneously. Several measures that required Spanish-translated versions were developed from an earlier study (see Swanson, Saez, Gerber, & Leafstedt, 2004; 2014, 2017, for further discussion). All participants were administered both English and Spanish versions of each measure by bilingual graduate students

MATH PROBLEM-SOLVING AND COGNITION

and staff researchers. Instructions were given in Spanish for all tasks requiring Spanish responses unless noted otherwise. There were some tasks (e.g., pseudoword reading) that required calibration for task difficulty. Three native Spanish speakers made judgments on the difficulty of the items in relation to the task presented in English. Interrater agreements exceeded 85%. The mean raw scores and the sample reliabilities for all measures described below are provided in Appendix A.

Criterion Measures

Applied problem-solving. The Applied Math Problem-solving subtest from the *Woodcock-Johnson III* (Woodcock, Grew, & Math, 2001) was administered for the English presentation and the *Problemas Aplicados* from the *Batería III Woodcock-Muñoz* (Muñoz-Sandoval Woodcock, McGrew, & Mather, 2005) was administered to establish normed referenced math levels in Spanish. Both of these subtests are individually administered and assess children's early mathematical operations (e.g., counting, addition, and subtraction) through practical problems. In order to solve each problem, the subject must listen to the formulation, recognize the procedures that must be followed, and then perform relatively simple calculations.

Mental computation of word problems. This individually administered task was taken from the English and Spanish arithmetic subtest of the *Wechsler Intelligence Scale for Children, Third Edition* (WISC-III; Psychological Corporation, 1991). The task was selected since it has been found predictive of MD (e.g., Swanson et al., 2008). For this task, each word problem was orally presented and was solved without paper or pencil. Questions ranged from simple addition (if I cut a banana in half, how many pieces will I have?) to more complex calculations (e.g., if three children buy tickets to the show for \$6.00 each, how much change do they get back from

MATH PROBLEM-SOLVING AND COGNITION

\$20.00?). This subtest requires children to answer simple to complex problems involving arithmetic concepts presented in a sentence context. The dependent measure was the number of problems correct. Although this task makes demands on memory, this task correlates more highly with the information and similarities subtest on the WISC-III than the other subtests (e.g., digit span). In fact, it has a *g* loading the same as vocabulary, information, and similarities, whereas its loading with digit span is substantially lower (see Keith et al., 2006, p. 122).

Measures Assumed Related to Math Status

Measures of Language, Reading and Nonverbal Reasoning

Woodcock-Muñoz Language Survey-Revised (WMLS-R). This test established a norm-referenced reading level in English and Spanish (Woodcock-Muñoz –Sandoval & Alverado, 2005). The WMLS-R Spanish and English Word Identification and Passage Comprehension subtests were administered. The subtests were administered individually to students in both English and Spanish. For the Word Identification Subtest, children were tested individually by presenting them with a list of words, which gradually increased in difficulty. The words followed regular spelling patterns in this non-timed test. For the passage comprehension subtest, children identified specific words that go in the blank spaces of various passages. Early passages were accompanied by a corresponding picture, and sentences gradually increased in complexity. The WMLS-R word identification and passage comprehension subtests yield a raw score that is converted to a standard score ($M = 100$, $SD = 15$).

Pseudoword reading task. The word attack subtest of the Woodcock Reading Mastery Test (Woodcock, 1998) was administered. The test was administered in English according to the standardized instructions. The measure required the child to orally read the list of pseudowords arranged in increasingly difficult order. The Woodcock technical manual reports internal

MATH PROBLEM-SOLVING AND COGNITION

reliability of the subtest at .88. A Spanish version of the task was also administered using the same rules. The Spanish version was developed using specific letter rules for each English item to ensure effective translation; for example "ift" was translated to "iyo".

Peabody Picture Vocabulary Test (PPVT) and Test de Vocabulario en Imagenes (TVIP). The Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) was administered in English. In this task, children were presented with four pictures and asked to select the picture that matched the word read aloud in English. Word presentation gradually increased in difficulty. The technical manual states parallel form reliability of .91. The TVIP measure is similar to the PPVT in the presentation and administration, except that words are read aloud in Spanish (Dunn, Lugo, Padilla & Dunn, 1986). The split-half reliability presented in the technical manual was .91-.94.

Expressive One-Word Picture Vocabulary Test -Spanish-Bilingual Edition. The Expressive One-Word Picture Vocabulary Test -Spanish-Bilingual Edition (EOWPVT-SBE: Brownell, 2001) was used as a measure of syntax and assesses English and Spanish speaking vocabulary. Items were administered in both languages, with the first language chosen for administration determining the order, until the child achieved ceiling. If the child achieved a ceiling in English before Spanish, Spanish alone was continued and vice versa. The technical manual reported reliability for the standardization sample at .95.

Fluid (nonverbal) intelligence. The Colored Progressive Matrices (Raven, 1976) was used as an indicator of nonverbal or Fluid intelligence. Children were given a booklet with patterns on each page, each pattern revealed a missing piece. For each pattern, six possible replacement pattern pieces were displayed. The dependent measure (range 0 to 36) was the number of problems solved correctly. The technical manual reports internal consistency

reliability ranging from .80 to .90.

Measures of Domain-specific Math Processes

Arithmetic calculation. The calculation subtests from the *Woodcock-Johnson III* (Woodcock, Grew, & Math, 2001) were administered for the English presentation and Cálculo from the *Batería III Woodcock-Muñoz* (Muñoz-Sandoval Woodcock, McGrew, & Mather, 2005) were administered to establish normed referenced math levels in Spanish. The Arithmetic subtest from the Wide Range Achievement Test (WRAT-3; Wilkinson, 2003) was also administered to assess arithmetic computational skills. The subtest required written computation of problems that increased in difficulty. Problems ranged from simple calculations (e.g., $2 + 2 = ?$) to algebra. The dependent measure was the number of problems solved correctly, which yielded a standard score ($M = 100$, $SD = 15$).

Word problem-solving components. The experimental measure assessed the child's ability to identify the components of word problems (Swanson & Beebe-Frankenberger, 2004). Each booklet contained three problems that included pages assessing the recall of text from the word problems. To control for reading problems, the examiner orally read each problem and all multiple-choice response options as the students followed along. After the problem was read, students were instructed to turn to the next page on which they were asked a series of multiple-choice questions requiring them to identify the correct propositions related to (1) question (2) number, (3) goal, (4) operation and (5) algorithm of each story problem. Children were also to identify the extraneous propositions for each story problem.

Estimation. Two number line estimation tasks adapted from Siegler and Opfer's (2003) and Siegler and Booth's (2004) study, were administered. For set 1 of the Estimation task, children were asked to examine five straight lines that were 25-cm long. Each line was identical

MATH PROBLEM-SOLVING AND COGNITION

in length and was marked with a zero at one end and one hundred on the other end, creating a blank number line. A single number (e.g., 50, 75, 45, 32, 6, 22) was placed above the center of each line. Children were asked to estimate where they thought the number presented should be placed on the line and indicated this by marking an X on the line. For set 2, children were asked to examine another set of five straight lines. For this set, however, each line was of a different length (25cm, 20cm, 12cm, 30cm, and 20cm) with endpoints of 0 and 100. The reason to manipulate the length of the line was related to issues raised as to whether spatial information or magnitude judgment underlines problem in estimation (Chew, Forte et al., 2016).

For each of the 10 lines (set 1 and 2), the point of accuracy was calibrated for each line. Accuracy was calculated by using a transparency template and counting how many units of measure the X was from the correct answer. For the five lines in set 1, the distance from the accuracy point was computed for each $\frac{1}{4}$ inch. For set 2, arithmetically equivalent distances were used to count off the distance between the participant's X and the where actual placement the correct answer should be on the line. The difference score (number of units from the exact point) was computed for each set. Thus, our estimate of the number line estimation varied from that of Siegler and Opfer (2003), in that they used group level median placements fitted to linear analog models to make inferences about the children's placements.

Magnitude comparisons. Two sets of digits were presented in 25 rows with three columns. Each row had the same number of digits (1 digit, 2 digits, and 3 digits) in each column. In the first set, children were asked to circle the largest number in each row as fast as they could in 30 seconds. The second set also had an additional 25 rows of numbers with three numbers in each row. Children were asked to circle the smallest number in each group as fast as they could in 30 seconds. The numerical distance between a symbolic magnitude comparison

MATH PROBLEM-SOLVING AND COGNITION

was alternated across rows so that each row had one comparison close in the numerical distance (e.g., 2 and 3) and one far in the numerical distance (2 and 9). Children were presented with 25 rows of numbers with three numbers (either in pairs or in three digits) in each row. The scores for set 1 were the number of correctly identified largest numbers (set 1) within 30 seconds, and the scores for set 2 were the smallest numbers correctly identified within 30 seconds.

