The Role of Mathematical Language Skills in Arithmetic Fluency and Word-problem Solving for First- and Second-Language Learners

Chang Xu[1] *

Anne Lafay[2] *

Heather Douglas[3]

Sabrina Di Lonardo Burr[3]

Jo-Anne LeFevre[1,3]

Helena P. Osana[2]

Sheri-Lynn Skwarchuk[4]

Judith Wylie[5]

Victoria Simms[6]

Erin A. Maloney[7]

* Co-first authors

[1] Department of Psychology, Carleton University

[2] Department of Education, Concordia University

[3] Department of Cognitive Science, Carleton University

[4] Faculty of Education, University of Winnipeg

[5] School of Psychology, Queen's University Belfast

[6] School of Psychology, Ulster University

[7] School of Psychology, University of Ottawa

Corresponding Author: Chang Xu

Email Address: ChangXu@cmail.carleton.ca Department of Psychology

Carleton University

1125 Colonel By Drive, Ottawa, K1S 5B6

Funding

Support for project was provided by the Social Sciences and Humanities Research Council (SSHRC) of Canada through an Insight Grant to J. LeFevre, E. Maloney, H. Osana, and S. Skwarchuk.

Acknowledgements

Thanks go to Renee Whittaker, Jill Turner, Stephanie Hadden, Sarah Macintosh, Marion Valat, and Sophie Lemieux for assistance in data collection and data entry.

This manuscript was accepted for publication in the *Journal of Educational Psychology* on February 6, 2021. This preprint is the peer-reviewed accepted version but has not yet been copyedited and may differ from the final version published in the journal.

Abstract

Language skills play an important role in mathematics development. Students (7 to 10 years of age) learning school mathematics either in the same language used at home (first-language learners; n = 103) or in a different language (second-language learners; n = 57) participated in the study. Relations among cognitive skills (i.e., receptive vocabulary, working memory, quantitative skills), domain-specific language skills (i.e., mathematical vocabulary, mathematical orthography), word-problem solving, arithmetic fluency, and word reading were investigated. Second-language learners had lower scores on measures with strong language components (i.e., receptive vocabulary, subitizing, and word-problem solving) than first-language learners, whereas they performed equally well on other tasks. Mathematical vocabulary and receptive vocabulary contributed to word-problem solving success for first-language learners, whereas only receptive vocabulary in the language of instruction related to mathematical outcomes for second-language learners. Mathematical vocabulary was related to arithmetic fluency for both groups, but mathematical orthography was not. For both groups, students' word reading was predicted by receptive vocabulary, but not by quantitative skills, highlighting the domain-specific nature of these skills. These findings have implications for supporting mathematical learning in second-language students.

Word count abstract: 180

Keywords: Language and mathematics; second-language learners; first-language learners; mathematical vocabulary; mathematical orthography.

Educational Impact and Implications Statement [30-70 words]

How do language skills affect mathematics for second-language learners? Vocabulary skills in the instructional language were more strongly related to word-problem solving and arithmetic skills for second- than first-language learners. Teachers of second-language learners may need to adapt materials or reduce language demands so that all students have equal opportunity to develop competent mathematical skills. For all students, explicit instruction on mathematical vocabulary and mathematical orthography may support their mathematical competency.

The Role of Mathematical Language Skills in Arithmetic Fluency and Word-Problem Solving for First- and Second-Language Learners

Students and adults use numbers and mathematics daily. They count sets of objects, name and write numbers, and participate in activities using time, dates, distance, and money. Mathematical skills predict school success for young learners (Duncan et al., 2007) and are strongly associated with adults' socio-economic level (Ritchie & Bates, 2013) and quality of life (Bynner & Parsons, 1997; Gerardi et al., 2010; Kutner et al., 2007; Lusardi, 2012). Because numbers are ubiquitous, and early mathematical skills predict later success, it is important to understand how students learn about mathematics. Moreover, in many countries, students are taught school mathematics in a different language than the one they speak at home. Thus, the goal of the present paper was to investigate the relations between language of instruction and mathematical learning. In particular, we explored the role of domain-specific language skills (i.e., mathematical vocabulary and mathematical orthography) in predicting individual differences in mathematical performance for first- and second-language learners.

Language and Mathematics

Language and mathematics are related in many ways (Cirino, 2011; Gjicali et al., 2019; Hooper et al., 2010; Kleemans et al., 2011). Specifically for children in kindergarten and early grades of elementary school (i.e., grades 1 through 3), receptive vocabulary and phonological awareness were related to a range of mathematical outcomes, including numeration knowledge, calculation, number line estimation, and geometry (LeFevre et al., 2010; Purpura & Ganley, 2014; Träff et al., 2017; Vukovic & Lesaux, 2013). LeFevre et al. situated their findings within a more general model of early mathematical development, called the "Pathways to Mathematics" model. This model described three categories of cognitive skills that contribute to mathematical performance: language, quantitative knowledge, and attention or working memory (LeFevre et al., 2010; Sowinski et al., 2015). In this model, language consistently predicted individual differences in performance for the majority of mathematical outcomes. Additional evidence of the relation between language and mathematics comes from students learning mathematics in their second language. In the present study, we defined first-language learners as students whose language of instruction at school was the same as the primary language they used at home and second-language learners as students whose language of instruction at school was different from the primary language they used at home.

Learning mathematics is more challenging for second- than for first-language learners (see a review in de Araujo et al., 2018). For example, Bonifacci et al. (2016) found that first-language learners in preschool had better early numeracy skills (i.e., number line estimation, size seriation, knowledge of digits) than second-language learners. Kleemans et al. (2011) found that 6- and 7-year-old second-language learners scored lower than first-language learners on early numeracy tasks, such as counting. Thus, performance on early numeracy tasks is related to students' language status.

Mathematical tasks that require a high degree of language processing, such as word-problem solving, may be more strongly related to students' language skills than mathematical tasks that require less language processing. To solve mathematical word problems, students need to have adequate knowledge of syntax and vocabulary so that they can interpret the word problem and produce an appropriate strategy (Daroczy et al., 2015; Fuchs et al., 2015). Consistent with this view, second-language learners in grades 2 and 3 (ages 7 to 10) scored lower than first-language learners on mathematical word-problem solving; the difference was larger for word problems that used passive language and those that included a conditional clause (Banks et al., 2016). In other studies, second-language learners experienced challenges when completing tasks that require knowledge of complex academic vocabulary (e.g., *therefore*, *based on*, and *substantial;* Abedi & Lord, 2001; Wolf & Leon, 2009). Moreover, students' vocabulary skill is related to their performance on mathematical word problems (e.g., Fuchs et al., 2008; LeFevre et al., 2010; Méndez et al., 2019). Finally, in some studies, second-language learners performed better when word problems were in the language spoken by the student at home (Ambrose & Molina, 2014; Bernardo, 2002; Telli et al., 2018). In summary, performance on the word problems was related to students' proficiency in the language of instruction and to their home-language skills.

Beyond word problems, which have explicit linguistic components, language proficiency may also be related to students' arithmetic skills. In some studies, participants were faster and more accurate when they were asked to verbally respond to arithmetic problems in their more proficient language than in their less proficient language (Geary et al., 1993; Marsh & Maki, 1976; McClain & Huang, 1982; Van Rinsveld et al., 2015). In Van Rinsveld et al. (2015), second-language learners in grades 7 to 11 (ages 12 to 22) verbally responded to arithmetic problems. For simple problems that presumably involved direct retrieval (i.e., sums < 10), students' performance was no different in their first versus second language. In contrast, for complex problems with two-digit operands (e.g., 56 + 23), students performed better when they responded in their first language than in their second language (Van Rinsveld et al., 2015). Furthermore, results from patients with calculation disorders support the view that arithmetic involves some level of language processing (Cohen et al., 2000; Dehaene & Cohen, 1991). In this view, language may be specifically involved when arithmetic problems are solved by activation of verbal associations (Cohen et al., 2000). Such findings may indicate that language-based processes are involved even when the mathematical problem is purely symbolic.

Domain-specific Language Skills

Beyond general language skills, language skills specific to mathematics are also related to mathematical learning (e.g., Forsyth & Powell, 2017; Powell et al., 2017; Purpura & Reid, 2016). In the present research, we considered two different aspects of domain-specific mathematical language: vocabulary and orthography.

Mathematical Vocabulary

Mathematical vocabulary comprises words or phrases that express mathematical concepts or procedures (e.g., *more than*, *take away*, *fewest*; Hebert & Powell, 2016). Mathematical vocabulary starts to become important very early in mathematics learning. In preschool and kindergarten (i.e., ages 3 to 5), students' knowledge of mathematical vocabulary predicts their performance on mathematical outcomes (e.g., counting, number identification, and number comparison (Purpura & Reid, 2016). Moreover, mathematical vocabulary mediated the relation between general language skills (e.g., print knowledge, vocabulary, and phonological awareness) and later digit knowledge in students aged 3- to 5-years of age (Purpura et al., 2017; Toll & Van Luit, 2014). For older students, mathematical vocabulary is related to some mathematical outcomes. For Chinese students in grade 4 (mean age 10.2 years), mathematical vocabulary was significantly related to word-problem solving, but not to arithmetic calculation (Peng & Lin, 2019). For American children in grades 3 and grade 5 (ages 8 to 11 years), mathematical vocabulary was correlated with arithmetic, however, word-problem solving was not assessed (Powell et al., 2017). In summary, relations between mathematical vocabulary and mathematical performance are consistently observed, although the effects may vary by students' age and the types of mathematical outcomes assessed.

Although the importance of mathematical vocabulary is well documented, few researchers have investigated mathematical vocabulary in second-language learners and the results of the research are not consistent. Even for first-language learners, mathematical vocabulary is challenging because it involves terms with technical definitions (e.g., *sum*, *value*, *product*), multiple terms can have the same meaning (e.g., subtraction can be described as many different ways, such as *minus*, *less*, *decrease*, *take away*), and homonyms and similar sounding words can be confusing (e.g., *whole/hole*, *half/have*; Roberts & Truxaw, 2013). Because of these complexities and the use of mathematics terms in non-mathematics context, mathematical vocabulary is often less familiar for second- than first-language learners. Accordingly, Powell, Berry, et al. (2020) found that second-language learners had significantly weaker knowledge of mathematical vocabulary than first-language learners in grade 3. However, word-problem solving performance did not differ between first- and second-language learners for third graders with mathematics difficulties (Powell, Urrutia, et al., 2020). The limited number of studies comparing mathematical vocabulary of first- and second-language learners highlights a need for further research.

Mathematical Orthography

Mathematical orthography is knowledge of mathematical symbols and the rules and conventions for writing mathematical expressions (Douglas et al., 2020; Headley, 2016; O'Halloran, 2005; Rubenstein & Thompson, 2001; Thompson & Rubenstein, 2014). For example, mathematical orthography includes the information that the addition symbol is placed between addends (e.g., 3 + 1 vs. + 3 1). Orthographic knowledge of *written* text entails knowledge of letter symbols and the conventions for combining those symbols into words and phrases (Apel, 2011) and is related to individual differences in literacy (Castles et al., 2018). Analogously, knowledge of mathematical orthography is related to individual differences in mathematical skills (Douglas et al., 2020; Headley, 2016). In summary, knowledge of mathematical symbols and the conventional patterns for arranging those symbols is critical to achievement in mathematics (O'Halloran, 2005, 2015; Rubenstein & Thompson, 2001).

Support for a link between specific aspects of mathematical orthography and mathematics learning comes from research on the equal sign and the negative sign. Students in elementary and middle school often assume the equal sign means "put the answer here" (Cobb, 1987; Crooks & Alibali, 2014; Rittle-Johnson & Alibali, 1999). Thus, students struggle to solve problems with numbers on both sides of the equal sign, such as 3 + 5 = 4 + ___ (Alibali, 1999; McNeil & Alibali, 2000; Sherman & Bisanz, 2009) and may judge equations, such as 3 + 5 = 2 + 6, as incorrect or meaningless (Carpenter et al., 2003; Li, 2008; Steinberg et al., 1991). Importantly, students' skills at correctly identifying sides of an equation, noticing the location of the equal sign, and the order of the mathematical symbols (i.e., encoding the equation; Alibali et al., 2018; McNeil et al., 2011) are related to their skills at solving equations.

Middle-school students find it challenging to solve equations that involve multiple minus/negative symbols (e.g., 5 – (-1) = 6) and equations where negative signs precede variables (e.g., solving $12 – x = 7$ by incorrectly writing $x = 7 – 12$; Herscovics & Linchevski, 1994; Vlassis, 2008). Because these students demonstrated a conceptual understanding of the negative sign, their mistakes were attributed to the conventions for writing negative signs. Based on the findings that knowledge of the conventions (i.e., the orthography) for notating the equal sign and the negative sign relate to solving equations, knowledge of mathematical orthography should be related to students' calculation skills (i.e., solving equations).