Measures of Phonological Storage and Naming Speed

Short-term memory (STM) measures (phonological loop). Four measures of STM were administered in Spanish and English: Forward Digit Span, Backward Digit Span, Word Span, and Pseudoword Span. The Forward and Backward Digit Span task (taken from the WISC-III; Wechsler, 1991) and a Spanish translated version were administered. The Forward Digit Span task required children to recall sequentially ordered sets of digits that increased in number, which were spoken by the examiner. The technical manual reported a test-retest reliability of .91. The Backward Digit Span task required children to recall sets of digits, but in reverse order. Backward span was included as part of STM since it loads on the same factor as forward digits and is associated with phonological memory (e.g., Colom, Shih, Flores-Mendoza & Quiroga, 2006; Rosen & Engle, 1997). Dependent measures for both tasks were the largest set of items recalled in order (range = 0 to 8 for Digits Forward; range = 0 to 7 for Digits Backward). For the translated Spanish versions of the Digit Span subtest (both Digits Forward and Backward), identical numbers were presented in the same order as the English version. There were no deviations in the procedure, except for language use.

The Word Span and Pseudoword Span tasks were presented in the same manner as the Forward Digit Span task. The Word Span task was previously used by Swanson and Beebe-Frankenberger (2004). Examiners read lists of one or two-syllable, high-frequency words that

MATH PROBLEM-SOLVING AND COGNITION

included unrelated nouns and then asked the children to recall the words. Word lists gradually increased in set size, from a minimum of two words to a maximum of eight. The Pseudoword Span task (Phonetic Memory Span task) uses strings of one-syllable nonsense words, which are presented one at a time in sets of 2 to 6 nonwords (e.g., DES, SEEG, SEG, GEEZ, DEEZ, DEZ). A parallel version was developed in Spanish for the Word Span and Pseudoword Span tests. The dependent measure for all STM measures was the highest set of items retrieved in the correct serial order (range 0 to 7).

Rapid Naming of digits and letters. The administration followed those specified in the manual of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 2000). For this task, the examiner presented the child with an array of items (e.g., six letters or six digits). Children were asked to name the items, speaking in either English or Spanish, as quickly as possible for each stimulus set. The examiner used a stopwatch to time the children as they named all 72 items (2 sets of digits and 2 sets of letters). The dependent measure was the amount of combined time it took for students to complete each set. The number of errors was also taken into account in creating the final score. The manual reports correlations between parallel forms ranging from .80 to .93.

Measures of Inhibition and Executive Component of WM

Random generation. The use of Random Generation tasks has been well articulated in the literature as a measure of inhibition (e.g., Towse & Cheshire, 2007). The task is considered to tap inhibition because participants are required to actively monitor candidate responses and suppress responses that would lead to well-learned sequences, such as 1-2-3-4 or a-b-c-d (Baddeley, 2007). For this study, each child was asked to write numbers (or letters) as quickly as

MATH PROBLEM-SOLVING AND COGNITION

possible, first in sequential order and then in random non-systematic order. For example, children were first asked to write numbers from 0 to 10 in order (i.e., 1, 2, 3, 4) as quickly as possible in a 30-second period. They were then asked to write numbers as quickly as possible “out of order” within a 30-second period. Scoring included an index for randomness, information redundancy, and percentage of paired responses in order to assess the tendency of children to suppress response repetitions.

Executive WM. The Conceptual Span, Listening Sentence Span, Digit Sentence and Updating task were administered in English and Spanish to capture the executive component of WM. Previous studies have shown that these measures load on the executive component of WM (see Swanson, 2017; Swanson et al., 2012, 2015). The WM tasks required children to hold increasingly complex information in memory while simultaneously responding to a question about the task. For example, after children listened to a list of words in the Conceptual Span task they were asked, “Which word from the list did I say, X or Y?” They were then asked to recall words from the list. The balance of simultaneous storage and processing is consistent with a number of studies of WM processing, including Daneman and Carpenter's (1980) seminal WM measure.

Specifically, the Conceptual Span task was used as an indicator of WM processing that involves the ability to organize sequences of words into abstract categories. Children listened to a set of words that, when re-organized, could be grouped into meaningful categories. For example, they were told a word set, such as, “shirt, saw, pants, hammer, shoes, nails.” After answering a distracter question, they were asked to recall the words that “go together” (i.e., shirt, pants, and shoes; saw, hammer, and nails). The range of set difficulty was two categories containing two words each to four categories with four words each. A Spanish- translated

MATH PROBLEM-SOLVING AND COGNITION

version was also administered. Care was taken in the development of the measure to keep the abstract categories the same in both languages (e.g., clothes and tools). The dependent measure for both versions was the number of sets recalled correctly (range 0 to 6).

The children's adaptation (Swanson, 1992; 2013) of Daneman and Carpenter's (1980) Listening Sentence Span task was also administered. The task required the presentation of groups of sentences, read aloud, for which children tried to simultaneously understand the sentence contents and to remember the last word of each sentence. The number of sentences in the group gradually increased from two to six. After each group of sentences was presented, the child answered a question about a sentence and then was asked to recall the last word of each sentence. The dependent measure was the total number of correctly recalled word items in order up to the largest set of items (e.g., set 1 contained 2 items, set 2 contained 3 items, set 3 contained 4 items, etc.), in which the process question was also answered correctly.

The digit sentence task measured the participant's ability to recall numerical information that is embedded within a short sentence. The numerical information referenced either a location or address. On this test, the examiner reads a sentence and then asks the examinee a process question and then asked the examinee to recall the numbers in the sentence.

Because WM tasks were assumed to tap a measure of controlled attention referred to as updating (e.g., Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), an experimental Updating task, adapted from Swanson and Beebe-Frankenberger (2004), was also administered. A series of one-digit numbers was presented that varied in a set length from 3, 5, 7, and 9. No digit appeared twice in the same set. The examiner told the child that the length of each list of numbers might be 3, 5, 7, or 9 digits long. Children were then told that they should only recall the last three numbers presented. Each digit was presented at approximately one-second

MATH PROBLEM-SOLVING AND COGNITION

intervals. After the last digit was presented the child was asked to name the last three digits, in order. The dependent measure was the total number of sets correctly repeated (range 0 to 16).

Visual-spatial sketchpad. Two measures were administered to assess visual-spatial WM: Visual Matrix and Mapping & Directions (Swanson, 1992; 2013). The Visual Matrix task assessed the ability of participants to remember visual sequences within a matrix. Participants were presented a series of dots in a matrix and were allowed 5 seconds to study the matrix. The matrix was then removed and participants were asked, in both English and Spanish, "Are there any dots in the first column?" To ensure the understanding of columns prior to the test, participants were shown the first column location and then practiced finding it on blank matrices. In addition, for each test item, the experimenter pointed to the first column on a blank matrix (a grid with no dots) as a reminder of first column location. After answering the process question, students were asked to draw the dots they remembered seeing in the corresponding boxes of their blank matrix response booklet. The task difficulty ranged from a matrix of 4 squares and 2 dots to a matrix of 45 squares and 12 dots. The dependent measure was the number of matrices recalled correctly (range of 0 to 11).

The Mapping and Directions task required the child to remember a sequence of directions on a map (Swanson, 1992; 2013). The experimenter presented a street map with dots connected by lines; the arrows illustrated the direction a bicycle would go to follow this route through the city. The dots represented stoplights, while lines and arrows mapped the route through the city. The child was allowed 10 seconds to study the map. After the map was removed, the child was asked a process question [i.e., "Were there any stop lights on the first street (column)?"]. The child was then presented a blank matrix on which to draw the street directions (lines and arrows) and stop lights (dots). Difficulty ranged on this subtest from 4 dots

MATH PROBLEM-SOLVING AND COGNITION

to 19 dots. The dependent measure was the highest set of correctly drawn maps (range = 0 - 9) and in which the distracter process question was also answered correctly.

Procedures

Children were tested individually and in groups after informed consent forms were obtained for participation. For each testing wave, two sessions of individual testing were conducted, lasting thirty to sixty minutes for each session. Group testing occurred over the course of two consecutive days for approximately one hour each day. One of four presentation orders related to the individually administered tasks was randomly assigned to each child. In addition, the presentation orders of Spanish and English tests were counterbalanced across all participants. No Spanish or English administration of the same measure followed each other. For the group-administered tests, the presentation order of English and Spanish measures for each type of task was also counterbalanced across groups.

Research Design and Approach to Analyses

Children in grades 1, 2 and 3, designated as English learners from their schools, were tested on the aforementioned battery of achievement, language and cognitive measures in English and Spanish. The results are organized into two parts. First, we determined if our categorization of the tasks provided a good fit to the data. Thus, a confirmatory factor model was computed on the total sample. This was done for measurement purposes (latent variables control for measurement error as different variables have different weightings on a construct) and also for practical reasons: some constructs (e.g., WM, STM) included several tasks. More important there are several advantages of using latent measures over-reliance on single task measures (e.g., focus on construct rather than the variance of a single task).