Knowledge of mathematical orthography is also related to performance on mathematical word problems. Headley (2016) found that grade 8 students' ability to discriminate between conventional (e.g., x^2) and non-conventional (e.g., ${}^2x$) mathematical expressions was related to their performance on a standardized measure of mathematical skills that included word-problem solving. Headley suggested that orthography skills tap into a broad understanding of symbolic mathematical language that influences all aspects of mathematical skills. Consistent with this view, using the same measure of mathematical orthography with adults, Douglas et al. (2020) found direct relations between mathematical orthography and both calculation and word-problem solving performance. In summary, mathematical orthography is an important mathematical language skill but its relations with other aspects of mathematical learning are not well understood (Headley, 2016). The current study is the first to explore if the relations between mathematical orthography and mathematics outcomes differ between first- and second-language learners.

The Current Study

The goal of the current study was to investigate the role of domain-specific language skills (i.e., mathematical vocabulary and mathematical orthography) in predicting individual differences in mathematics performance for first- and second-language learners. Accordingly, we had two key research questions:

Research Question 1: Are there performance differences between first- and second-language learners?

Research Question 2: Do domain-specific language skills relate to mathematics performance differently for first- and second-language learners? Extending this question beyond mathematics, we explored whether domain-specific language skills relate to a non-mathematical outcome (i.e., word reading) differently for first- and second-language learners.

Prior to any modeling, we needed to confirm that first- and second-language learners did not differ with respect to nonverbal intelligence. Thus, the Matrix Reasoning task from the Weschler Intelligence Scale for Children was administered as a proxy measure for nonverbal intelligence (Wechsler, 2014). In accordance with the Pathways Model (LeFevre et al., 2010), students also completed measures of cognitive skills that have known relations with mathematics performance. More specifically, students completed measures of receptive vocabulary, working memory (i.e., digit forward span, digit backward span, and spatial span), and quantitative skills (i.e., subitizing and number comparison). The working memory measures were selected based on findings described in previous meta-analyses (Friso-Van Den Bos et al., 2013; Peng et al., 2016). The included quantitative skills have also been used in other studies (e.g., Cirino, 2011; Bugden et al., 2020; LeFevre et al., 2010; Sowinski et al., 2014). Although both subitizing and number comparison are quantitative skills, we note that subitizing, which involves naming small quantities, has higher verbal demands because students need to orally provide the correct response (i.e., 'one', 'two', 'three').

Students also completed two measures of domain-specific language skills (i.e., mathematical vocabulary and mathematical orthography) and three outcome measures (i.e., arithmetic fluency, word-problem solving, and word reading). These outcomes were chosen because of the varying level of language processing involved: Word-problem solving requires a high degree of language processing; arithmetic fluency is less language-demanding, nevertheless may involve verbal processes; and word reading only requires language skills, not mathematical skills. Word reading was included to test discriminant validity, verifying that domain-specific (i.e., mathematical) language skills were not related to reading for either group of learners.

The analyses were framed around two research questions and five corresponding hypotheses.

Research Question 1

Are there performance differences between first- and second-language learners?

Hypothesis 1. Based on the assumptions that second-language learners are less proficient in the language of instruction than first-language learners and that language proficiency is related to language-rich outcomes such as word-problem solving (Banks et al., 2016; Fuchs et al., 2016), we hypothesized that first-language learners would perform better than second-language learners on tasks that have a strong language component (i.e., subitizing, receptive vocabulary, mathematical vocabulary, and word-problem solving).

Hypothesis 2. Based on the assumption that the Pathways to Mathematics model is equally applicable to both first- and second-language learners, we hypothesized that, for both groups, word-problem solving (Hypothesis 2a) and arithmetic fluency (Hypothesis 2b) would be predicted by cognitive skills (i.e., receptive vocabulary, working memory, and quantitative skills; LeFevre et al., 2010; Sowinski et al., 2015).

Research Question 2

Do domain-specific language skills relate to mathematics performance and/or word-reading performance differently for first- and second-language learners?

Hypothesis 3. Based on the assumption that learners' developing knowledge becomes more domain specific as they gain mastery in the domain, we hypothesized that domain-specific language skills (i.e., mathematical vocabulary and mathematical orthography) would mediate the relation between general language skills (i.e., receptive vocabulary) and word-problem solving (Powell et al., 2017).

Hypothesis 4. Because mathematical vocabulary (Powell et al., 2015, 2017) and mathematical orthography (Douglas et al., 2020) are related to arithmetic fluency in first-language learners, we predicted that for both groups of learners, mathematical vocabulary and orthography would be related to arithmetic fluency.

Hypothesis 5. In support of domain-specificity, for both groups, we predicted that students' word reading would be predicted by receptive vocabulary but not by quantitative skills (Sowinski et al., 2015).

Method

Participants

Following ethics approval from Carleton University and Concordia University, school principals were contacted. One-hundred and sixty students in second- and third grade were recruited from nine public schools in or near (three) large multi-cultural Canadian cities. Two different provinces were selected for data collection because of the first-language differences between these provinces. By combining datasets from two large and multicultural locations with different first languages, we were able to more broadly compare mathematics performance between first- and second-language learners without limiting any effects to a single first language. In Quebec, students received mathematics instruction in French and in Ontario, students received mathematics instruction in English. These groups are described separately below.

French Language Students

Eighty-one second-grade students (M = 7:11, SD = 0:5; range = 7:4 to 10:0; 58% female) were recruited from six French-language schools in Quebec. Notably, age was not reported by two students. Parents were asked what language their child spoke at home (i.e., *Langue parlée à la maison?*). For 50 students, parents reported that their child spoke French at home, whereas for the other 31 students, parents reported that their child spoke another language at home (i.e., English: 18.3%; Spanish: 6.1%; Arabic: 3.1%; Creole: 2.4%; Chinese: 1.2%; Polish: 1.2%; Persian: 1.2%; Wolof: 1.2%; Swahili: 1.2%). According to the "Indice de milieu socio-économique," a school-level index of socioeconomic status (http://www.education.gouv.qc.ca/references/indicateurs-et-statistiques/indices-de-defavorisation/), 43% of students were from low-SES schools, 36% were from middle-SES schools, and 21% were from high SES schools. There were more second-language learners in the lower- than in the higher-SES schools, $\chi^2(2, N = 81) = 19.65$, $p < .001$. More generally, students in the first- and second-language groups were equally distributed across schools, $\chi^2(5, N = 81) = 9.92$, $p = .078$.

English Language Students

The other 79 students were recruited from three English-language schools in the province of Ontario. Of these, 21 were in second grade (*M* = 7:7, *SD* = 0:3, range = 7:2 to 8:1; 57% female) and 58 were in third grade (*M* = 8:7, *SD* = 0:5; range = 8:2 to 9:11; 57% female). These schools were selected because they had a higher-than-average (within the school district) proportion of second-language learners. All students were receiving mathematics instruction in English. Parents were asked "What is the primary language of your child?" Within this group, 53 students spoke English as the primary language, whereas 26 students spoke another language (i.e., Arabic: 7.6%; Burmese/Lai: 5.1%; Spanish: 3.8%; Portuguese: 2.5%; Tagalog: 2.5%; Aramaic: 1.3%; Gujarati: 1.3%; Creole: 1.3%; French: 1.3%; Hindi: 1.3%; Inuktitut: 1.3%; Kinyarwanda: 1.3%; Swahili: 1.3%; Vietnamese: 1.3%). Sixty-eight of the responding parents provided their highest level of education (i.e., less than high school, high school, technical/applied college, university, or postgraduate). The median level was technical/applied college for both first- and second-language learners and did not differ across language groups, $\chi^2(4, N = 68) = 3.18$, $p = .528$. Based on provincial reports of family socioeconomic status by school (https://www.app.edu.gov.on.ca/eng/sift/index.asp), participating schools had a higher-than-average proportion of low-income households. Students in the first- and second-language groups were not equally distributed across schools, $\chi^2(2, N = 79) = 6.71$, $p = .035$, however, this difference in distribution was due to one particular school where more first-language learners were recruited than second-language learners.

To get a sense of the language experience of the students, parents in Ontario were also asked about second and third languages, and, for each reported language, they were asked to indicate how frequently the language was spoken at home by their child, by them, and by the other parent. Due to an oversight, parents in Quebec were not asked these additional questions. Parents responded on a five-point scale to each question, coded as 1 = occasionally, 2 = sometimes, 3 = about half the time, 4 = often, and 5 = always. These ratings were used to calculate a language exposure index for each reported language by summing the ratings across the three questions, and thus could range from 3 to 15. Of the 53 students in Ontario for whom English was reported as their primary language, 24 also spoke a second language at home. For these students, language exposure at home was significantly higher for English (M = 12.5, SD = 3.2) than for the second language (M = 7.2, SD = 4.0), t(23) = 3.90, p = .001. Thus, we were confident that for these students, their primary language at home was English, consistent with the language of school.

For the 26 students whose first language was not English, 21 reported English as the second language, four parents did not indicate a second language for their child, and one reported that English was their child's third language. Three other students spoke a third language at home, but infrequently. For the 17 students whose second language was English and whose parents provided the information about language exposure at home, exposure to the home language (M = 10.3, SD = 3.4) was not significantly higher than to English (M = 10.0, SD = 2.7), t(16) = .239, p = .814. However, these data may overestimate some students' exposure to English, because some parents in this group did not complete the exposure questions for a second language, suggesting that only the first language was spoken at home.

Exposure to Multiple Languages

As described above, the first-language students in this study were those where the primary language spoken at home was consistent with the language of instruction for mathematics in school, whereas the second-language students were exposed at home equally or more frequently to a language other than the one used in school. However, in Canada, all students study a second language (French in Ontario; English in Quebec), starting in elementary school. Additionally, in Quebec, students may have some exposure to English on television, the internet, or in daily life. Thus, the students in this study were exposed to two or more languages at home or at school, which is consistent with the status of Canada as a bilingual country with a large immigrant population. Our focus was on the extent to which the language of mathematics as taught in school was also used at home.

Provincial Curricula

Education is a provincial responsibility in Canada. Thus, students in Quebec and Ontario follow a different curriculum. The curricula are similar in organization in that both include learning expectations related to numeration, operations, measurement, geometry and spatial sense, probability, and data management. Curricular expectations in both provinces include concepts and procedures that students should learn in each grade. As well, both provincial curricula provide similar guidelines for solving multi-step problems in authentic contexts that invite multiple solution strategies. That is, students are expected to develop, select, apply, and communicate their problem-solving strategies. Moreover, both documents indicate students should “assimilate concepts, processes and strategies” (Éducation Loisir et Sport Québec, 2009, p. 3) and instruction should “help <students> deepen their mathematical understanding” (Ontario Ministry of Education, 2005, p. 65). There are, however, minor differences in the curricula specific to grade level. For example, in the Ontario curriculum, students in third grade are expected to add and subtract 3-digit numbers, whereas in the Quebec curriculum, students in second grade are expected add and subtract 3-digit numbers. Expectations related to fluency with addition facts, however, are the same by grade in each province.

There are also differences in how mathematical vocabulary is organized in the curricula for the two provinces. Both curricula include mathematical words that students must learn, understand, and use to express their mathematical thinking. However, in the Ontario curriculum, the mathematical words that students have to learn are included in a glossary that covers grades 1 to 8. In contrast, the Quebec curriculum specifies the mathematical words that students are expected to learn in each grade. Importantly however, words are common across curricula (e.g., odd number, sum, difference, cube, less than, etc.). To reduce any effect of curricular differences (i.e., expectations by grade and organization of mathematical vocabulary) on mathematical performance, all scores were standardized within grade and within provincial groups.

In Ontario, children who turn five between January 1st and December 31st start kindergarten in September of that year. In Quebec, children who turn five between January 1st and September 30th start kindergarten in September of that year. Those born after September 30th will start kindergarten in September of the following year. Thus, some Ontarian children could be up to three months younger starting a grade level compared to children in Quebec.

Measures

The data were collected as part of a larger project. In addition to the measures used in the current analysis, all of the students completed some additional measures. Three of these measures were not used because they were not directly related to the predictions of the study: An assessment of the language used to count, rapid automatic naming of letters, and number line estimation. Two others, transcoding and number ordering, were not included because of inconsistencies in data collection between English- and French-speaking students.

Nonverbal Intelligence

To confirm that first- and second-language learners did not differ with respect to nonverbal intelligence, students completed the Matrix Reasoning task from the Weschler Intelligence Scale for Children - Fifth Edition (Wechsler, 2014). They were presented with an incomplete grid and they were asked to identify the missing part that properly completes the matrix. There were 34 trials, and testing was discontinued after three consecutive errors. The total score was the number of correct trials. The internal reliability (Cronbach's α) was .83 and .85, for students who learned mathematics in English and French, respectively.