The model for the confirmatory analysis was based on an earlier study with monolingual

MATH PROBLEM-SOLVING AND COGNITION

children (e.g., Swanson & Beebe-Frankenburg, 2004; Swanson & Fung, 2016). Overall, the factors include math problem-solving (WJ-applied problems, WISC-III word problems), reading (word identification, passage comprehension, WRMT-word attack), calculation (WJ-calculation, WRAT-Calculation), vocabulary (PPVT receptive, EOWVT- expressive), numeracy (large numbers, small numbers), estimation (line estimation-same line length, line length varied), STM (span=nonword, real words, digit forward, digit backward), naming speed (letters, numbers), executive component of WM (conceptual span, sentence span, listening span, updating), and inhibition (random generation= letters and numbers). For the present study, unique factors were computed for the English and Spanish presentation. Also included were factors that measured nonverbal reasoning (Raven Colored Progressive Matrices Test-sections A, AB, and B) and visual-spatial WM (Mapping/directions, visual matrix). For these last two factors, the presentation of instructions was in both English and Spanish. We used the SAS CALIS (2012) program to create factor scores (latent variables) for each set of measures with two or more variables. The procedure allowed us to calculate standardized beta weights. Task weightings for the latent measures used in the analysis (as well as means and *SDs* and task sample reliabilities) are presented in Appendix A. Based on the standardized loadings in Appendix A, latent scores were computed by multiplying the z-score of the target variable by the standardized factor loading weight based on the total sample (see Nunnally & Bernstein, 1994, p. 508 for calculation procedures). All measures were scaled to have a mean of 0 and a standard deviation of 1. Normality of the data was considered. Except for the naming speed tasks, measures met standard criteria for univariate normality with skewness for all measures less than 3 and kurtosis less than 4. Transformations were conducted on these variables and the results of each model with and without these transformations were compared. Since there were no substantial differences in the

model results, the untransformed scores were used for analyses.

Second, we used a multilevel logistic model, via SAS PROC GLIMMIX software (SAS, 2010), to analyzed differences between children at risk or not at risk for MD. Of interest was whether the odds of being identified as MD increased as a function of cognitive measures and bilingual proficiency. Given the non-normal nature of our outcome variables, the use of ML estimation was not appropriate for this analysis. Instead, for this analysis, we used a common estimation technique with PROC GLIMMIX referred to as Laplace estimation.

The equation for estimating the unconditional model was:

$$\eta_{ij} = \beta_{0j} + \beta_{1j} X_{ij} \quad (\text{Eq. 2})$$

Equation 2 represented a simple level-1 model with one student-level predictor, where η_{ij} represented the log odds of not being designated MD for student i in classroom j , β_{0j} is the intercept or the average log odds of not being designated at risk for MD at classroom j , i_j is a student-level predictor for student i in classroom j , and β_{1j} represents the slope associated with X_{ij} , showing the relationship between the student-level variable and the log odds of not being designated at risk for MD. It is important to notice that unlike hierarchical linear models used to analyze the total sample, this model has no error variance at level-1 (see Hox, 1999, pp. 225-227).

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + u_{0j} \quad (\text{Eq. 3})$$

$$\beta_{1j} = \gamma_{10}$$

Equation 3 represented a simple classroom level-2 predictor, where γ_{00} provided the log odds of “not” being designated MD at a typical classroom, W_j was a level 2 predictor for classroom j ,

MATH PROBLEM-SOLVING AND COGNITION

γ_{01} was the slope associated with this predictor, u_{0j} was the level-2 error term representing a unique effect associated with the nested effect of classroom j , and γ_{10} was the average effect of the student-level predictor. As the effect of the student-level predictor was modeled as fixed or constant across classrooms, this was represented as a random intercept-only model.

$$\eta_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} W_j + u_{0j} \quad (\text{Eq. 4})$$

The combined level-1 and level-2 model (Equation 4) represented the log odds of not being designated as MD for student i in classroom j (η_{ij}) at a typical classroom (γ_{00}), at the student-level ($\gamma_{10} X_{ij}$) and classroom-level predictor ($\gamma_{01} W_j$), as well as the classroom-level error [u_{0j} , $u_{0j} \sim N(0, \tau_{00})$].

Results

Normative Sample Representation

Table 1 shows the means and standard deviations on normed referenced measures as a function of the total sample and as a function of the math group. The normative scores yielded three important patterns. First, variations emerged on receptive and expressive vocabulary scores. Receptive language scores in Spanish were higher than receptive language scores in English, while the reverse pattern emerged for expressive language scores. Overall, the mean vocabulary scores were in the below average range and/or low average range suggesting the sample is best characterized as emerging bilinguals rather than Spanish dominant. Second, the means for the norm-referenced scores were substantially higher on basic skill measures (calculation, word identification) than on higher level skill measures such as problem-solving accuracy and reading comprehension. Finally, a general pattern across the children at risk and

those not at risk was that normative scores in problem-solving accuracy were lower in children at risk.

Data Preparation

Confirmatory factor analysis. To determine if our a priori categorization of measures fit the data, the Comparative Fit Index (CFI), Bentler-Bonett non-normed fit index (NNFI), and root mean square error of approximation (RMSEA) were computed. Values at .90 and over on the CFI, and NNFI and RMSEA values of .05 or less indicate an acceptable fit. The model provided an acceptable fit to the data (CFI = .92, NNFI=.90 RMSEA = .044; 90% CI: .039 to .048). Clearly, a second order model could have been tested that reflected constructs that overlapped between the two languages (i.e., language independent constructs), however, as stated in our research question we were primarily interested in individual differences in accessing information within and across the language systems and if such differences played a significant role in the predictions of math problem-solving accuracy. Thus, we assumed that individual differences that emerge on each measure were related to the ease of access within the preferred language, and not a language-specific cognitive system (also see Swanson, Kudo, & Van Horn, 2018). Based on the standardized loadings in Appendix A, latent scores were computed by multiplying the z-score of the target variable by the standardized factor loading weight based on the total sample.

Given these preliminary analyses, the following analyses focused on the first research question.

1. Do measures of WM uniquely predict MD status when measures of vocabulary, reading and domain-specific measures are entered into the regression model?

Mean comparisons. Prior to answering this question, a comparison was made across the measures as a function of MD status. As expected, children with and without MD varied significantly on classification variable ($M=80.95$, $SD=9.33$ vs. $M=101.42$, $SD=6.97$) that was the median score across normative measures of English and Spanish math problem-solving, $F(1,387)=192.49$, $p < .0001$. A mixed ANOVA compared the two groups on the array of latent measures. Table 2 reports the means and standard deviations on the array of measures as a function of the two groups. Because of the number of comparisons and thereby increasing Type 1 error, a Bonferroni correction procedure set the alpha level to .002.

As shown in Table 2, significant differences emerged between the groups on the majority of measures. Significant group differences emerged on English and Spanish measures of reading, STM, and WM. Significant group differences also occurred on English measures of vocabulary and estimation and Spanish measures of numeracy. Because several variables significantly differentiated children with and without MD, a multilevel logistic regression was computed to determine those variables that uniquely predicted MD status.

Multilevel logistic model. The estimates for the multilevel logistic unconditional and conditional models in predicting the odds of being classified as MD vs. non-MD are shown in Table 3. The unconditional multilevel logistic model was assumed to have no error at level-1 (Snijders & Bosker, 1999). That is, the level-1 residual follows a logistic distribution with a mean of 0 and a variance of 3.29 for dichotomous outcomes (Snijders & Bosker, 1999, p. 227). The intraclass correlation was computed as .17 $(.68/.68 + 3.29)$, suggesting that approximately 17% of the variability was accounted for by children nested in classrooms, leaving approximately 83% of the variability to be accounted for by the latent measures (or other unknown factors).

As shown in Table 3, the intercept for the unconditional model indicated a significant amount of variability in the log odds of being classified as “Non-MD” in our sample (log odds=1.11). Grade level and bilingual proficiency were significant predictors of the odds of “not” being classified MD in Model 1. (Initially, we entered children’s age in months into the regression model; however, this variable was not a significant predictor and therefore for parsimony was removed from the regression model).

Model 1 showed that as grade level increased, the odds of being designated “non-MD” decreased significantly in the upper grades. In addition, Model 1 showed that increased bilingual proficiency was related to the increased odds of “not” being identified as MD.

Along with measures related to the executive processing, Model 2 entered the phonological storage variables in predictions of the classification variable. Consistent with several studies we assumed that the residual variance related to WM tapped an executive system after measures related to STM (storage) were entered into the analysis (e.g., Engle, Tuholski, Laughlin, & Conway, 1999). The non-MD designation was significantly increased by increased performance on both English and Spanish latent measures of STM and WM. Further, Model 2 reduced the variance in the MD vs. non-MD designation related to the classroom by 33% ($.34/.34 + .68$).

Model 3 entered variables related to domain-specific processes. As shown, this model eliminated significant variance related to the nested effect of the classroom assignment. Further, the model eliminated the significant contribution of the English memory measures in predicting MD status. The model also indicated that increased performance on both English and Spanish

numeracy and the Spanish Estimation latent measures increased the odds of “not” being classified as MD.

The complete model (Model 4) entered the variables related to reading and fluid intelligence. Significant predictors in this model were grade level, bilingual proficiency, Spanish reading, Spanish numeracy, and Spanish WM. The important finding was that the executive component of WM within the Spanish language system was both significant and positive, indicating that as a student’s WM increased within the Spanish language system, the predicted log odds of a student “not” being designated as MD also increased. The final model addressed the second question of the study.

2. Does bilingual proficiency influence the role of cognitive measures in predictions of MD status?

Because bilingual proficiency has been associated with executive processing, as a follow-up analysis we removed the bilingual variable from the full model (Model 4) to determine if variations in the parameters emerged. Thus, a comparison was made between Model 4 that included the bilingual variable and Model 5 the removed the bilingual variable. A comparison was made by estimating the reduction in the parameter estimates between the two models (Models 4 and 5). For example, as shown in Table 3, the bilingual effect on visual-spatial WM was 20 % $(0.40 - 0.32)/.40$. Thus, 20% of the influence of visual-spatial WM in predicting math problem-solving groups was moderated by bilingual proficiency. As shown at the far right of Table 3, however, the largest reduction related to bilingual proficiency (ratios > 50%) was on the executive component of English WM.