Known Cognitive Predictors of Mathematical Performance

Receptive Vocabulary. An adapted form of the *Peabody Picture Vocabulary Test–Revised*, Form B (PPVT; Dunn & Dunn, 2012) and an adapted form of the *Échelle de vocabulaire en images Peabody*, Form A (EVIP; Dunn et al., 1993) were administered to students in Ontario and Quebec, respectively, to assess students' receptive vocabulary. Students were presented with a card with four pictures. Upon hearing an English word (PPVT) or a French word (EVIP), students were asked to point to the picture that best corresponded to the word they had just heard. For the PPVT, students were asked to complete the 60 items that are appropriate for students in grades 2 and 3. For the EVIP, students were asked to complete the 36 items that are appropriate for students in grade 2. The testing was discontinued when students made eight or more errors in a set. To account for potential language-related differences, for both measures, the dependent variable was the within-language standardized total number of correct items. The internal reliability (Cronbach's α) for the PPVT (among the five subsets) and EVIP (among the three subsets) was .91 and .78, respectively.

Working Memory. Students' working memory was assessed through digit span tasks, forward and backward, and a spatial span task. Digit span tasks assessed verbal short-term and working memory, whereas the spatial span task assessed visual-spatial memory (Alloway et al., 2008).

Digit Span Forward. In the *Digit Span Forward* (WISC-5; Wechsler, 2014), students heard a series of numbers (e.g., 3-8-2) and were asked to repeat the numbers back in order (e.g., 3-8-2). There were two trials for each span. Students started with a span length of two digits. If students correctly repeated all the numbers in the correct order for at least one of the two sequences per span, then the span length increased by one digit. Testing discontinued when students were incorrect on both spans of a given length. The total score was the number of sequences that each student repeated correctly. Based on the subscores of the first and second sequences at each length, reliability was Cronbach's $\alpha = .82$ and .86, for English and French, respectively.

Digit Span Backward. The procedure for the *Digit Span Backward* (WISC-5; Wechsler, 2014) was identical to the *Digit Span Forward*, except that students heard a series of numbers (e.g., 3-8-2) and were asked to repeat the numbers they heard in the reverse order (i.e., 2-8-3). A reliability analysis was calculated based on the subscores of first and second sequences at each length. For both English and French versions of the test, Cronbach's $\alpha = .75$ and .57, respectively. Further examination revealed that the performance distribution for each group was comparable to the norms reported by a larger national survey for children between the ages of 7 and 9 (www.nlsinfo.org), and thus, the lower reliability of the French version is likely not problematic, nor should it influence the pattern of results.

Spatial Span. In this task, nine green dots were presented on an iPad screen (Path Span: https://hume.ca/ix/pathspan/). On each trial, a series of dots lit up one by one, after which students were asked to touch the dots that were lit up in the same order. Students started with a sequence length of two dots. There were three trials for each sequence length. If students correctly reproduced at least one sequence, the sequence length increased by one dot. The task terminated when students made errors on all three sequences for each sequence length. The total score was the number of sequences completed correctly. A reliability analysis was calculated based on the subscores of first, second, and third sequences at each length; Cronbach's α was adequate for students who learned mathematics in English (.76) and good for students who learned mathematics in French (.86).

Quantitative Skills

Subitizing. To assess students' knowledge of the mappings between small quantities and number words (i.e., subitizing), students were asked to name quantities 1, 2, and 3 as quickly as possible. Stimuli and instructions are available at https://carleton.ca/cacr/math-lab/measures/speeded-processing-tasks/. This subitizing task consisted of two pages (8 ½ x 11 in. white paper) with sets of dots. Each page consisted of 24 stimuli presented in four rows of six sets of small dots, for a total of 48 items. The dots in each set were randomly arranged. Students were given practice trials. On the cover page for each task, students were shown a single row of stimuli (six sets of one, two, or three dots) and they were asked to say how many were in each set from left to right as quickly as possible. Feedback was provided if students named the stimuli incorrectly.

After the six practice trials, students completed two full pages of stimuli, one page at a time. For each page, students were timed, and errors were noted. The score was correct per second rate, calculated by dividing the number of correct items by the response time in seconds for each page. The internal reliability for students who learned mathematics in English was calculated based on performance on each page, Cronbach's α = .92. Internal reliability was not available for students who learned mathematics in French because of an administration error in which only one page of the task was administered. We compared the distribution on this task for each group and found that they were highly similar and there were no violations of normality. The administration issue for the French-speaking children on this measure did not impact their overall score and thus was not a concern for data interpterion.

Number Comparison. A digit comparison task (Bigger Number; https://carleton.ca/cacr/math-lab/apps/bigger-number-app/) was administered to assess students' knowledge of the magnitude of Arabic numbers. Students were presented with two single-digit numbers (ranging from 1 to 9) on an iPad, and they were asked to choose the numerically larger number as quickly as possible. If a response was not made within three seconds, the program automatically advanced to the next trial. There were 26 experimental trials. The distance between the two numbers was manipulated such that half of the trials had a small distance ranging from 1 to 3 (e.g., 5 vs. 6), whereas the other half had a large distance ranging from 4 to 7 (e.g., 2 vs. 8; Bugden & Ansari, 2011). The score was the correct per second rate (i.e., mean accuracy/mean RT in seconds). The internal reliability (Cronbach's α) based on RT for correct trials was .96 and .93 for English and French learners, respectively.

Domain-specific Language Skills

Two novel mathematical language tasks were introduced in the present study: mathematical vocabulary and mathematical orthography.

Mathematical Vocabulary. The mathematical vocabulary task was designed to measure receptive vocabulary knowledge of mathematical words for students in grades 2 and 3. The measure was adapted from the pre-school mathematical vocabulary task developed by Purpura and Reid (2016) but was designed to extend the breadth of terms included in the measure and the applicable age range. In the mathematical vocabulary task, the child indicated which picture (from a set of two to four pictures) corresponded to a given mathematical word or expression. For example, students saw the number 425 and were asked to name the digit in the tens column. Target words were chosen from second-grade curriculum documents and are shown in Appendix A (Grade 2 word list; Manitoba Education, 2013). Target words tapped into five strands of mathematics: geometry (e.g., *cube*), spatial sense (e.g., *between*), numeration (e.g., *hundreds column*, *sum*), measurement (e.g., *mass*), and patterning (e.g., *ascending order*). The task consists of 19 items, of which three entail demonstrating knowledge of the target word in the context of very simple addition or subtraction. However, these arithmetic items had accompanying pictures of objects, so that selecting the correct target word reflected mathematical vocabulary, not arithmetic skills. Scoring is the mean proportion correct. Internal reliability based on the individual items was Cronbach's $\alpha = .67$ for the French version and .69 for the English version. Notably, the selection of items was verified for this group of students using item response modelling.

Mathematical Orthography. The mathematical symbol decision task was used to index students' knowledge of mathematical orthography – that is, their knowledge of the conventions for notating the mathematical symbol system. A similar task was developed for students in the seventh and eighth grades (ages 12 - 15) by Headley (2016). In the mathematical symbol decision task, students judged whether a string of mathematical symbols was conventional/readable (e.g., 3 – 1 = 2) or non-conventional/non-readable (e.g., 3 _ 1 = 2). Half the stimuli were orthographically correct (i.e., conventional) and half were not (i.e., non-conventional). Conventional stimuli were based on the mathematical symbols currently used in Ontario classrooms (Ontario Ministry of Education, 2005). Non-conventional stimuli were created by transforming conventional stimuli in one of three ways: (a) character substitutions (e.g., replacing + with ⊕, see Figure 1), (b) changing the symbol order (e.g., 2 + 1 = 3 changed to + 2 1 = 3), or (c) changing the spatial orientation of the symbol (e.g., a minus sign "–", changed to an underscore "_"). The full set of stimuli and a detailed description of the task are available in Appendix B.

The task is administered on an iPad (https://carleton.ca/cacr/math-lab/apps/sdt-app/). Figure 1 shows an example trial. Students were instructed to press the green check mark if the mathematical symbols look "readable or right, the way you see them on the blackboard or a textbook, and to press the red x if they look wrong or non-readable." For each trial, the symbol stimulus remained on the screen for three seconds. The green check and red x below the stimuli remained on the screen until the child responded. To ensure students understood the task, they were given feedback on a set of eight word-based practice items (i.e., cat, cta, the, hte, fly, lyf, and nde). The researcher also showed a mathematical example to the child and instructed that any arithmetic that was presented was correct (i.e., they would not see incorrect mathematical expressions, such as 1 + 1 = 3). Following the practice trials, students completed 44 test trials. Trials were presented in a random order. Trial order was randomized for each child. Response decision and response time were recorded.

Sensitivity scoring (dPrime) was used to capture the variability across all items and to assess any response bias, *c*, such as participants uniformly accepting or rejecting all items. Sensitivity scoring is recommended for two-choice tasks, especially where there is a range of item familiarity (Diependaele et al., 2012). We considered sensitivity scoring to be appropriate because the use of some symbol combinations is more frequent than others in primary classrooms (e.g., addition vs. multiplication; whole numbers vs. fractions). dPrime was calculated by taking the difference between the standardized proportion of hits (correctly accepted items) and false alarms (incorrectly accepted items). To calculate response bias, *c*, the sum of the standardized proportion of hits and false alarms is divided by -2 (Stanislaw & Todorov, 1999). A negative *c* value indicates a bias to accept all stimuli, whereas a positive *c* value indicates a bias to reject all stimuli. Response bias for the sample was low (M = 0.07, SD = 0.51) indicating that on average, students were not biased toward either accepting or rejecting all items. Internal consistency for accuracy across the 44 items using Spearman Brown split-half reliability was acceptable for both English and French speakers (ρ= .71 for each group).

<Insert Figure 1 here>

Outcomes

Word-Problem Solving. This task was an adapted form of the Applied Problem Solving subtest of the *KeyMath*, third edition (Connolly, 2007). Students were asked to solve 10 mathematical word problems that increased in difficulty (i.e., items 1 – 10). Word problems were presented verbally with a matching image. Testing discontinued if the child could not successfully solve three consecutive word problems. The dependent variable was mean accuracy. Among the 10 items, one assessed reasoning, in which students were asked to complete a sequence of numbers, one item assessed linguistic analysis in which students had to identify the information that was not useful to solve the word problem, one item assessed the arithmetic reasoning and numerical knowledge needed to find the coins that would total a given amount, one item assessed students' abilities to invent a mathematical word problem from an arithmetic equation, five items assessed students' skills in solving a range of addition word problems, and one assessed students' skills in solving a word problem that required multiplication. Students in Ontario completed the original English version of the task, whereas students in Quebec completed a version translated to French. The internal reliability (Cronbach's α) for the English and French versions of the task was .78 and .72, respectively.

Arithmetic Fluency. A paper-and-pencil calculation fluency test was administered to assess students' arithmetic fluency (Chan & Wong, 2019). The first page of the test consisted of 60 single-digit addition items arranged in three columns of 20 whose sums were less than or equal to 17 (e.g., 2 + 1 to 8 + 9). The second page of the test consisted of 60 subtraction items with minuends less than or equal to 17 (i.e., the complements of the addition problems; 3 – 2 to 17 – 9). Students were given one minute to solve as many items as possible on each page, without skipping any items. More than 90% of the students completed the first 20 items of the task. The Cronbach's α for these items for the English and French speakers of the task was .93 and .91, respectively.

Word Reading. In Ontario, word reading was assessed using the *Wechsler Individual Achievement Test (II) word reading subtest* (WIAT-II; Wechsler, 2005). On the first seven items, students were presented with a picture with four words underneath. For the first four items, students were asked to point to the word that begins with the same sound as the word in the picture. For the next three items, students selected the word that matched a given picture. After Item 7, students were presented with a page with rows of words. They were asked to read each word out loud by row from left to right. As recommended by the WIAT-II testing manual, students in grade 2 started with Item 1, and students in grade 3 started with Item 8 and received "correct" scores for the first seven items (i.e., they started with a score of 7). The task was discontinued when students make six consecutive errors or if they reached Item 41, as recommended by the manual. If the students made an error on any of the first five items, they were instructed to complete the first seven items and started with a score of 0. Scoring was the standardized total number of correct responses. The internal reliability (Cronbach's α) was .94.

In Quebec, word reading was assessed using an adapted form of the *Lecture de mots* from the *Test de rendement individuel de Wechsler, deuxième édition* (WIAT-II CDN-f; Wechsler, 2005). Students began the task at Item 22, as recommended by the testing manual. On the first eight items, the experimenter said a letter out loud and the students had to identify the letter in written form among eight letters. After Item 30, for the next nine items, students were read a set of three words and had to point to the word that did not rhyme or to the word that did not have the same starting or ending sounds. After Item 39, the next three items required students to fuse phonemes together to create a short word (for example, /ch/ and /a/ forms the word "chat", cat). Finally, after Item 42, for the last six items, the experimenter made a sound and students were asked to point the letter or group of letters associated with that sound. The task was stopped when students made six consecutive errors. At most, students were presented with items 22 to 47, for a total of 26 items. To address task-related differences, the dependent variable was the within-language standardized number of correct responses. The internal reliability (Cronbach's α) was .95.

Procedure

Written consents from parents were obtained for each child. Testing took place during the 2017-2018 school year from February to April. Research assistants who had either completed or were working toward completion of a bachelor's degree (i.e., in Psychology, Cognitive Science, or Education) administered the tests. All of the experimenters were provided with a detailed testing manual, including specific testing and scoring procedures. The experimenters completed multiple training sessions (two hours per session) during which they practiced the testing and scoring procedures and general principles of working with students. Senior researchers carefully evaluated each experimenter's ability to properly administer the tasks prior to data collection.

Each child was tested individually in a quiet area of the school by a research assistant. In Ontario, students were tested in two 20- to 30-minute sessions in a fixed order by two experimenters. More specifically, in the first session, students completed measures of receptive vocabulary, working memory, number comparison, mathematical orthography, and word reading. In the second session, students completed measures of subitizing and mathematical vocabulary. In Quebec, students were tested in one 60- to 75-minute session. The order of the test administration was randomized. Tasks were administered in English to students in Ontario and in French to students in Quebec. Apart from being administered in different languages and with the exceptions noted above, the tasks were equivalent across the two provinces. After the testing session(s), students received a sticker as an appreciation for their participation. Two additional research assistants who did not participate in data collection entered the data independently. The datasets were then compared by a computer program and any discrepancies were resolved by a third research assistant.

Results

Hypothesis 1: Comparisons of First- and Second-Language Learners' Performance

Averaging across provinces, first- and second-language learners did not differ in age (M_{age} 8:1 vs. 8:2 years), $t(156) = 1.08$, $p = .284$; gender, $\chi^2(1, N = 160) = 1.16$, $p = .281$; distribution by grade, $\chi^2(1, N = 160) = .013$, $p = .908$; or distribution across language of instruction (i.e., French vs. English), $\chi^2(1, 160) = .501$, $p = .479$. Furthermore, first- and second-language learners did not differ in their nonverbal intelligence, measured by the matrix reasoning task (Wechsler, 2014): $M = 17.0$ vs. 16.1, $t(154) = -1.35$, $p = .178$, Cohen's $d = .22$. Thus, the first- and second-language learners were comparable on demographic factors and had similar nonverbal intelligence. The mean descriptive scores for all measures, separated by French and English instruction, are provided in Appendix C. For each task, scores were converted to standardized (*z*) scores within each grade, separated by instructional language (English vs. French) to facilitate data interpretation.

The results partially supported our first hypothesis: First-language learners performed better than the second-language learners on all tasks that require a strong language component except for mathematical vocabulary. As shown in Table 1, the second-language learners had lower means than the first-language learners on receptive vocabulary, subitizing, and word-problem solving, whereas no significant difference was found on any other measures. This result occurred regardless of whether the instructional language was English or French. Additional correlational and regression analyses were also conducted, controlling for students' instructional language. We found no difference in the pattern of results with or without controlling for instructional language. Thus, for parsimony, instructional language was excluded from the reported results in this paper.

Data Reduction: Principal Component Analyses

Data were missing for individual students on some of the variables for eight tasks (3.3% across all tasks; ranged from 0% to 6.3%). Little's test for missingness showed that data were missing completely at random, $\chi^2(60) = 76.095$, $p = .079$. In the present paper, multiple imputations (i.e., five imputations using Expectation Maximization in SPSS) were used to estimate missing data. Sensitivity analyses showed that the patterns of results remained the same before and after imputations (see results without imputations in Appendix D). Thus, in the following analyses, pooled results based on multiple imputations are reported for principal component analyses, correlations, and regression analyses.

There were significant correlations among the three measures of working memory (*p*s <

.05) and the two measures of quantitative skills ($p < .001$), respectively. Thus, principal component analyses (PCA) were conducted to reduce the data and create two component scores subsequently labelled *working memory* and *quantitative skills*. The *working memory* PCA included three measures: digit span forward, digit span backward, and spatial span (factor loadings of .71, .67, and .66, respectively), accounting for 46.5% of the variance. The *quantitative* PCA included two measures: number comparison and subitizing (factor loadings = .85), accounting for 71.93% of the variance. Component scores were saved and used in subsequent analyses.

<Insert Table 1 here>

Correlations

Correlations among the variables are shown in Table 2 separately for first- and second-language learners. Language of instruction (English vs. French) did not correlate with any of the measures. Gender was only related to one variable: For first-language learners, boys performed better than girls on quantitative skills.

For first-language learners, receptive vocabulary was not correlated with either working memory or quantitative skills, whereas the latter measures were correlated with each other. Nevertheless, receptive vocabulary, working memory, and quantitative skills were correlated with all mathematical outcomes. Similarly, mathematical vocabulary and mathematical orthography were correlated with all other measures, with one exception: Mathematical orthography was not related to quantitative skills. These results are generally consistent with the findings described in the Pathways to Mathematics model such that the cognitive predictors are all correlated with mathematical outcomes (LeFevre et al., 2010; Sowinski et al., 2015). Furthermore, the lack of relation between mathematical orthography and quantitative skills suggests that a student's ability to identify mathematical written symbols is a unique component skill that is separate from their quantitative knowledge.

For the second-language learners, all cognitive predictors were correlated with each other. Receptive vocabulary and working memory skills were correlated with mathematical vocabulary and word-problem solving but not with arithmetic fluency, whereas quantitative skills were correlated with all of the mathematical measures. Nevertheless, similar to the results for first-language learners, mathematical vocabulary was correlated with all cognitive predictors. Mathematical orthography was not correlated with either quantitative skills or working memory for this group. Thus, mathematical orthography appears to capture aspects of students' mathematical knowledge distinct from their quantitative skills for both groups.

Comparing the correlational patterns across groups, it is notable that receptive vocabulary, working memory, and quantitative skills in the Pathways to Mathematics model were strongly inter-correlated for the second-language learners, but weakly related or uncorrelated for the first-language learners. Given that all of the tasks were administered in the language of instruction, these patterns presumably mean that instructional language skills, in general, are more strongly related to students' performance on a range of cognitive skills for second- than for first-language learners.

<Insert Table 2 here>

Hypotheses 2-5: Mathematical and Reading Performance for First- and Second-language Learners

We conducted hierarchical regressions to test whether the domain-specific language skills (i.e., mathematical vocabulary and mathematical orthography) would uniquely predict problem solving, arithmetic fluency, and reading performance for first- and second-language learners, controlling for cognitive predictors (i.e., receptive vocabulary, working memory, and quantitative skills). Model 1 included the domain-general cognitive predictors (linguistic, working memory, and quantitative skills). In Model 2, the domain-specific language skills were added to test the hypothesis that the linguistic pathway has become more closely tuned to the domain-specific aspects of mathematical language (Powell et al., 2017) and thus account for unique variance in the outcomes.

Hypotheses 2a and 3: Word-problem Solving

We found support for our second hypothesis: Receptive vocabulary, working memory, and quantitative skills predicted word-problem solving for both first- and second-language learners (see Table 3). In Model 1, all three cognitive skills predicted word-problem solving for both groups, with a better model fit for second-language learners. Consistent with the correlations among these precursors, Model 1 accounted for more shared variance for second- (~32%) than for first-language learners (~10%). Word-problem solving directly involves the linguistic pathway because it requires students use general language comprehension skills. The cognitive predictors from the Pathways to Mathematics model captured more individual differences for second- than for first-language learners.

Our third hypothesis was that mathematical vocabulary and mathematical orthography would mediate the relation between receptive vocabulary and word-problem solving. This hypothesis was supported for first-language learners: Both mathematical vocabulary and mathematical orthography accounted for additional unique variance in word-problem solving in Model 2. Some of the variance accounted for by Model 2 is shared among the predictors (~30%), but each predictor also accounts for unique variance. In comparing the changes from Model 1 to Model 2, it is clear that the unique r^2 associated with each of the three cognitive predictors has decreased, reflecting shared variance with the two domain-specific predictors. Further mediation analyses were conducted on the paths from receptive vocabulary to word-problem solving through mathematical vocabulary and mathematical orthography using bias-corrected bootstrapping in M*plus* (1000 samples; Muthén & Muthén, 1998). Results indicated that the indirect paths from receptive vocabulary through mathematical vocabulary and mathematical orthography were significant, $\beta = .083$, $SE = .039$, $p = .034$, CI = [.008, .166], and $\beta = .082$, $SE = .041$, $p = .047$, CI = [.022, .180], respectively. Thus, mathematical vocabulary and mathematical orthography partially mediated the relation between receptive vocabulary and word-problem solving for first-language learners, controlling for working memory and quantitative skills.

In contrast, for second-language learners, in Model 2 the two domain-specific language measures did not account for additional significant variance. There was some variance shared with receptive vocabulary, as shown by the decrease in unique variance in receptive vocabulary from Model 1 to Model 2, however, the small effect sizes for each of the predictors in Model 2 (i.e., unique r^2) indicated that these predictors did not provide further explanatory power for this group of students. Although quantitative skill was not a unique predictor in Model 2 for the second-language learners, the effect size was very similar to that for the first-language learners. Furthermore, neither of the indirect paths through mathematical vocabulary and mathematical orthography was significant, $\beta = .04$, $SE = .08$, $p = .58$, CI = [-.07, .23], and $\beta = .06$, $SE = .06$, $p = .32$, CI = [-.04, .20], respectively. Thus, for second-language learners, the relation between receptive vocabulary and word-problem solving was not mediated by either mathematical vocabulary or mathematical orthography, after controlling for working memory and quantitative skills.

In summary, our third hypothesis was supported for the first-language learners but not for the second-language learners. The relation between receptive vocabulary and word-problem solving was partially mediated by the domain-specific language skills for first-language learners, whereas there was no significant mediation for second-language learners. Further, performance on word-problem solving suggested that by the second grade, mathematics knowledge has started to differentiate for first-language learners, with domain-specific language skills beginning to account for some variance in the individual differences for word-problem solving. In contrast, for second-language learners, word-problem solving remained strongly linked to the cognitive predictors.

<Insert Table 3 here>

Hypotheses 2b and 4: Arithmetic Fluency

In addition to word problems, our second hypothesis was that the cognitive predictors would be related to arithmetic fluency for both first- and second-language learners. Notably, however, the initial analyses in which receptive vocabulary was included in Model 1 showed that there was a strong suppressor relation among arithmetic fluency, receptive vocabulary, and mathematical vocabulary for the second-language learners. Specifically, despite high correlations between receptive and mathematical vocabulary, and between mathematical vocabulary and arithmetic fluency, arithmetic fluency and receptive vocabulary were only weakly correlated. Accordingly, receptive vocabulary acted as a suppressor (Ludlow & Klein, 2014) and distorted the interpretation of the pattern of results (see Table E1 in Appendix E). Because our goal was to assess direct relations between the predictors and arithmetic fluency, we omitted receptive vocabulary from Model 1 for these analyses.

As shown in Table 4, for Model 1, working memory and quantitative skills were both unique predictors of arithmetic fluency for first-language learners, supporting Hypothesis 2b. In Model 2, mathematical vocabulary accounted for unique variance, whereas mathematical orthography did not. These results partially support our fourth hypothesis. Overall, the results were similar for the second-language learners, except that working memory was not a unique predictor. The results suggested, therefore, that variability in arithmetic performance for second-language learners was related to language skills, but it was more strongly related to their developing mathematical vocabulary than for first-language learners.

<Insert Table 4 here>

Hypothesis 5: Word Reading

We analysed word reading to test whether, as hypothesized, quantitative knowledge would only predict mathematical outcomes, not reading outcomes. In support of our fifth hypothesis, receptive vocabulary was a unique predictor of word reading for both first- and second-language learners whereas quantitative skills were not (see Table 5). This pattern supports the domain-specific nature of quantitative skills. The amount of variance accounted for by Model 1 was similar for the two groups, although working memory was not a unique predictor for the second-language learners.

<Insert Table 5 here>

Discussion

The goal of the present paper was to investigate the relations between language of instruction and mathematical learning. We had two research questions. First, are there performance differences between first- and second-language learners? Second, do domain-specific language skills relate to mathematical performance and to non-mathematical outcomes differently for first- versus second-language learners? To answer these questions, we tested the relations between language and mathematics for first- and second-language learners, considering both known cognitive predictors and domain-specific language skills. Students completed measures of cognitive skills (i.e., receptive vocabulary, working memory, and quantitative skills) and domain-specific language skills (i.e., mathematical vocabulary and mathematical orthography) in their language of instruction at school. We used multiple regression to assess the relations of cognitive and domain-specific language skills to three outcome measures:

mathematical word-problem solving, arithmetic fluency, and word reading. The results showed that, for both groups of learners, (a) all three cognitive skills were relevant for students' word-problem solving, (b) quantitative skills were important for arithmetic fluency, and (c) receptive vocabulary was important for word reading. Working memory was important for arithmetic fluency and word reading, but only for first-language learners. Moreover, mathematical vocabulary and mathematical orthography partially mediated the relation between receptive vocabulary and word-problem solving for first-language learners. In contrast, for second-language learners, receptive vocabulary in the language of instruction was more important for all aspects of mathematical outcomes. A summary of the results is shown in Table 6.