Table 3 also showed no reduction related to the odds of predicting MD status when the bilingual proficiency variable was removed from the regression model. That is, the odds of predicting MD status were 2.39 in the Full Model (Model 4) when compared to the odds of 2.35 when bilingual proficiency was removed from the Full Model (Model 5). A similar pattern of “non-influence” (as indicated by removing bilingual proficiency from the model) occurred on measures of achievement (reading and calculation).

In general, a comparison of Model 4 and 5 shows that bilingual proficiency yielded a positive influence on latent measures of fluid intelligence and executive processing within the English language system. These findings are consistent with others suggesting that bilingual proficiency has a direct impact on measures of executive processing. However, bilingual proficiency did not have a direct impact on MD status.

Discussion

The study examined the cognitive processes that were significantly related to math problem-solving accuracy in emerging bilingual children as well as those processes that underlie the probability of being at risk for MD. To this end, children in grades 1, 2 and 3 were assessed in both Spanish and English on a large array of cognitive and achievement measures. Two important findings emerged. First, the results directly supported the notion that the executive component of WM uniquely predicted MD status when latent measures of fluid intelligence, reading, domain-specific processes and phonological storage were entered into the regression model. Although the executive component of WM was not the only significant predictor, the results clearly indicated that the executive component of WM in Spanish uniquely predicted MD status across.

Second, clear cognitive advantages related to bilingual proficiency emerged. The results

MATH PROBLEM-SOLVING AND COGNITION

clearly showed that increased bilingual proficiency played a role in mediating the influence of cognitive variables in predicting the probability of not being designated as MD. However, removal of bilingual proficiency from the full regression model yielded no reduction in the odds of being designated MD. Thus, when all the variables are taken together, the influence of bilingual proficiency on predicting MD status may be indirect. As shown in Table 3, a comparison of Model 4 and 5 suggested that changes (reduction in parameter estimation) emerged in the parameter estimates of fluid intelligence and executive processes when the bilingual variable was removed from the regression model. No such reduction occurred related to the odds of predicting MD status. In contrast, the results showed that at least 50% of changes in the parameter estimates between MD status and executive processing were related to bilingual proficiency. Therefore, we assume that increased bilingual proficiency increased the chances of executive processes within the English language system as playing a major role in decreasing the odds of children being label MD. These findings are consistent with others who find a strong relationship between proficiency in L1 and L2 (i.e., bilingualism) and executive processing (e.g., flexibility and intentional control, Bialystok, 2011).

Competing Models

No doubt, there are several competing interpretations as to whether the executive component of WM underlies math problem-solving performance; two are considered. First, individual differences in executive processing are merely an artifact of individual differences in the phonological loop. That is, one of the components of WM commonly attributed to problem-solving performance is the phonological loop. The model suggests that the phonological system has a bottom-up influence on WM. Because the phonological loop (STM) has been found to play a major role in L2 reading among emerging bilingual children (e.g.

MATH PROBLEM-SOLVING AND COGNITION

Swanson et al. 2004) and reading performance is highly correlated with math problem-solving in the younger grades, it can be argued that the influence of higher order processes, such as executive processing, in predictions of problem-solving is an artifact of phonological storage (e.g., Peng et al., 2016). A key finding of this study, however, was that MD status was significantly related to WM performance even when specific processes (STM, inhibition) or more general processes (naming speed) were entered into the analysis.

A second possibility is that resistance to interference between the two language systems underlies the influence of WM on math performance in emerging bilingual children. The model suggests that interference may be related to the competing influence of Spanish on math measures and/or a general ability to suppress competing linguistic information. As a consequence, children with MD, when compared to children not at risk, may have trouble preventing unnecessary information in the Spanish language system from entering memory on English measures. However, this interpretation does not appear to be a viable alternative to the present results, for three reasons.

First, removal of bilingual proficiency from the full model had minimal influence on latent measures of inhibition as well as the odds of being designated as MD. As shown in Table 3, bilingual proficiency yielded its greatest influence in terms of parameter reductions on measures of fluid intelligence, estimation, and WM when compared its influence on inhibition measures. Second, neither latent measures of inhibition in Spanish nor English were significant covariates in the final mixed regression models predicting the odds of being classified as MD. Thus, our measure of inhibition did not play a major role in explaining individual differences in math problem-solving. Finally, and perhaps more importantly, the Spanish and English inhibition measures may reflect different monitoring activities than the executive component of WM. The

executive component of WM may be more closely aligned with updating than inhibition activities. Thus, updating may be more closely aligned with problem-solving in children with MD than inhibition.

Theoretical Implications

What is the theoretical importance of the findings? We provide three applications to theory. First, our findings qualify bottom-up models suggesting that low-order processes (e.g., phonological STM) mediate the influence of executive processing (WM) on children's mathematical problem-solving performance. Although our findings indicate that math problem-solving is associated with the phonological loop (i.e. STM), this latent measure did not completely account for the influence of the executive component of WM on math performance. Thus, the influence of low-order processes on executive processing was minimal for children in this study. Of course, these results apply only to the age and ability groups represented in this sample.

A second implication is that only two of the components in Baddeley's three components model independently predicted math problem-solving performance. Both Spanish STM and WM predicted MD status (see Table 3). Finding that WM and STM tasks contributed unique variance to the MD classification is important outcomes since it varies from studies suggesting the two memory measures tap the same construct (e.g., load onto the same factor, Hutton & Towse, 2001) as well as differs from other studies that have found WM is a nonsignificant predictors when measures of reading and fluid intelligence have been entered into the analysis (e.g., Fuchs et 2005; Peng et al., 2016).

Finally, the findings do not align well with the hypothesis that the relationship between WM and problem-solving are primarily moderated by language interference and/or proficiency in reading (e.g., Fuchs et al., 2005; Peng et al., 2016). The influence of WM in predicting math problem-solving status was not partialled out when reading, STM, and inhibition measures were entered into the regression analysis. What the results do suggest is that high performance on the executive or controlled attention component of WM is related to increased performance on problem-solving measures.

Practical Implications

There are two practical implications related to our findings. First, there is a language specific effect that influences performance the odds of being classified with MD. We find that Spanish measures of numeracy and the executive component of working memory are particularly well-suited to identify children at risk for MD who are emerging bilinguals and/or learning English as a second language. These findings are important because confounds exist in the assessment of children with potential MD who are second language learners. These confounds are due in part to attributing difficulties in second language acquisition and reading acquisition to the same cognitive processes that are involved in mathematical problem-solving. In practice, these confounds may lead to English language learners being inappropriately diagnosed with MD and placed in special education. The opposite condition is also true that children who are at potential risk for MD are being overlooked and not being provided intervention. In order to circumvent some of these problems, it is necessary to identify the processes in children at risk for MD from other processes related to second language acquisition. Our findings clearly show that children at risk for MD but of average intelligence are deficient on Spanish measures of numeracy and the executive component of WM.

A second implication is that bilingual proficiency appears to have its most profound effect for emerging bilinguals on cognitive measures of English rather than Spanish. These findings are of interest because the normative receptive language scores in English and Spanish were in the same range and English expressive language scores were higher than Spanish expressive language scores. As shown in Table 3, the largest reductions in the parameter estimates when the bilingual variable was removed from the regression model were on measures of English calculation, English estimation, English STM, English naming speed and English WM. There was a minimal reduction in parameter estimates related to bilingual proficiency on measures of Spanish calculation, Spanish estimation, and Spanish WM. Thus, the results suggest that bilingual proficiency within the current sample has more to do with performance in the second language (English) than the first language (Spanish).

Limitations

There are at least four limitations to this study. First, the design of the study was cross-sectional instead of longitudinal. In order to investigate language dominance and language shift in bilingual children with MD, longitudinal studies in which the development of linguistic skills is monitored in the course of time are necessary. Second, the sample yielded minimal variance in SES (98% of the sample was on Federal assistance programs) and therefore the influence of high versus low SES could not be evaluated.

Third, although our primary focus is on the relationship between WM and its effects on math problem-solving performance, no doubt other variables (covariates) should have been considered in the analysis. We selected covariates from variables that have been found significantly related to academic achievement and language acquisition in the literature as

MATH PROBLEM-SOLVING AND COGNITION

potential mediators that influence phonological storage (phonological knowledge, naming speed) and executive processing (inhibition, inattention), but other variables need to be considered.

Finally, the sample reflected sequential bilingualism (L2 follows L1 development) and therefore may not reflect bilingualism when two languages are learned simultaneously. It is important to note that the majority of these studies on executive processing and bilingualism have focused on children who learned L1 and L2 simultaneously. However, emerging bilingual children in U.S. public schools frequently represent children who learn L1 first and L2 later (as they enter school). Thus, few studies have focused on sequential bilinguals (who learn their L1 first, then L2 later) with different levels of language proficiency on executive processing and math. If bilingualism influences cognitive processes, it is possible that a positive cognitive impact on math performance appears as bilinguals gain higher degrees of bilingual proficiency.

Summary

Taken together, we interpret our findings as suggesting that math problem-solving performance in emerging bilingual children is directly tied to the WM system. The results suggest that when the effects of vocabulary, reading, fluid intelligence, naming speed and inhibition were partialled out, the executive system component of WM was related to problem-solving ability. Further, the results suggested that the probability of risk for MD within the current emerging bilingual sample of children was significantly tied to poor performance on Spanish measures of numeracy and the executive component of WM. The results also suggest that the influence of bilingual proficiency was primarily directed to executive processing measures within the English language system and not to measures of academic achievement (e.g., reading, calculation).

References

- Baddeley, A. (2007). *Working memory, thought, and action*. New York, NY, US: Oxford University Press. doi:10.1093/acprof:oso/9780198528012.001.0001
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1-29. doi: [/10.1146/annurev](https://doi.org/10.1146/annurev)
- Baddeley, A. D., & Logie, R. H. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). Cambridge, U.K.: Cambridge University Press.
- Bialystok, E. (2011). Coordination of executive functions in monolingual and bilingual children. *Journal of Experimental Child Psychology*, 110, 461-468. doi:10.1016/j.jecp.2011.05.005
- Bialystok, E., & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and language*, 109(2), 93-100. doi: 10.1016/j.bandl.2008.09.001
- Bjork, I. M., & Bowyer-Crane, C. (2013). Cognitive skills used to solve mathematical word problems and numerical operations: A study of 6- to 7-year-old children. *European Journal of Psychology of Education*, 28(4), 1345-1360. doi:/10.1007/s10212-012-0169-7
- Bonifacci, P., Giombini, L., Bellocchi, S., & Contento, S. (2011). Speed of processing, anticipation, inhibition and working memory in bilinguals. *Developmental Science*, 14(2), 256-269.
- Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, 79(4), 1016-1031. doi:/10.1111/j.1467-8624.2008.01173.x

MATH PROBLEM-SOLVING AND COGNITION

- Branum-Martin, L., Fletcher, J. M., & Stuebing, K. K. (2013). Classification and identification of reading and math disabilities: The special case of comorbidity. *Journal of Learning Disabilities, 46*(6), 490-499.
- Brownell, K. (2001). *Expressive One-Word Picture Vocabulary Test* (3rd Edition). New York: Academic Therapy Publications.
- Censabella, S., & Noel, M. P. (2008). The inhibition capacities of children with mathematical disabilities. *Child Neuropsychology, 14*(1), 1-20. doi:10.1080/09297040601052318
- Cirino, P. T., Fuchs, L. S., Elias, J. T., Powell, S. R., & Schumacher, R. F. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities, 48*(2), 156-175. doi:[10.1177/0022219413494239](https://doi.org/10.1177/0022219413494239)
- Chew, C. S., Forte, J. D., & Reeve, R. A. (2016). Cognitive factors affecting children's nonsymbolic and symbolic magnitude judgment abilities: A latent profile analysis. *Journal of Experimental Child Psychology, 152*, 173-191. doi:[10.1016/j.jecp.2016.07.001](https://doi.org/10.1016/j.jecp.2016.07.001)
- Colom, R., Shih, P. C., Flores-Mendoza, C., & Quiroga, M. Á. (2006). The real relationship between short-term memory and working memory. *Memory, 14*(7), 804-813. doi:10.1080/09658210600680020
- Cooper, R. P. (2016). Executive functions and the generation of "random" sequential responses: A computational account. *Journal of Mathematical Psychology, 73*, 153-168. doi:10.1016/j.jmp.2016.06.002
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222-251. doi:10.3102/00346543049002222

- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
doi:10.1016/S0022-5371(80)90312-6
- David, C. V. (2012). Working memory deficits in math learning difficulties: A meta-analysis. *International Journal of Developmental Disabilities*, 58(2), 67-84.
doi:[10.1179/2047387711Y.00000000007](https://doi.org/10.1179/2047387711Y.00000000007)
- de Abreu, Pascale M. J. Engel. (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory*, 19(5), 529-537. doi:10.1080/09658211.2011.590504
- Dunn, L. M., & Dunn, L. M. (2007). *The Peabody Picture Vocabulary Test-4*. NY: Pearson.
- Dunn, L. M., Lugo, D. E., Padilla, E. R., & Dunn, L. M. (1986). *Test de Vocabulario Imágenes Peabody*. Circle Pines, MN: American Guidance Service.
- Engel, D. A., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism enriches the poor: Enhanced cognitive control in low-income minority children. *Psychological Science*, 23(11), 1364-1371. doi:/10.1177/0956797612443836
- Engel, D. A., & Gathercole, S. E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology*, 104(4), 974-986.
Doi:10.1037/a0028390
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1), 19-23. doi: 10.1111/1467-8721.00160
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, 128, 309-331. doi:10.1037//0096-3445.128.3.309
- Farnia, F., & Geva, E. (2011). Cognitive correlates of vocabulary growth in English language

learners. *Applied Psycholinguistics*, 32, 711-738. doi:10.1017/S0142716411000038

Friedman, N. P., Haberstick, B. C., Willcutt, E. G., Miyake, A. Young, S. Corely, R. P. & Hewitt, J. K. (2007). Greater attention problems during childhood predict poorer executive functioning in late adolescence. *Psychological Science*, 18, 893-900.

doi:10.1111/j.1467-9280.2007.01997

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(3), 493-513. doi:/10.1037/0022-0663.97.3.493

Fuchs, L. S., Compton, D. L., Fuchs, D., Powell, S. R., Schumacher, R. F., Hamlett, C. L., . . . Vukovic, R. K. (2012). Contributions of domain-general cognitive resources and different forms of arithmetic development to pre-algebraic knowledge. *Developmental Psychology*, 48(5), 1315-1326. doi:/10.1037/a0027475

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, 98(1), 29-43. doi:10.1037/0022-0663.98.1.29

Fuchs, L. S., Gilbert, J. K., Powell, S. R., Cirino, P. T., Fuchs, D., Hamlett, C. L., . . . Tolar, T. D. (2016). The role of cognitive processes, foundational math skill, and calculation accuracy and fluency in word-problem-solving versus prealgebraic knowledge. *Developmental Psychology*, 52(12), 2085-2098. doi:/10.1037/dev0000227

MATH PROBLEM-SOLVING AND COGNITION

- Gathercole, S. E., Pickering, S. J., Ambridge, B., Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40, 177-190.
doi:10.1037/0012-1649.40.2.177
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology*, 47(6), 1539-1552. doi:/10.1037/a0025510
- Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology*, 104(1), 206-223.
doi:/10.1037/a0025398
- Geary, D. C., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology*, 109(5), 680-693.
doi:/10.1037/edu0000159
- Gorman, B. K. (2012). Relationships between vocabulary size, working memory, and phonological awareness in Spanish-speaking English language learners. *American Journal of Speech-Language Pathology*, 21(2), 109-123. doi:/10.1044/1058-0360(2011/10-0063
- Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language*, 92, 183-201. doi:10.1016/j.jml.2016.06.0
- Han, W. (2012). Bilingualism and academic achievement. *Child Development*, 83(1), 300-321.
doi: 10.1111/j.1467-8624.2011.01686.x

MATH PROBLEM-SOLVING AND COGNITION

- Harvey, H. A., & Miller, G. E. (2017). Executive function skills, early mathematics, and vocabulary in head start preschool children. *Early Education and Development*, 28(3), 290-307. doi:/10.1080/10409289.2016.1218728
- Hox, J. (2010). *Multilevel Analysis: Techniques and Applications* (2nd Ed.) New York, NY: Routledge/Taylor & Francis.
- Hutton, U. M. Z., & Towse, J. N. (2001). Short-term memory and working memory as indices of children's cognitive skills. *Memory*, 9(4-6), 383-394. doi:[10.1080/09658210042000058](https://doi.org/10.1080/09658210042000058)
- Johnson, E. S., Humphrey, M., Mellard, D. F., Woods, K., & Swanson, H. L. (2010). Cognitive processing deficits and students with specific learning disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly*, 33(1), 3-18. doi: 10.1177/073194871003300101
- Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler Intelligence scale for Children--fourth edition: What does it measure? *School Psychology Review*, 35(1), 108-127. Retrieved from <https://search.proquest.com/docview/621623099?accountid=14521>
- Kempert, S., Saalbach, H., & Hardy, I. (2011). Cognitive benefits and costs of bilingualism in elementary school students: The case of mathematical word problems. *Journal of Educational Psychology*, 103(3), 547-561. doi:/10.1037/a0023619
- Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2014). Involvement of working memory in longitudinal development of number–magnitude skills. *Infant and Child Development*, 23(1), 36-50. doi:[10.1002/icd.1834](https://doi.org/10.1002/icd.1834)