<Insert Table 6 here>

Research Question 1: Performance Differences between First- and Second-language Learners

In support of our first hypothesis, we found that, with the exception of mathematical vocabulary, second-language learners scored lower on measures with a strong language component (i.e., receptive vocabulary, subitizing, and word-problem solving) than first-language learners. These patterns support the view that second-language learners are at a disadvantage compared to first-language learners in language-demanding measures (Banks et al., 2016; Bonifacci et al., 2016; Kleemans et al., 2011). In contrast to Powell et al. (2020), we did not find significant differences in mathematical vocabulary between first- and second-language learners. One potential explanation for this pattern is that most of the students in this study had been receiving instruction in their second language since kindergarten. Thus, they presumably acquired their mathematical vocabulary at school, in the language of instruction. Results from training studies support the view that individuals who are instructed in a language that is different from their home language tend to favor the instructed language for mathematics (Spelke & Tsivkin, 2001). Thus, the lack of difference between the first- and second-language learners may indicate that both groups acquire mathematical vocabulary primarily from instruction at school. Consistent with this possibility, the two groups also performed equally well on other tasks that mainly reflected school-based instruction (i.e., mathematical orthography, mathematical vocabulary, arithmetic fluency, and word reading).

Compared to first-language learners, second-language learners had less extensive receptive vocabulary in their second language (see similar results in Bedore & Peña, 2008; Beech & Keys, 1997; Bialystok et al., 2010). This finding is consistent with the findings of a meta-analysis that showed deficits in vocabulary knowledge in second-language learners (Melby-Lervåg & Lervåg, 2014). Moreover, word-problem solving not only requires knowledge of receptive vocabulary, but also of other linguistic features of the text, such as lexical complexity (e.g., word length and familiarity; Martiniello, 2009), grammatical complexity (e.g., number of noun phrases; Haag et al., 2013; Walkington et al., 2019), phonological awareness (Bonifacci et al., 2016; LeFevre et al., 2010), expressive vocabulary (Méndez et al., 2019), and receptive syntax (Chow & Ekholm, 2019). Young second-language learners may not have mastered these skills. Furthermore, second-language learners did not perform as well as first-language learners on subitizing, but both groups performed equally well on number comparison. Presumably, second-language learners have not developed automaticity in producing number words in their second language, whereas number comparison does not involve any expressive language skills. In summary, compared to first-language learners, second-language learners may be at a disadvantage in both receptive and expressive language, both of which are related to the acquisition of mathematics. More research into the specific role of different types of linguistic skills in mathematical learning for second-language learners is needed to fully understand the complex relations between mathematical performance and language of instruction.

We expected that mathematical outcomes would be predicted by cognitive skills (i.e., receptive vocabulary, working memory, and quantitative skills) equally in first- and second-language learners. For word-problem solving, the results of the present research supported the patterns of relations in the Pathways to Mathematics model (LeFevre et al., 2010): As shown in Table 6, all of the cognitive predictors were related to word-problem solving for both groups of learners. Quantitative skills draw on a student's ability to develop associative networks of mathematical knowledge that are involved in calculation (Merkley & Ansari, 2016; Núñez, 2017; Xu & LeFevre, 2020). Working memory may be important for extracting the meaning of the text and for maintaining intermediate calculations during problem solving (Fuchs et al., 2006; Raghubar et al., 2010; Swanson, 2011). Receptive vocabulary is also required to comprehend the meaning of the text and construct a situation model (Fuchs et al., 2015). In the present study, working memory and receptive vocabulary together supported word-problem solving for both groups of learners, although they accounted for more shared variance (31.5% vs. 9.7%) and overall a larger portion of explained variance (55.9% vs. 37.1%) for second- than for first-language learners. We interpret this pattern as support for our second hypothesis that, for word-problem solving, individual differences in language skills are even more important for second- than for first-language learners.

Receptive vocabulary was not a direct predictor of arithmetic fluency. The simple correlations between these variables were low for both groups and the relations were subsumed by working memory and quantitative skills. In previous research, arithmetic fluency was most closely linked to individual differences in phonological awareness, rather than receptive vocabulary, suggesting that it is students' ability to develop mappings between symbolic knowledge and mental representations that is shared between arithmetic and language tasks (Hecht et al., 2001; Krajewski & Schneider, 2009; Singer et al., 2019). In contrast, mathematical vocabulary was significantly correlated with arithmetic fluency and this correlation was stronger for second-language learners. These patterns resulted in a suppressor relation among receptive vocabulary, mathematical vocabulary, and arithmetic and so the final model is presented without receptive vocabulary, to enhance interpretation of the patterns.

Working memory predicted arithmetic fluency only for first-language learners. It is worth noting that second-language learners completed the working memory tasks in their second language. Thus, for second-language learners, their working memory scores may reflect their language abilities as well as their general cognitive abilities. First- and second-language learners performed equally well on working memory tasks, regardless of whether the task had a verbal demand (i.e., digit span forward and backward) or had no verbal demand (i.e., spatial span; see Table 1). Nevertheless, for second-language learners, mathematical tasks that require counting, holding, and interpreting numbers in a second language may require additional processing. In previous research, adults showed qualitatively different brain responses and more extensive neural activation when they were processing numbers in their second versus their first language (Salillas & Wicha, 2012; Wang et al., 2007). Thus, although working memory capacity did not differ for first- and second-language learners as measured in the present study, it may reflect the additional language processing required to complete the task for second-language learners. Thus, working memory was only a unique predictor of arithmetic fluency for first-language learners.

Research Questions 2: The Role of Domain-specific Language Skills in Mathematical Performance

Mathematical Vocabulary

Mathematical vocabulary is critical for students learning mathematics in their first language (e.g., Hornburg et al., 2018; Kung et al., 2019; Peng & Lin, 2019; Powell et al., 2017; Purpura et al., 2017; Purpura & Reid, 2016; Toll & Van Luit, 2014). However, much less is known about the relation between mathematical vocabulary and mathematical skills for second-language learners. The results of the present study showed that mathematical vocabulary uniquely contributed to word-problem solving for first-language learners, but not for second-language learners, despite a significant zero-order correlation between the two variables in the latter group. Rather than mathematical vocabulary, cognitive skills (i.e., receptive vocabulary, working memory, and quantitative skills) accounted for most of the explained variability in word-problem solving for second-language learners. Note that for second-language learners, the cognitive predictor measures were administered in their second language. Given the presence of the strong intra-domain correlations among the cognitive predictors for the second-language learners (but not for the first-language learners), much of the variance in word-problem solving may be captured by students' skills in the language of instruction. These results further highlight the importance of general language skills in mathematical learning for second-language learners. For first-language learners, the present results are consistent with other research showing that students develop differentiated mathematical vocabulary which is related to mathematical proficiency (e.g., Kung et al., 2019; Powell et al., 2017; Purpura et al., 2017; Riccomini et al., 2015). For second-language learners, the present findings suggest that students' proficiency in the language of instruction is critical in the early grades, presumably until students have developed more skill in both the instructional language and in mathematics.

Mathematical Orthography

Mathematical language knowledge also entails an understanding of mathematical symbols. Although students often struggle to make sense of mathematical text that involves symbols (Bardini & Pierce, 2015; Rubenstein & Thompson, 2001), there is limited research examining the relations between knowledge of mathematical orthography and mathematical skills. In this study, we used a novel task to examine the role of orthography in mathematical skill development among first- and second-language learners. The results showed that, for first-language learners, mathematical orthography uniquely contributed to variability in word-problem solving. In contrast, for second-language learners, mathematical orthography was only indirectly related to word-problem solving, presumably because individual differences in receptive vocabulary continued to be the most critical language-dependent skill for these students. Mathematical orthography did not uniquely predict arithmetic fluency for either group.

The Mediating Roles of Mathematical Vocabulary and Mathematical Orthography

In support of our third hypothesis, mathematical vocabulary and mathematical orthography partially mediated the relation between receptive vocabulary and word-problem solving for first-language learners. In contrast, for second-language learners, there was no significant mediation: Individual differences in receptive vocabulary, but not in mathematical vocabulary, directly predicted word-problem solving. These findings indicated that variability in performance was tied more strongly to domain-general language skills for second- than for first-language learners. First-language learners presumably had more experience with general vocabulary skills outside of school, compared to second-language learners, and thus their language skills become differentiated earlier in development (Powell et al., 2017; Purpura & Reid, 2016; Toll & Van Luit, 2014).

The results of the present study partially supported our fourth hypothesis that mathematical vocabulary and mathematical orthography were related to arithmetic fluency. For arithmetic, mathematical vocabulary accounted for unique variance, in support of a model in which access to mathematical knowledge is segregated from more general cognitive skills. The strong suppressor pattern observed for second-language learners in which receptive vocabulary was negatively related to arithmetic suggests that, for this group, students may require a base level of general vocabulary to be able to access mathematics-specific vocabulary in their second language. Individual differences in mathematical vocabulary skills may be more closely tied to school experiences than are general vocabulary skills (Riccomini et al., 2015).

In contrast, mathematical orthography knowledge did not directly predict arithmetic for either first- or second-language learners. Our prediction that these skills would be related was based on research with adults and youth in grades 7 and 8 (Headley, 2016; Douglas et al., 2020). Notably, the children's version of the mathematical orthography task developed for this study included a broad set of mathematical symbols (i.e., procedural symbols such as operators and equations; abbreviation symbols for time, temperature, and money), whereas the adult/youth version of the task is focused specifically on more advanced symbols (i.e., pre-algebra relations and functions; Headley, 2016). This difference between tasks may explain why the adult version of the task was more strongly correlated to arithmetic ($r = 0.50$; Douglas et al., 2020) than the children's version. Presumably, familiarity with orthography for more-advanced procedural symbols is linked to arithmetic. Notably, mathematical orthography predicted word reading for both groups in the present study, suggesting that visual processes used to recognize words also may be important for learning about mathematical symbols. This was an interesting finding that needs more research.

Word Reading

To extend the findings beyond mathematics, the relation between the two domain-specific language skills and a non-mathematical outcome (i.e., word reading) was explored for first- and second-language learners. Our fifth hypothesis was supported: For both groups, students'word reading was predicted by receptive vocabulary, but not quantitative skills, highlighting the domain-specific nature of these skills. Mathematical vocabulary did not contribute to word reading above-and-beyond general vocabulary for both groups of students, suggesting that mathematical vocabulary is specific to mathematics. In contrast, the extent to which mathematical orthography is domain specific is less clear. The shared relation between mathematical orthography and word reading is stronger for the second-language learners than first-language learners, suggesting that mathematical orthography may be capturing some domain-general cognitive skills.

Limitations and Future Directions

The present study has several limitations. First, the discrepant sample sizes of first- versus second-language learners is one limitation, although this difference is an accurate reflection of the proportion of students learning in a second language in these school systems. Recruiting second-language learners has some additional difficulties, especially in a multi-cultural country like Canada. For example, it can be difficult to translate measures and test students in multiple languages because both researcher and participant must be fluent in testing languages. As always, further replication studies with larger sample sizes are needed.

Second, the only linguistic skill included in the present study was receptive vocabulary. Other aspects of linguistic skills are also related to early mathematical development. For example, phonological awareness is related to the developing numerical representations and stored arithmetic facts in the early school years (Hecht et al., 2001; Krajewski & Schneider, 2009; LeFevre et al., 2010; Singer et al., 2019). Moreover, oral comprehension may be required when students need to understand the linguistic contexts of mathematical word problems (Fuchs et al., 2015). Thus, future research should use multiple measures to investigate the contribution of linguistic skills in mathematical outcomes for second- and first-language learners.

Third, the second-language learners reported many different first languages because they were recruited from immigrant families and thus, we could not control for the specific home language. It is possible that there were differences in students' performance depending on the characteristics of the languages and their verbal number systems (for example, French-Spanish vs. French-Arabic). The specific language that students use at home, as well as how long they have received formal education in a second language, may be important factors in understanding how mathematical learning proceeds for second-language learners (Crespo et al., 2019; Jasińska & Petitto, 2018; Swanson et al., 2015). In addition, many of the first-language students were exposed to another language at home or at school, or in their community. These language complexities are a limitation inherent in the multi-cultural, multi-lingual context of our research. Despite these complexities, we captured evidence of broad similarities among diverse second-language learners and overall differences between second- and first-language learners.

Educational Implications

The present study has important educational implications. First, we observed that second-language learners were not as competent as first-language learners in subitizing – a fundamental quantitative skill that predicts mathematical outcomes. Thus, teachers may need to recognize that processing even simple numerical information may be slower for second- than for first-language learners. Providing additional time and supportive strategies may help second-language learners recognize numerical patterns and process numerical information.