MATH PROBLEM-SOLVING AND COGNITION

- Landerl, K., & Moll, K. (2010). Comorbidity of learning disorders: Prevalence and familial transmission. *Journal of Child Psychology and Psychiatry*, 51(3), 287-294.
doi:[10.1111/j.1469-7610.2009.02164.x](https://doi.org/10.1111/j.1469-7610.2009.02164.x)
- Lee, K., Ng, S-F., Ng, E-L, & Lim, Z-Y. (2004). Working memory and literacy as predictors of performance on algebraic word problems. *Journal of Experimental Child Psychology*, 89(2), 140-158. doi:10.1016/j.jecp.2004.07.001
- Lonigan, C. J., Allan, D. M., Goodrich, J. M., Farrington, A. L., & Phillips, B. M. (2017). Inhibitory control of Spanish-speaking language-minority preschool children: Measurement and association with language, literacy, and math skills. *Journal of Learning Disabilities*, 50(4), 373-385. doi:[10.1177/0022219415618498](https://doi.org/10.1177/0022219415618498)
- Macizo, P., Herra, A., Roman, P., & Marin, M.C. (2011). Proficiency in a second language influences the processing of number words. *Journal of Cognitive Psychology*, 23, 915-921. Doi:10.1080/20445911.2011.586626
- Martin, R. B., Cirino, P. T., Barnes, M. A., Ewing-Cobbs, L., Fuchs, L. S., Stuebing, K. K., & Fletcher, J. M. (2013). Prediction and stability of mathematics skill and difficulty. *Journal of Learning Disabilities*, 46(5), 428-443. Retrieved from <https://search.proquest.com/docview/1651841294?accountid=14521?accountid=14521>
- Martiniello, M. (2008). Language and the performance of ELL in math word problems. *Harvard Educational Review*, 78(2), 333-368. doi:[10.17763/haer.78.2.70783570r1111t32](https://doi.org/10.17763/haer.78.2.70783570r1111t32)
- Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment*, 14(3-4), 160-179. doi:[10.1080/10627190903422906](https://doi.org/10.1080/10627190903422906)

MATH PROBLEM-SOLVING AND COGNITION

- Mayer, R. E., & Hegarty, M. (1996). The process of understanding mathematical problem-solving. In R.J. Sternberg & T. Ben-Zeev (Eds.). *The Nature of Mathematical Thinking* (pp.29-54). Mahwah, NJ: Erlbaum.
- Menon, V. (2016). Working memory in children's math learning and its disruption in dyscalculia. *Current Opinion in Behavioral Sciences*, 10, 125-132. doi:/10.1016/j.cobeha.2016.05.014
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100.
doi:10.1006/cogp.1999.0734
- Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, 130(4), 621-640.
doi:10.1037/0096-3445.130.4.621
- Murphy, M.M., Mazzocco, M.M.M., Hanich, L.B., & Early, M.C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, 40(5), 458-478.
- National Mathematics Advisory Panel (2008). The final report of the National Mathematics Advisory Panel. U.S. Department of Education. Retrieved September, 2014, from <http://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf>
- National Assessment of Educational Progress (2011). Achievement gap: How Hispanics and white students in public schools perform in mathematics and reading on the national assessment of educational progress. Washington DC: US Department of Education.

MATH PROBLEM-SOLVING AND COGNITION

National Assessment of Educational Progress (2013). Math report Card. Washington DC: US Department of Education.

National Assessment of Educational Progress (2017). The condition of education (update 2017) Washington DC: US Department of Education.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd Ed.). NY: McGraw-Hill.

Ockey, G. J. (2007). Investigating the validity of math word problems for English Language

Learners with DIF. *Language Assessment Quarterly*, 4(2), 149-164.

doi:[10.1080/15434300701375717](https://doi.org/10.1080/15434300701375717)

OECD (2012). Programme for International Student assessment (PISA): Results from PISA

2012. Retrieved September, 2014, from [http://www.oecd.org/pisa/keyfindings/PISA-](http://www.oecd.org/pisa/keyfindings/PISA-2012-results-US.pdf)

[2012-results-US.pdf](http://www.oecd.org/pisa/keyfindings/PISA-2012-results-US.pdf)

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem-solving. *Journal of*

Experimental Child Psychology, 80, 44-57. doi:10.1006/jecp.2000.2626

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research.

Child Development, 78(4), 1255-1264. doi:[10.1111/j.1467-8624.2007.01064.x](https://doi.org/10.1111/j.1467-8624.2007.01064.x)

Peña, E. D., Bedore, L. M., & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one.

International Journal of Language & Communication Disorders, 51(2), 192-202.

doi:[10.1111/1460-6984.12199](https://doi.org/10.1111/1460-6984.12199)

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and

working memory: Moderating effects of working memory domain, type of mathematics

skill, and sample characteristics. *Journal of Educational Psychology*, 108(4), 455-473.

doi:[10.1037/edu0000079](https://doi.org/10.1037/edu0000079)

Peng, P., Namkung, J. M., Fuchs, D., Fuchs, L. S., Patton, S., Yen, L., Hamlett, C. (2016). A longitudinal study on predictors of early calculation development among young children at risk for learning difficulties. *Journal of Experimental Child Psychology*, 152, 221-241. doi:[10.1016/j.jecp.2016.07.017](https://doi.org/10.1016/j.jecp.2016.07.017)

Raven, J. C. (1976). *Colored Progressive Matrices*. London, England: H. K. Lewis & Co. Ltd.

Rosen, V. M., & Engle, R. W. (1997). Forward and backward serial recall. *Intelligence*, 25(1), 37-47.

Rosselli, M., Ardila, A., Lalwani, L. N., & Vélez-Urbe, I. (2016). The effect of language proficiency on executive functions in balanced and unbalanced Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 19(3), 489-503. doi:[10.1017/S1366728915000309](https://doi.org/10.1017/S1366728915000309)

Rousselle, L., & Noël, M. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude. *Cognition*, 102(3), 361-395. doi:[10.1016/j.cognition.2006.01.005](https://doi.org/10.1016/j.cognition.2006.01.005)

SAS Institute. (2010). *SAS/STAT software: Changes and Enhancements through release 9.3*. Cary, NC: SAS Institute Inc.

Siegler, R. S. & Booth, J. (2004). Development of numerical estimation in young children. *Child Development*, 75(2), 428-444. doi:[10.1111/j.1467-8624.2004.00684.x](https://doi.org/10.1111/j.1467-8624.2004.00684.x)

Siegler, R. S., & Opfer, J. (2003). The development of numerical estimation: Evidence For multiple representation of numerical quantity. *Psychological Science*, 14(3), 237-243. doi:[10.1111/1467-9280.02438](https://doi.org/10.1111/1467-9280.02438)

MATH PROBLEM-SOLVING AND COGNITION

- Snijders, T., & Bosker, R. (1999). *Multilevel modeling: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage Press.
- Swanson, H. L. (1992). Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology*, 84(4), 473-488. doi: 10.1037/0022-0663.84.4.473
- Swanson, H. L. (2013). *Abbreviated test of working memory*. Washington, DC: American Psychological Association. PsycTESTS.
- Swanson, H. L. (2017). Verbal and visual-spatial working memory: What develops over a life span? *Developmental Psychology*, 53(5), 971-995. doi:10.1037/dev0000291
- 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(3), 471-491. doi:10.1037/0022-0663.96.3.471
- Swanson, H. L., & Fung, W. (2016). Working memory components and problem-solving accuracy: Are there multiple pathways? *Journal of Educational Psychology*, 108(8), 1153-1177. doi:10.1037/edu0000116
- Swanson, H. L., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research*, 76(2), 249-274. doi:10.3102/00346543076002249
- Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, 100(2), 343-379. doi:10.1037/0022-0663.100.2.343

MATH PROBLEM-SOLVING AND COGNITION

- Swanson, H. L., Kong, J., & Petcu, S. (2018). Math difficulties and working memory growth in English language learner children: Does bilingual proficiency play a significant role? *Language, Speech, and Hearing Services in Schools, 49*(3), 379-394. doi:10.1044/2018_LSHSS-17-0098
- Swanson, H. L., Kong, J. E., & Petcu, S. D. (2019). Individual differences in math problem solving and executive processing among emerging bilingual children. *Journal of Experimental Child Psychology, 187*, 25. doi:10.1016/j.jecp.2019.06.006
- Swanson, H. L., Kudo, M., & Guzman-Orth, D. (2016). Cognition and literacy in english language learners at risk for reading disabilities: A latent transition analysis. *Journal of Educational Psychology, 108*(6), 830-856. doi:10.1037/edu0000102
- Swanson, H. L., Kudo, M. F., & Van Horn, M. L. (2018). Does the structure of working memory in el children vary across age and two language systems? *Memory*, doi:10.1080/09658211.2018.1496264
- Swanson, H. L., Olide, A. F., & Kong, J. E. (2018). Latent class analysis of children with math difficulties and/or math learning disabilities: Are there cognitive differences? *Journal of Educational Psychology, 110*(7), 931-951. doi: 10.1037/edu0000252
- Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2012). Cognition and literacy in English language learners at risk for reading disabilities. *Journal of Educational Psychology, 104*(2), 302-320. doi:10.1037/a0026225
- Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2015). Growth in literacy, cognition, and working memory in English language learners. *Journal of Experimental Child Psychology, 132*, 155-188. doi: 10.1016/j.jecp.2015.01.001

MATH PROBLEM-SOLVING AND COGNITION

- Swanson, H. L., Sáez, L., Gerber, M., & Leafstedt, J. (2004). Literacy and cognitive functioning in bilingual and nonbilingual children at or not at risk for reading disabilities. *Journal of Educational Psychology, 96*(1), 3-18. doi:10.1037/0022-0663.96.1.3
- Thorn, A. S. C., & Gathercole, S. E. (2001). Language differences in verbal short-term memory do not exclusively originate in the process of subvocal rehearsal. *Psychonomic Bulletin & Review, 8*, 357-365.
- Toll, S. W. M., Van, d. V., Kroesbergen, E. H., & Van Luit, Johannes E. H. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities, 44*(6), 521-532. doi:10.1177/0022219410387302
- Towse, J., & Cheshire, A. (2007). Random generation and working memory. *European Journal of Cognitive Psychology, 19* (3), 374-394. doi:10.1080/09541440600764570
- Vukovic, R. K., & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology, 115*(2), 227-244. doi: 10.1016/j.jecp.2013.02.002
- Wagner, R., Torgesen, J., & Rashotte, C. (2000). *Comprehensive Test of Phonological Processes*. Austin TX: Pro-ED.
- Weber, R. C., Johnson, A., Riccio, C. A., & Liew, J. (2016). Balanced bilingualism and executive functioning in children. *Bilingualism: Language and Cognition, 19*(2), 425-431. doi:[10.1017/S1366728915000553](https://doi.org/10.1017/S1366728915000553)
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-Third Edition*. San Antonio, TX: Psychological Corporation.
- Wilkinson, G. S. (2003). *The Wide Range Achievement Test*. Wilmington DE: Wide Range, Inc.