Second, both receptive vocabulary and domain-specific language skills (i.e., mathematical vocabulary and mathematical orthography) were predictors of mathematical outcomes for first-language students. The vocabulary used in instruction, whether it is specific to mathematics or more general, plays an important role in how students learn new concepts. Thus, teachers should use consistent terminology to help support the development of students' mathematical vocabulary. Additionally, it is important to be explicit about which terms are mathematics specific (e.g., multiplication, radius) and which words may have different meanings outside of mathematics class (e.g., odd, even; Powell et al., 2020). Providing clear definitions and making use of examples could be beneficial for students' mathematical learning. For example, the use of external representations (e.g., manipulatives, drawings, schema) support the encoding, storing, and retrieving of information in mathematical learning (see review papers in Carbonneau et al., 2013; Lafay et al., 2019). Further, students' understanding of mathematical symbols is important for acquiring mathematical competency. Thus, explicit instruction on notating symbols may support students' fluency with the mathematical symbol system (Powell, 2015; Rubenstein & Thompson, 2001).

Mathematical vocabulary is vital for students to communicate, represent, and retrieve mathematics knowledge and concepts. Thus, support for both groups in gaining mathematics-language skills should be a target in the classroom. However, it is also important for educators to understand that proficiency in the language of instruction may be more important for second- than for first-language learners. Students who learn mathematics in a second language may be unfamiliar with the general vocabulary and linguistic structures encountered in the classroom. Second-language learners may require additional support so that their mathematical performance is less strongly tied to their second-language skills.

In conclusion, for second language learners, mathematical skills are closely tied to students' second-language skills and language skills appear to play a more important role for second- than first-language learners. The results of the present study highlight the importance of instructional language for students learning mathematics in a second language and show that language processes are important for mathematical learning in general.

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*(3), 219–234. https://doi.org/10.1207/S15324818AME1403_2

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, *35*(1), 127–145.

Alibali, M. W., Crooks, N. M., & McNeil, N. M. (2018). Perceptual support promotes strategy generation: Evidence from equation solving. *British Journal of Developmental Psychology*, *36*(2), 153–168. https://doi.org/10.1111/bjdp.12203

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the Automated Working Memory Assessment. *Educational Psychology*, *28*(7), 725–734. https://doi.org/10.1080/01443410802243828

Ambrose, R., & Molina, M. (2014). Spanish/English bilingual students' comparison of comprehension of arithmetic story problem texts. *International Journal of Science and Mathematics Education*, *12*(6), 1469–1496. https://doi.org/10.1007/s10763-013-9472-2

Banks, K., Jeddeeni, A., & Walker, C. M. (2016). Assessing the effect of language demand in bundles of math word problems. *International Journal of Testing*, *16*(4), 269–287. https://doi.org/10.1080/15305058.2015.1113972

Bardini, C., & Pierce, R. (2015). Assumed knowledge: The challenge of symbols. *International Journal of Innovation in Science and Mathematics Education*, *23*(1), 1–9.

Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for ddentification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, *11*(1), 1–29. https://doi.org/10.2167/beb392.0

Beech, J. R., & Keys, A. (1997). Reading, vocabulary and language preference in 7- to 8-year-old bilingual Asian children. *British Journal of Educational Psychology*, *67*(4), 405–414. https://doi.org/10.1111/j.2044-8279.1997.tb01254.x

Bernardo, A. B. I. (2002). Language and mathematical problem solving among bilinguals. *The Journal of Psychology*, *136*(3), 283–297. https://doi.org/10.1080/00223980209604156

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525–531. https://doi.org/10.1017/S1366728909990423

Bonifacci, P., Tobia, V., Bernabini, L., & Marzocchi, G. M. (2016). Early literacy and numeracy skills in bilingual minority children: Toward a relative independence of linguistic and numerical processing. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01020

Bugden, S., & Ansari, D. (2011). Individual differences in children's mathematical competence are related to the intentional but not automatic processing of Arabic numerals. *Cognition*, *118*(1), 32–44. https://doi.org/10.1016/j.cognition.2010.09.005

Bugden, S., Peters, L., Nosworthy, N., Archibald, L., & Ansari, D. (2020). Identifying children with persistent developmental dyscalculia from a 2-min test of symbolic and nonsymbolic numerical magnitude processing. *Mind, Brain, and Education*, https://doi.org/10.1111/mbe.12268

Bynner, J., & Parsons, S. (1997). *Does Numeracy Matter? Evidence from the National Child Development Study on the Impact of Poor Numeracy on Adult Life*. The Basic Skills Agency.

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology*, *105*(2), 380–400. https://doi.org/10.1037/a0031084

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school*. Heinemann.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Chan, W. W. L., & Wong, T. T.-Y. (2019). Subtypes of mathematical difficulties and their stability. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000383

Chow, J. C., & Ekholm, E. (2019). Language domains differentially predict mathematics performance in young children. *Early Childhood Research Quarterly*, *46*, 179–186. https://doi.org/10.1016/j.ecresq.2018.02.011

Cirino, P. T. (2011). The interrelationships of mathematical precursors in kindergarten. *Journal of Experimental Child Psychology*, *108*(4), 713–733. https://doi.org/10.1016/j.jecp.2010.11.004

Cobb, P. (1987). An investigation of young children's academic arithmetic contexts. *Educational Studies in Mathematics*, *18*(2), 109–124. https://doi.org/10.1007/BF00314722

Cohen, L., Dehaene, S., Chochon, F., Lehéricy, S., & Naccache, L. (2000). Language and calculation within the parietal lobe: A combined cognitive, anatomical and fMRI study. *Neuropsychologia*, *38*(10), 1426–1440. https://doi.org/10.1016/S0028-3932(00)00038-5

Connolly, A. J. (2007). *KeyMath-3 Diagnostic Assessment*. Pearson.

Crespo, K., Gross, M., & Kaushanskaya, M. (2019). The effects of dual language exposure on executive function in Spanish–English bilingual children with different language abilities. *Journal of Experimental Child Psychology*, *188*, 104663. https://doi.org/10.1016/j.jecp.2019.104663

Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, *34*(4), 344–377. https://doi.org/10.1016/j.dr.2014.10.001

Cross, A. M., Joanisse, M. F., & Archibald, L. M. D. (2019). Mathematical abilities in children with developmental language disorder. *Language, Speech, and Hearing Services in Schools*, *50*(1), 150–163. https://doi.org/10.1044/2018_LSHSS-18-0041

Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H.-C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology*, *06*. https://doi.org/10.3389/fpsyg.2015.00348

de Araujo, Z., Roberts, S. A., Willey, C., & Zahner, W. (2018). English Learners in K–12 Mathematics Education: A Review of the Literature. *Review of Educational Research*, *88*(6), 879–919. https://doi.org/10.3102/0034654318798093

Diependaele, K., Brysbaert, M., & Neri, P. (2012). How noisy is lexical decision? *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00348

Dehaene, S., & Cohen, L. (1991). Two mental calculation systems: A case study of severe acalculia with preserved approximation. *Neuropsychologia*, *29*(11), 1045–1074. https://doi.org/10.1016/0028-3932(91)90076-K

Douglas, H., Headley, G., Hadden, S., LeFevre, J.-A. (2020). Knowledge of mathematical symbols goes beyond numbers. *Journal of Numerical Cognition, 6*(3), doi: 10.5964/jnc.v6i3.293

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Dunn, L. M., & Dunn, D. M. (2012). *Peabody Picture Vocabulary Test—Fourth Edition*. American Psychological Association. https://doi.org/10.1037/t15144-000

Dunn, L. M., Dunn, L. M., & Thériault-Whalen, C. M. (1993). *EVIP: échelle de vocabulaire en images Peabody*. PSYCAN.

Éducation Loisir et Sport Québec. (2009). *Progression of Learning Mathematics*.

Forsyth, S. R., & Powell, S. R. (2017). Differences in the mathematics-vocabulary knowledge of fifth-grade students with and without learning difficulties: Fifth-grade mathematics vocabulary. *Learning Disabilities Research & Practice*. https://doi.org/10.1111/ldrp.12144

Fuchs, L. S., Fuchs, D., Compton, D. L., Hamlett, C. L., & Wang, A. Y. (2015). Is word-problem solving a form of text comprehension? *Scientific Studies of Reading*, *19*(3), 204–223. https://doi.org/10.1080/10888438.2015.1005745

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. https://doi.org/10.1037/0022-0663.98.1.29

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology*, *100*(1), 30–47. https://doi.org/10.1037/0022-0663.100.1.30

Friso-Van Den Bos, I., Van Der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Geary, D. C., Cormier, P., Goggin, J. P., Estrada, P., & Lunn, M. C. E. (1993). Mental arithmetic: A componential analysis of speed-of-processing across monolingual, weak bilingual, and strong bilingual adults. *International Journal of Psychology*, *28*(2), 185–201. https://doi.org/10.1080/00207599308247184

Gerardi, K. S., Goette, L. F., & Meier, S. (2010). Financial literacy and subprime mortgage delinquency: Evidence from a survey matched to administrative data. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1600905

Gjicali, K., Astuto, J., & Lipnevich, A. A. (2019). Relations among language comprehension, oral counting, and numeral knowledge of ethnic and racial minority young children from low-income communities. *Early Childhood Research Quarterly*, *46*, 5–19. https://doi.org/10.1016/j.ecresq.2018.07.007

Haag, N., Heppt, B., Stanat, P., Kuhl, P., & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction*, *28*, 24–34. https://doi.org/10.1016/j.learninstruc.2013.04.001

Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The Relations between Phonological Processing Abilities and Emerging Individual Differences in Mathematical Computation Skills: A Longitudinal Study from Second to Fifth Grades. *Journal of Experimental Child Psychology*, *79*(2), 192–227. https://doi.org/10.1006/jecp.2000.2586

Hebert, M. A., & Powell, S. R. (2016). Examining fourth-grade mathematics writing: features of organization, mathematics vocabulary, and mathematical representations. *Reading and Writing, 29*(7), 1511-1537. https://doi.org/10.1007/s11145-016-9649-5

Headley, M. G. (2016). *What is Symbolic Mathematics Language Literacy? A Multilevel Mixed Methods Study of Adolescents in a Middle School* [Electronic Dissertation, University of Cincinnati, Education]. http://rave.ohiolink.edu/etdc/view?acc_num=ucin1470045155

Herscovics, N., & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics*, *27*(1), 59–78. https://doi.org/10.1007/BF01284528

Hooper, S. R., Roberts, J., Sideris, J., Burchinal, M., & Zeisel, S. (2010). Longitudinal predictors of reading and math trajectories through middle school for African American versus Caucasian students across two samples. *Developmental Psychology*, *46*(5), 1018–1029. https://doi.org/10.1037/a0018877

Hornburg, C. B., Schmitt, S. A., & Purpura, D. J. (2018). Relations between preschoolers' mathematical language understanding and specific numeracy skills. *Journal of Experimental Child Psychology*, *176*, 84–100. https://doi.org/10.1016/j.jecp.2018.07.005

Hughes, E. M., Powell, S. R., & Stevens, E. A. (2016). Supporting clear and concise mathematics language: Instead of that, say this. *TEACHING Exceptional Children*, *49*(1), 7–17. https://doi.org/10.1177/0040059916654901

Jasińska, K. K., & Petitto, L.-A. (2018). Age of bilingual exposure is related to the contribution of phonological and semantic knowledge to successful reading development. *Child Development*, *89*(1), 310–331. https://doi.org/10.1111/cdev.12745

Kleemans, T., Segers, E., & Verhoeven, L. (2011). Cognitive and linguistic precursors to numeracy in kindergarten: Evidence from first and second language learners. *Learning and Individual Differences*, *21*(5), 555–561. https://doi.org/10.1016/j.lindif.2011.07.008

Kotsopoulos, D. (2007). Like a foreign language. *Mathematics Teacher*, *101*(4), 301–305.

Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual–spatial working memory, and preschool quantity–number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology*, *103*(4), 516–531. https://doi.org/10.1016/j.jecp.2009.03.009

Kung, M., Schmitt, S. A., Zhang, C., Whiteman, S. D., Yang, F., & Purpura, D. J. (2019). The role of mathematical language in mathematics development in China and the US. *International Journal of Educational Research*, *95*, 131–142. https://doi.org/10.1016/j.ijer.2019.02.008

Kutner, M., Greenberg, E., Jin, Y., Boyle, B., Hsu, Y., & Dunleavy, E. (2007). *Literacy in everyday life: Results from the 2003 National Assessment of Adult Literacy*. Department of Education, National Center for Education Statistics.

Lafay, A., Osana, H. P., & Valat, M. (2019). Effects of Interventions with Manipulatives on Immediate Learning, Maintenance, and Transfer in Children with Mathematics Learning Disabilities: A Systematic Review. *Education Research International*, *2019*, 1–21. https://doi.org/10.1155/2019/2142948

Swanson, H., 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. https://doi.org/10.1016/j.jecp.2015.01.001

LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to Mathematics: Longitudinal predictors of performance. *Child Development*, *81*(6), 1753–1767. https://doi.org/10.1111/j.1467-8624.2010.01508.x

Li, Y. (2008). What do students need to learn about division of fractions? *Mathematics Teaching in Middle School*, *6125*(9), 546–552.

Ludlow, L., & Klein, K. (2014). Suppressor variables: The difference between 'is' versus 'acting as'. *Journal of Statistics Education*, *22*(2), 6. https://doi.org/10.1080/10691898.2014.11889703

Lusardi, A. (2012). Numeracy, financial literacy, and financial decision-making. *Numeracy*, *5*(1). https://doi.org/10.5038/1936-4660.5.1.2

Manitoba Education. (2013). *Kindergarten to Grade 8 Mathematics*.

Marsh, L. G., & Maki, R. H. (1976). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory & Cognition*, *4*(4), 459–464. https://doi.org/10.3758/BF03213203

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. https://doi.org/10.1080/10627190903422906

McClain, L., & Huang, J. Y. S. (1982). Speed of simple arithmetic in bilinguals. *Memory & Cognition*, *10*(6), 591–596. https://doi.org/10.3758/BF03202441

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology*, *92*(4), 734–744. https://doi.org/10.1037/0022-0663.92.4.734

McNeil, N. M., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Brletic-Shipley, H. (2011). Benefits of Practicing 4 = 2 + 2: Nontraditional Problem Formats Facilitate Children's Understanding of Mathematical Equivalence: Nontraditional Arithmetic Practice. *Child Development*, *82*(5), 1620–1633. https://doi.org/10.1111/j.1467-8624.2011.01622.x

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin*, *140*(2), 409–433. https://doi.org/10.1037/a0033890

Méndez, L. I., Hammer, C. S., Lopez, L. M., & Blair, C. (2019). Examining language and early numeracy skills in young Latino dual language learners. *Early Childhood Research Quarterly*, *46*, 252–261. https://doi.org/10.1016/j.ecresq.2018.02.004

Merkley, R., & Ansari, D. (2016). Why numerical symbols count in the development of mathematical skills: Evidence from brain and behavior. *Current Opinion in Behavioral Sciences*, *10*, 14–20. https://doi.org/10.1016/j.cobeha.2016.04.006

Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide*. Muthén & Muthén.

Núñez, R. E. (2017). Is there really an evolved capacity for number? *Trends in Cognitive Sciences*, *21*(6), 409–424. https://doi.org/10.1016/j.tics.2017.03.005

O'Halloran, K. L. (2005). *Mathematical discourse: Language, symbolism and visual images*. Continuum.

O'Halloran, K. L. (2015). The language of learning mathematics: A multimodal perspective. *The Journal of Mathematical Behavior*, *40*, 63–74. https://doi.org/10.1016/j.jmathb.2014.09.002

Ontario Ministry of Education. (2005). *The Ontario Curriculum, Grades 1-8: Mathematics*. http://www.edu.gov.on.ca/eng/curriculum/elementary/math.html

Peng, P., & Lin, X. (2019). The relation between mathematics vocabulary and mathematics performance among fourth graders. *Learning and Individual Differences*, *69*, 11–21. https://doi.org/10.1016/j.lindif.2018.11.006

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. https://doi.org/10.1037/edu0000079

Powell, S. R. (2015). The Influence of Symbols and Equations on Understanding Mathematical Equivalence. *Intervention in School and Clinic*, *50*(5), 266–272. https://doi.org/10.1177/1053451214560891

Powell, S. R., Berry, K. A., & Tran, L. M. (2020). Performance differences on a measure of mathematics vocabulary for english learners and non-english learners with and without mathematics difficulty. *Reading & Writing Quarterly*, *36*(2), 124–141. https://doi.org/10.1080/10573569.2019.1677538

Powell, S. R., Driver, M. K., Roberts, G., & Fall, A.-M. (2017). An analysis of the mathematics vocabulary knowledge of third- and fifth-grade students: Connections to general vocabulary and mathematics computation. *Learning and Individual Differences*, *57*, 22–32. https://doi.org/10.1016/j.lindif.2017.05.011

Powell, S. R., & Fluhler, S. K. (2018). A brief measure of mathematics symbols: Performance at grades 1, 3, and 5. *Assessment for Effective Intervention*, 153450841880988. https://doi.org/10.1177/1534508418809884

Powell, S. R., & Nelson, G. (2017). An investigation of the mathematics-vocabulary knowledge of first-grade students. *The Elementary School Journal*, *117*(4), 664–686. https://doi.org/10.1086/691604

Powell, S. R., Stevens, E. A., & Hughes, E. M. (2019). Math language in middle school: Be more specific. *TEACHING Exceptional Children*, *51*(4), 286–295. https://doi.org/10.1177/0040059918808762

Powell, S. R., Urrutia, V. Y., Berry, K. A., & Barnes, M. A. (2020). The word-problem solving and explanations of students experiencing mathematics difficulty: A comparison based on dual-language status. *Learning Disability Quarterly*, 073194872092219. https://doi.org/10.1177/0731948720922198

Purpura, D. J., & Ganley, C. M. (2014). Working memory and language: Skill-specific or domain-general relations to mathematics? *Journal of Experimental Child Psychology*, *122*, 104–121. https://doi.org/10.1016/j.jecp.2013.12.009

Purpura, D. J., Logan, J. A. R., Hassinger-Das, B., & Napoli, A. R. (2017). Why do early mathematics skills predict later reading? The role of mathematical language. *Developmental Psychology*, *53*(9), 1633–1642. https://doi.org/10.1037/dev0000375

Purpura, D. J., & Reid, E. E. (2016). Mathematics and language: Individual and group differences in mathematical language skills in young children. *Early Childhood Research Quarterly*, *36*, 259–268. https://doi.org/10.1016/j.ecresq.2015.12.020

Ministère de l'Éducation et de l'Enseignement Supérieur (MEES). (2018). Indices de défavorisation. Québec, QC, Canada: Gouvernement du Québec.

Ministère de l’Éducation, du Loisir et du Sport (MELS) (2009). Progression des apprentissages au primaire: Mathématique. Québec, QC, Canada. *http://www1.education.gouv.qc.ca/progressionPrimaire/mathematique*.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Riccomini, P. J., Smith, G. W., Hughes, E. M., & Fries, K. M. (2015). The Language of mathematics: The importance of teaching and learning mathematical vocabulary. *Reading & Writing Quarterly*, *31*(3), 235–252. https://doi.org/10.1080/10573569.2015.1030995

Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, *24*(7), 1301–1308. https://doi.org/10.1177/0956797612466268

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*(1), 175–189. https://doi.org/10.1037/0022-0663.91.1.175

Roberts, N., & Truxaw, M. (2013). For ELLs: Vocabulary beyond the Definitions. *The Mathematics Teacher*, *107*(1), 28. https://doi.org/10.5951/mathteacher.107.1.0028

Rubenstein, R. N., & Thompson, D. R. (2001). Learning mathematical symbolism: Challenges and instructional strategies. *The Mathematics Teacher*, *94*(4), 265–271.

Salillas, E., & Wicha, N. Y. Y. (2012). Early learning shapes the memory networks for arithmetic: Evidence from brain potentials in bilinguals. *Psychological Science*, *23*(7), 745–755. https://doi.org/10.1177/0956797612446347

Sherman, J., & Bisanz, J. (2009). Equivalence in symbolic and nonsymbolic contexts: Benefits of solving problems with manipulatives. *Journal of Educational Psychology*, *101*(1), 88–100. https://doi.org/10.1037/a0013156

Singer, V., Strasser, K., & Cuadro, A. (2019). Direct and indirect paths from linguistic skills to arithmetic school performance. *Journal of Educational Psychology*, *111*(3), 434–445. https://doi.org/10.1037/edu0000290

Sowinski, C., LeFevre, J.-A., Skwarchuk, S.-L., Kamawar, D., Bisanz, J., & Smith-Chant, B. (2015). Refining the quantitative pathway of the Pathways to Mathematics model. *Journal of Experimental Child Psychology*, *131*, 73–93. https://doi.org/10.1016/j.jecp.2014.11.004

Spelke, E. S., & Tsivkin, S. (2001). Language and number: A bilingual training study. *Cognition*, *78*(1), 45–88. https://doi.org/10.1016/S0010-0277(00)00108-6

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Steinberg, R. M., Sleeman, D. H., & Ktorza, D. (1991). Algebra students' knowledge of equivalence of equations. *Journal for Research in Mathematics Education*, *22*(2), 112. https://doi.org/10.2307/749588

Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology*, *103*(4), 821–837. https://doi.org/10.1037/a0025114

Telli, S., Rasch, R., & Schnotz, W. (2018). A Phenomenological perspective to bilingual students' word problems solving behaviours. *International Journal of Research in Education and Science*, 517–533. https://doi.org/10.21890/ijres.428302

Thompson, D. R., & Rubenstein, R. N. (2000). Learning mathematics vocabulary: Potential pitfalls and instructional strategies. *Mathematics Teacher*, *93*(7), 568–574.

Thompson, D. R., & Rubenstein, R. N. (2014). Literacy in language and mathematics: More in common than you think. *Journal of Adolescent & Adult Literacy*, *58*(2), 105–108. https://doi.org/10.1002/jaal.338

Toll, S. W. M., & Van Luit, J. E. H. (2014). The developmental relationship between language and low early numeracy skills throughout kindergarten. *Exceptional Children*, *81*(1), 64–78. https://doi.org/10.1177/0014402914532233

Träff, U., Olsson, L., Skagerlund, K., & Östergren, R. (2018). Cognitive mechanisms underlying third graders' arithmetic skills: Expanding the pathways to mathematics model. *Journal of Experimental Child Psychology*, *167*, 369–387. https://doi.org/10.1016/j.jecp.2017.11.010

Van Rinsveld, A., Brunner, M., Landerl, K., Schiltz, C., & Ugen, S. (2015). The relation between language and arithmetic in bilinguals: Insights from different stages of language acquisition. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00265

Vlassis, J. (2008). The role of mathematical symbols in the development of number conceptualization: The case of the minus sign. *Philosophical Psychology*, *21*(4), 555–570. https://doi.org/10.1080/09515080802285552

Vukovic, R. K., & Lesaux, N. K. (2013). The relationship between linguistic skills and arithmetic knowledge. *Learning and Individual Differences*, *23*, 87–91. https://doi.org/10.1016/j.lindif.2012.10.007

Walkington, C., Clinton, V., & Sparks, A. (2019). The effect of language modification of mathematics story problems on problem-solving in online homework. *Instructional Science*, *47*(5), 499–529. https://doi.org/10.1007/s11251-019-09481-6

Wang, Y., Lin, L., Kuhl, P., & Hirsch, J. (2007). Mathematical and linguistic processing differs between native and second Languages: An fMRI study. *Brain Imaging and Behavior*, *1*(3–4), 68–82. https://doi.org/10.1007/s11682-007-9007-y

Wechsler, D. (2005). *Wechsler Individual Achievement Test 2nd Edition (WIAT II)*. The Psychological Corp.

Wechsler, D. (2014). *WISC-V: Technical and Interpretive Manual*. NCS Pearson, Inc.

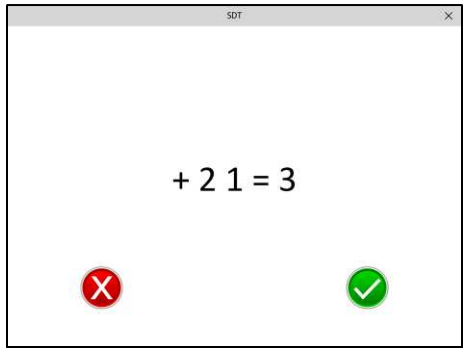
Wolf, M. K., & Leon, S. (2009). An Investigation of the Language Demands in Content Assessments for English Language Learners. *Educational Assessment*, *14*(3–4), 139–159. https://doi.org/10.1080/10627190903425883

Xu, C., & LeFevre, J.-A. (2020). Children's knowledge of symbolic number in Grades 1 and 2: Integration of associations. *Child Development,* doi: 10.1111/CDEV.13473

Figure and Tables

Figure 1

Screenshot of the Children's Symbol Decision Task



Note. In this non-conventional example, the + sign and operands are incorrectly ordered.