MATH PROBLEM-SOLVING AND COGNITION

Woodcock, R. W. (1998). *Woodcock Reading Mastery Test- Revised* (Form G). Circle Pines, MN: American Guidance Service, Inc.

Woodcock, R. W., Muñoz-Sandoval, A. F. & Alverado, C. G. (2005). *Woodcock-Muñoz Language Survey*. Itasca, IL: Riverside Publishing.

Table 1

Descriptive Information on Norm-reference Measures for Total Sample and Children with and Without MD.

Variable	N	<u>Total Sample</u>		<u>Children at Risk</u>		<u>Not at Risk</u>		N	M	SD
		M	SD	N	M	SD				
Age (Mos)	394	92.46	13.43	108	95.07	13.03		286	91.47	13.47
Fluid Intell	360	101.09	16.63	97	94.35	15.2		263	103.58	16.47
English										
Problem-solving & Calculation										
WJ-AP	380	96.77	13.67	102	85.70	13.15		278	100.83	11.45
WISC-III	382	92.33	14.91	103	77.52	15.42		279	97.8	10.31
WJ-Cal	382	101.55	18.00	103	98.29	19.3		276	106.4	10.94
WRAT	381	106.85	10.99	103	101.32	11.33		274	95.18	16.91
Reading										
WordID	379	103.6	16.73	100	92.28	15.93		279	102.75	17.38
E-Comp	379	92.98	17.23	101	82.38	16.62		278	108.9	10.14
Vocabulary										
E-ppvt	375	90.94	17.39	102	84.18	17.16		280	98.88	11.18
E-expressive	379	98.85	24.74	102	89.29	18.62		270	108.92	10.29
Spanish										
Problem-solving & Calculation										
BWJ-AP	383	101.42	13.95	107	88.58	12.65		273	93.46	16.82
WISC-III	381	88.71	19.42	107	72.15	15.19		277	102.36	25.79
BWJ-Cal	383	96.21	12.65	103	88.93	13.59		275	95.82	14.64
WRAT	367	107.18	10.96	97	76.10	17.55		277	102.31	11.32
Reading										
Word-ID	383	117.98	17.23	107	107.43	16.51		279	107.66	15.09
S-Comp	378	91.81	16.39	106	81.79	16.96		278	96.83	15.79
Vocabulary										
S-ppvt	382	92.73	15.46	107	84.78	14.72		276	122.07	15.72
S-expressive	379	74.93	17.06	102	71.78	15.30		272	95.72	14.41

Note. E-English; S-Spanish; Fluid intelligence=Raven Colored Progressive Test; WJ-=Woodcock-Johnson; AP=applied problem subtest; WISC-III=Wechsler Intelligence Scale for Children, Arithmetic subtest; BWJ=Bateria; WJ Cal=calculation subtest; Word ID = Letter Word identification subtest from Woodcock-Muñoz Language Survey-Revised; WRAT= computation subtest from Wide Range Achievement Test; E-Comp = English Passage Comprehension subtest from Woodcock-Muñoz Language Survey-Revised; Word attack = Word attack subtest from

MATH PROBLEM-SOLVING AND COGNITION

Woodcock-Muñoz Language Survey-Revised; PPVT = Peabody Picture Vocabulary Test in English; E-Expressive= Expressive One-Word Picture Vocabulary Test-English-Bilingual Edition; S-ppvvt = Test de Vocabulario en Imágenes; S-Comp = Spanish Passage Comprehension subtest from Woodcock-Muñoz Language Survey-Revised; S-Expressive= Expressive One-Word Picture Vocabulary Test -Spanish-Bilingual Edition (EOWPVT-SBE: Brownell, 2001).

MATH PROBLEM-SOLVING AND COGNITION

Table 2 *Comparison of Children With and Without MD on Latent Measures of Achievement, Vocabulary and Cognition.*

	N	<u>At risk MD</u>		<u>Not at Risk</u>			<u>F-ratio</u>
		Mean	SD	N	Mean	SD	
Age(mos)	108	106.30	13.83	286	102.67	14.15	5.21
E-Vocab	102	-0.49	1.53	271	0.18	1.60	13.46***
S-Vocab	102	-0.31	1.33	273	0.12	1.19	9.19
S-Read	100	-0.78	1.99	271	0.31	1.84	24.75***
E-Read	99	-1.00	2.02	274	0.36	2.31	26.52***
FIQ	97	-0.51	1.87	262	0.19	1.82	10.27
E-cal	102	-0.35	1.56	276	0.13	1.54	7.12
S-cal	96	-0.25	1.69	269	0.08	1.61	2.80
E-Num	99	-0.38	1.85	266	0.1	1.81	4.83
S-Num	103	-0.78	1.57	271	0.31	1.88	27.27***
E-Estim	108	0.44	1.6	286	-0.17	1.55	11.98***
S-Estim	108	0.29	1.58	286	-0.11	1.51	5.19
E-STM	106	-0.83	1.48	272	0.32	1.59	41.81***
S-STM	102	-0.86	1.66	273	0.36	1.56	43.82***
E-Speed	104	0.42	2.05	275	-0.17	1.32	10.54
S-Speed	99	0.18	1.65	271	-0.08	1.45	2.08
E-Inhib	99	-0.22	0.92	268	0.1	0.98	7.96
S-Inhib	97	-0.12	0.69	263	0.05	0.71	4.52
E-WM	98	-0.59	0.98	271	0.25	1.50	26.38***
S-WM	105	-0.87	1.25	271	0.35	1.62	48.38***
Vis-WM	108	-0.36	1.09	286	0.14	1.16	14.60***

Alpha set at ***p < .002, Note. E-English, S-Spanish Vocab=vocabulary, Read=Reading, FIQ=fluid intelligence, Cal=calculation, num=magnitude number comparison, Estim=estimation, , STM=short-term memory or phonological loop, speed=naming speed, inhib=inhibition, WM=executive component of Working memory, Vis-WM=visual-spatial –visual-spatial sketch pad.

MATH PROBLEM-SOLVING AND COGNITION

Table 3

Multilevel Logistic Models Predicting MD Status

	<u>Unconditional</u>		<u>Model 1</u>		<u>Model 2</u>		<u>Model 3</u>		<u>Model 4</u>	
Fixed Effects										
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	1.11***	0.21	1.49	0.22	1.89***	0.28	2.05***	0.28	2.39***	0.36
GRADE			-0.82	0.27	-2.26***	0.40	-3.54***	0.58	-4.48***	0.76
Bilingual			0.12**	0.02	0.08**	0.02	0.08**	0.03	0.06*	0.03
E-Read									0.05	0.22
S-Read									0.46*	0.22
FIQ									0.10	0.16
E-Cal							0.52	0.28	0.37	0.33
S-Cal							0.12	0.25	0.20	0.28
E-Num							-0.54*	0.23	-0.50	0.27
S-Num							0.76**	0.24	0.79**	0.26
E-Estim							0.0009	0.18	-0.07	0.21
S-Estim							-0.41*	0.19	-0.35	0.21
E-STM					0.30*	0.16	0.19	0.17	0.09	0.20
S-STM					0.31*	0.14	0.35*	0.17	0.26	0.18
E-Speed					-0.16	0.18	-0.07	0.20	-0.06	0.23
S-Speed					-0.18	0.17	-0.007	0.19	0.17	0.22
E-Inhib					0.28	0.21	0.20	0.25	0.12	0.28
S-Inhib					0.19	0.30	0.16	0.36	0.48	0.43
E-WM					0.43*	0.21	0.16	0.22	0.03	0.26
S-WM					0.51**	0.15	0.77***	0.20	1.09***	0.27
Vis-WM					0.03	0.19	0.01	0.22	0.32	0.27
Error Variance			Variance	SE	Variance	SE	Variance	SE	Variance	SE
Level 2	0.68*	0.34	0.63**	0.26	0.35	0.37	0		0	
Deviance		444.02		364.05	219.87		160.42		132.58	
AIC		448.02		372.05	245.87		196.42		174.58	

Note. %reduction, >^a = estimates larger in full model than model without bilingual proficiency, E-English, S-Spanish, Word problems=classification variable, SE= standard error, * $p < .05$, ** $p < .01$, *** $p < .001$, Deviance= Chi-square value for the correspondence between model and data, AIC= Akaike's Information Criterion.

E-English, S-Spanish, Read=Reading, FIQ=fluid intelligence, Cal=calculation, num=magnitude number comparison, Estim=estimation, , STM=short-term memory or phonological loop, speed=naming speed, inhib=inhibition, WM=executive component of working memory, Vis-WM=visual-spatial –visual-spatial sketch pad.

MATH PROBLEM-SOLVING AND COGNITION

Appendix A

Descriptive Information on Raw Scores, Reliability, and Standardized Estimates.

	<u>Mean</u>	<u>SD</u>	<u>Kurtosis</u>	<u>Skewness</u>	<u>KR₂₀</u>	<u>Estimates</u>	<u>SE</u>	<u>t-ratio</u>
Math Problem-solving								
English								
WJ-AP	21.31	6.13	1.79	0.43	0.78	0.77	0.04	20.51**
WISC-III	10.87	3.42	1.14	-0.97	0.67	0.70	0.04	16.74**
Spanish								
WJ-AP	23.76	6.65	0.61	-0.52	0.83	0.86	0.03	34.36**
WISC-III	10.05	4.55	-0.47	-0.62	0.83	0.78	0.03	24.67**
Math Calculation								
English								
WJ-Cal	15.32	4.91	1.11	-0.55	0.78	0.83	0.03	32.25**
WRAT	20.61	3.01	0.52	0.53	0.42	0.87	0.02	39.11**
Spanish								
WJ-Cal	9.29	3.38	0.29	-0.45	0.71	0.85	0.02	37.50**
WRAT	20.81	2.91	-0.2	0.36	0.44	0.89	0.02	45.16**
Vocabulary								
English								
Receptive	103.89	29.09	-0.29	-0.14	0.96	0.84	0.03	31.89**
Expressive	49.46	19.36	0.1	-0.08	0.95	0.87	0.02	35.08**
Spanish								
Receptive	50.57	15.64	1.03	-0.07	0.92	0.85	0.04	21.45**
Expressive	30.80	15.31	-0.77	-0.18	0.96	0.56	0.05	10.58**
Reading								
English								
Word-ID	32.45	14.61	-1.07	0.18	0.95	0.96	0.01	86.83**
Compreh.	11.51	5.98	-0.97	-0.04	0.90	0.89	0.02	50.62**
Word-Att	12.48	10.82	-0.52	0.67	0.96	0.7	0.04	18.88**
Spanish								
Word-ID	35.34	10.58	0.41	0.2	0.89	0.81	0.03	29.62**
Compreh.	10.72	5.03	3.29	0.68	0.80	0.79	0.03	26.22**
Word-Att	21.62	11.31	-1.03	-0.57	0.95	0.69	0.04	17.83**
Fluid Intelligence								
Eng/Spain								
RAV_A	8.29	2.05	3.26	-1.27	0.39	0.65	0.05	13.05**
RAV__AB	7.45	2.94	-0.25	-0.62	0.73	0.77	0.04	18.79**
RAV_B	5.88	3.01	-0.5	0.04	0.79	0.79	0.04	19.57**
Numeracy								
English								
PS-Components	11.67	4.23	0.03	-0.58	0.80	0.51	0.05	9.39**
Num-L	7.52	4.81	-0.13	0.26	0.87	0.88	0.02	39.76**
Num-S	5.78	4.36	-0.09	0.58	0.87	0.83	0.03	32.47**

Spanish								
PS-Components	10.73	4.19	0.08	-0.58	0.81	0.51	0.05	9.26**
Num-L	7.72	4.66	-0.14	0.15	0.86	0.87	0.02	37.07**
Num-S	5.74	4.07	0.48	0.65	0.85	0.83	0.03	31.49**
Estimation								
English								
Set 1-	39.36	29.06	-1.32	0.39	0.99	0.84	0.02	34.70**
Set 2-	21.43	18.78	-0.55	0.96	0.98	0.85	0.02	34.95**
Spanish								
Set 1-	40.39	26.29	-1.37	0.19	0.98	0.77	0.04	21.24**
Set 2-	20.69	16.24	-0.85	0.74	0.98	0.89	0.03	30.12**
STM-Phonological Loop								
English								
Non words	5.64	2.98	-0.03	0.26	0.76	0.51	0.06	8.68**
Real Words	9.82	3.79	-0.11	0.22	0.77	0.67	0.05	13.97**
Digit-Forward	6.46	1.97	0.73	0.55	0.42	0.67	0.05	13.93**
Digit-Backward	2.92	1.42	0.14	0.14	0.47	0.42	0.06	6.56**
Spanish								
Non words	4.87	2.91	0.21	0.38	0.79	0.54	0.06	9.49**
Real Words	7.42	3.57	-0.68	-0.22	0.80	0.75	0.04	17.57**
Digit-Forward	6.10	1.76	2.23	-0.86	0.30	0.56	0.06	10.02**
Digit-Backward	2.86	1.5	2.01	0.47	0.41	0.51	0.06	8.75**
Naming Speed								
English								
Letters	59.41	24.61	5.29	1.91	0.96	0.94	0.02	40.65**
Digits	55.28	22.85	8.97	2.41	0.95	0.77	0.03	22.92**
Spanish								
Letters	69.98	25.94	4.41	1.59	0.96	0.79	0.03	23.58**
Digits	55.13	18.62	8.52	2.25	0.94	0.87	0.03	30.29**
WM-Executive Processing								
English								
Concept-Span	3.59	3.42	2.71	1.63	0.84	0.43	0.06	6.92**
Sentence-Span	2.42	2.96	3	1.63	0.85	0.62	0.05	11.88**
Listen-Span	1.21	1.38	5.23	1.67	0.52	0.50	0.06	8.41**
Update	3.08	2.27	1.35	0.96	0.80	0.50	0.06	8.52**
Spanish								
Concept-Span	3.29	2.86	1.11	1.17	0.83	0.64	0.05	13.00**
Sentence-Span	2.84	3.31	2.59	1.6	0.86	0.63	0.05	12.7**
Listen-Span	0.92	1.06	1.16	1.04	0.52	0.52	0.06	9.21**
Update	3.22	2.21	2.27	1.13	0.70	0.52	0.06	9.23**
Inhibition								
English								
Letters	6.09	3.25	-0.48	0.05	0.80	0.51	0.07	7.40**
Numbers	3.43	2.63	0.05	0.58	0.77	0.66	0.07	9.33**
Spanish								

MATH PROBLEM-SOLVING AND COGNITION

Letters	3.52	2.72	-0.48	0.49	0.81	0.47	0.08	5.80**
Numbers	5.8	3.4	-0.4	0.2	0.82	0.43	0.08	5.45**
Visual-Spatial Sketchpad								
Eng/Span								
Matrix	9.11	7.77	-0.33	0.66	0.95	0.74	0.07	11.17**
Mapping	3.29	3.11	7.21	2.30	0.80	0.63	0.06	9.68**

*Note. Eng/Span=administer in both languages. ** $ps < .01$

Fluid intelligence=three parts of the Raven Colored Progressive Test, WJ=Woodcock-Johnson; AP=applied problem subtest; WISC-III=Wechsler Intelligence Scale for Children, Arithmetic subtest; WJ Cal=calculation subtest; WRAT= computation subtest from Wide Range Achievement Test; E-Receptive = Peabody Picture Vocabulary Test in English; S-Receptive = Test de Vocabulario en Imagenes; Expressive= Expressive One-Word Picture Vocabulary Test - Spanish-Bilingual Edition (EOWPVT-SBE: Brownell, 2001), Word ID = Letter Word identification subtest from Woodcock-Muñoz Language Survey-Revised; Word-att= Word attack subtest from Woodcock Reading Mastery Test; PS-Components=Problem-solving components, Num=number comparisons-large vs. small, L=letters, N=numbers.

Footnotes

¹The literature is unclear as to what terms appropriately capture our sample (e.g., English language learners, English learners, limited English-proficient, balance vs. unbalanced bilingual, emerging bilinguals). We used the term English learner (EL) to align with the literature, but realize the sample is best described as emerging bilinguals to emphasize children's strengths as well as their language proficiency. English learner status in this study was determined by a state and district-wide test that assesses ELs in the following areas: listening, speaking, reading and writing.

²One framework to capture executive processing in emerging bilingual children is Baddeley's multicomponent WM model (Baddeley & Logie, 1999). The multicomponent model characterizes WM as comprising a central executive controlling system that interacts with a set of two subsidiary storage systems: the speech-based phonological loop and the visual-spatial sketchpad. According to Baddeley (Baddeley, 2012; Baddeley & Logie, 1999), the central executive coordinates the two systems, focusing and switching attention, and activating representations within long-term memory (LTM). The three component model has been revised to include an episodic buffer (Baddeley, 2012), but support for the tripartite model has been found across various age groups of children (Gathercole Pickering, Ambridge & Wearing, 2004; Gray et al., 2017). Thus, the current study will focus on the three components of WM consistent with Baddeley's earlier model.