Table 1

Performance on Measures for First-Language Learners and Second-Language Learners

	M (*SD*)		*t*	*df*	*CI*s	Cohen's *d*
	First	Second				
Receptive Vocabulary[a]	.24 (.71)	-.44 (1.25)	3.78***	76.43	[.32, 1.04]	.67
Digit Span Forward	.07 (.97)	-.13 (1.04)	1.17	150.00	[-.14, .52]	.20
Digit Span Backward	.03 (.95)	-.06 (1.08)	0.53	150.00	[.24, .42]	.09
Spatial Span	.03 (.93)	-.05 (1.11)	0.49	149.00	[-.25, .42]	.08
Subitizing	.14 (.95)	-.25 (1.02)	2.41*	152.00	[.07, .72]	.40
Number Comparison[a]	.05 (.89)	-.09 (1.17)	-.81	92.37	[-.21, .50]	.13
Mathematical Vocabulary[a]	.01 (.91)	-.02 (1.14)	.16	92.42	[-.32, .38]	.03
Mathematical Orthography	.01 (.94)	-.03 (1.09)	.24	148.00	[-.29, .38]	.04
Word-problem Solving	.13 (.97)	-.23 (1.01)	2.16*	152.00	[.03, .69]	.36
Arithmetic Fluency	.01 (.99)	-.01 (1.00)	.13	152.00	[-.31, .35]	.02
Word Reading	.03 (.90)	-.05 (1.14)	.50	158.00	[-.41, .24]	.08

Note. Scores were standardized within grade (grade 2, grade 3) and language (English, French). [a]Adjusted *df* was used to correct for unequal variance; $*p < .05$, $***p < .001$

Table 2

Correlations among Measures for First-Language Learners and Second-language Learners

	1	2	3	4	5	6	7	8	9	10
1. Gender[a]	-	.03	.20	.14	-.17	.07	.17	-.02	-.15	.17
2. Language Group[b]	-.01	-	.05	.19	.09	.25	.05	.18	.20	.09
3. Receptive Vocabulary[c]	.10	-.01	-	.45**	.30*	.60***	.46***	.63***	.13	.44***
4. Working Memory[d]	-.14	-.10	.16	-	.47**	.42***	.26	.60***	.34*	.34**
5. Quantitative[d]	-.24*	-.05	.05	.24*	-	.42**	.18	.49***	.61***	.27*
6. Mathematical Vocabulary[c]	-.11	-.15	.40***	.28**	.32**	-	.45**	.55***	.56***	.43***
7. Mathematical Orthography[e]	.07	-.04	.34***	.29*	.13	.43***	-	.45***	.20	.45***
8. Word Problem Solving[c]	-.11	-.05	.42***	.43***	.35***	.54***	.52***	-	.38**	.38**
9. Arithmetic Fluency[c]	-.18	-.09	.21*	.45***	.54***	.48***	.34**	.61***	-	.26*
10. Word Reading[c]	.02	-.06	.30**	.36***	.25*	.38*	.38***	.36***	.42***	-

Note. $p < .05$*, $p < .01$**, $p < .001$***; Correlations among Measures for first-language learners (below the diagonal) and second-language learners (above the diagonal).[a]Gender is coded as 1=male and 2=female; [b]Instructional Language is coded as 1=English and 2=French; [c]Standardized total score; [d]Standardized PCA component score; [e]Standardized dPrime score

Table 3

Hierarchical Regression of Word Problem Solving

	First-Language Learners [a]				Second-Language Learners [b]			
	Beta	*t*	*p*	*Unique r^2*	*Beta*	*t*	*p*	*Unique r^2*
Model 1								
Receptive Vocabulary	**.352*** **	**4.274*** **	**<.001**	**.121*** **	**.427*** **	**4.144*** **	**<.001**	**.144*** **
Working Memory	**.314*** **	**3.730*** **	**<.001**	**.091*** **	**.302***	**2.284***	**.027**	**.062***
Quantitative Skills	**.257** **	**3.077** **	**.002**	**.062** **	**.222***	**2.101***	**.036**	**.038***
Averaged ΔR^2 across imputations	**.371*** **				**.559*** **			
Model 2								
Receptive Vocabulary	**.188***	**2.278***	**.023**	**.028***	**.317***	**2.579***	**.010**	**.056***
Working Memory	**.216** **	**2.644** **	**.008**	**.040** **	**.285***	**2.214***	**.031**	**.054***
Quantitative Skills	**.181***	**2.301***	**.021**	**.029***	.195	1.809	.071	.028
Mathematical Vocabulary	**.227***	**2.526***	**.012**	**.034***	.097	.772	.440	.005
Mathematical Orthography	**.271** **	**2.903** **	**.005**	**.055** **	.147	1.378	.168	.016
Averaged ΔR^2 across imputations	**.126*** **				.027			

Note. Squared semi-partial correlations indicate unique r^2 within that specific model tested. $p < .05$*, $p < .01$**, $p < .001$***. Numbers in boldface indicate significant results at $p < .05$.

[a]Model R^2s across imputations ranged from 36.0% to 37.9% in Model 1, $F(3, 99)$ = from 18.58 to 20.15, *p*s < .001, and in Model 2, from 47.1% to 51.8%, $M = 49.7\%$, $F(5, 97)$ = from 17.28 to 20.87, *p*s < .001 for first-language learners.

[b]Model R^2s across imputations ranged from 52.0% to 59.5% in Model 1, $F(3,53)$ = from 19.11 to 25.97, *p*s < .001, and in Model 2, from 55.0% to 62.0%, $M = 58.6\%$, $F(5, 51)$ = from 12.48 to 16.63, *p*s < .001 for second-language learners.

Table 4

Hierarchical regression of Arithmetic Fluency

	First-Language Learners [a]				Second-Language Learners [b]			
	Beta	*t*	*p*	*Unique r^2*	*Beta*	*t*	*p*	*Unique r^2*
Model 1								
Working Memory	**.338*****	**4.088*****	**<.001**	**.108*****	.064	.486	.628	.003
Quantitative Skills	**.457*****	**5.565*****	**<.001**	**.197*****	**.580*****	**4.455*****	**<.001**	**.263*****
Averaged ΔR^2 across imputations	**.397*****				**.377*****			
Model 2								
Working Memory	**.254****	**3.102****	**.002**	**.056****	-.043	-.330	.724	.001
Quantitative Skills	**.389*****	**4.849*****	**<.001**	**.132*****	**.473*****	**3.840*****	**<.001**	**.161*****
Mathematical Vocabulary	**.231****	**2.616****	**.009**	**.038****	**.400*****	**3.213*****	**.001**	**.104*****
Mathematical Orthography	.117	1.338	.181	.011	-.051	-.384	.702	.002
Averaged ΔR^2 across imputations	**.076****				**.114****			

Note. Squared semi-partial correlations indicate unique r^2 within that specific model tested. $p < .05$*, $p < .01$**, $p < .001$***. Numbers in boldface indicate significant results at $p < .05$.

[a]Model R^2s across imputations ranged from 38.5% to 41.0% in Model 1, $F(2, 100)$ = from 31.23 to 34.81, *p*s < .001, and in Model 2, from 45.3% to 49.1%, M = 47.4%, $F(4, 98)$ = from 20.28 to 23.59, *p*s < .001 for first-language learners.

[b]Model R^2s across imputations ranged from 35.9% to 40.2% in Model 1, $F(2,54)$ = from 15.11 to 18.15, *p*s < .001, and in Model 2, from 47.7% to 51.3%, M = 49.1%, $F(4, 52)$ = from 11.86 to 13.72, *p*s < .001 for second-language learners.

Table 5

Hierarchical Regression of Word Reading

	First-Language Learners [a]				Second-Language Learners [b]			
	Beta	*t*	*p*	*Unique* r^2	*Beta*	*t*	*p*	*Unique* r^2
Model 1								
Receptive Vocabulary	**.244***	**2.695***	**.007**	**.058***	**.350***	**2.567***	**.010**	**.097**
Working Memory	**.282***	**2.970***	**.003**	**.073***	.136	.887	.375	.013
Quantitative Skills	.170	1.822	.068	.027	.103	.740	.459	.008
Averaged ΔR^2 across imputations	**.217*****				**.229****			
Model 2								
Receptive Vocabulary	.132	1.360	.174	.014	.168	1.050	.294	.016
Working Memory	**.215***	**2.217***	**.027**	**.040***	.108	.730	.466	.008
Quantitative Skills	.120	1.266	.205	.013	.066	.472	.637	.003
Mathematical Vocabulary	.149	1.387	.166	.015	.132	.831	.406	.009
Mathematical Orthography	.191	1.891	.059	.027	**.275***	**2.011***	**.044**	**.056***
Averaged ΔR^2 across imputations	**.059***				.080			

Note. Squared semi-partial correlations indicate unique r^2 within that specific model tested. $p < .05$*, $p < .01$**. Numbers in boldface indicate significant results at $p < .05$.

[a]Model R^2s across imputations ranged from 20.9 to 23.1% in Model 1, $F(3, 99)$ = from 8.70 to 9.89, *p*s < .001, and in Model 2, from 25.7% to 29.3%, $M = 27.6\%$, $F(5, 97)$ = from 6.72 to 8.04, *p*s < .001 for first-language learners.

[b]Model R^2s across imputations ranged from 22.1% to 23.7% in Model 1, $F(3,53)$ = from 5.00 to 5.50, *p*s < .01, and in Model 2, from 30.6% to 31.4%, $M = 30.9\%$, $F(5, 51)$ = from 4.50 to 4.67, *p*s < .01 for second-language learners.

Table 6

Summary of Patterns of Significance for Regression Models of Mathematical Outcomes and Reading for First and Second-language Learners

	Word-Problem Solving		Arithmetic Fluency		Word Reading	
	First	Second	First	Second	First	Second
Model 1						
Receptive Vocabulary	✓	✓	-	-	✓	✓
Working Memory	✓	✓	✓	×	✓	×
Quantitative Skills	✓	✓	✓	✓	†	×
Model 2						
Receptive Vocabulary	✓	✓	-	-	×	×
Working Memory	✓	✓	✓	×	✓	×
Quantitative Skills	✓	†	✓	✓	×	×
Mathematical Vocabulary	✓	×	✓	✓	×	×
Mathematical Orthography	✓	×	×	×	†	✓

Note. The "✓" indicates significant unique variance in that model. The "-" indicates that the variable was omitted from the model (see text for details). The "×" indicates that the variable did not predict unique variance. The † indicates that $.05 < p < .08$.

Appendix A

Evaluation of Mathematical Vocabulary Measure

The mathematical vocabulary task was modelled on the pre-school mathematical vocabulary task developed by Purpura and Reid (2016) and was designed to extend the age range and breadth of terms included in the Purpura task. To understand and validate the measure, we used item response theory for a nuanced analysis of the individual test items.

Using an item response model, we assessed how individual test items relate to the latent trait, mathematical vocabulary. This knowledge can inform our understanding of task validity (Kimberlin & Winterstein, 2008; Reise et al., 1993). A two-parameter logistic (2-PL) item response analysis was run using Stata/SE 15.1 (StataCorp, 2017b), estimating item discrimination and item difficulty. Item discrimination estimates the probability of success on an individual item based on overall mathematical vocabulary (i.e., the latent trait estimate). A high positive discrimination level indicates that the item differentiates well between low and high levels of mathematical vocabulary. In contrast, a negative discrimination level indicates that students with strong mathematical vocabulary are less likely to choose the correct item than students with weak mathematical vocabulary. The difficulty parameter indicates how challenging the item is; if the difficulty parameter is low, the probability of success on the item is high. Reviewing these properties, we can ensure test items have a range of difficulty and positively discriminate between low and high levels of mathematical vocabulary knowledge.

As shown in Table A1, the task items captured a range of difficulty (b = -3.71 to 1.09) and no two items had the same difficulty parameters. Item 5 (“nearest”) had the highest discrimination parameter (a = 7.34) and was one of the least difficult (b = -2.45) items suggesting that if a student could not answer this question, they would not perform well on the measure. On the other end of the scale, item 18 (“difference”) also had a relatively high discrimination parameter (a = 1.88) but was difficult (b = 1.01) suggesting that if a student could answer this question, they had a strong mathematical vocabulary. Importantly, discrimination levels were positive for all items indicating that the probability of getting each item correct increased with the student's skills. Based on the IRT model, the task provided a reasonable range of item difficulty and positively discriminated between skilled and less-skilled students.

The test information function graph (TIF) (see Figure A1) shows that the vocabulary measure provides the most information for students located near θ=-.2; the point where test information is greatest and estimation error is lowest (StataCorp, 2017a). This point corresponds to a score of approximately 13.5 out of 19 (see Figure A2), slightly below the mean score of all participants (M= 13.69, $S.D.$ = 2.78) indicating that the measure is well-suited for differentiating mathematical language skills amongst typically performing students. Importantly, we are not focused on struggling or gifted students in this study, so the range of questions is appropriate. If, however, we want a measure that accurately estimates ability across a larger ability range, a flatter TIF graph is recommended. Findings from the IRT analysis suggest the mathematical vocabulary task is a valid measure for first- and dual-language (i.e., English and French) learners in grades 2 and 3.

Task reliability was moderate but acceptable (Cronbach's α =0.69). It was acceptable amongst grade 2 students (Cronbach's α =0.70) but lower amongst grade 3 students (Cronbach's α =0.65). Notably, we had a small group of grade 3 students (N=58) who performed better on the task than the grade 2 students. To improve the task for future use with grade 3 students, we suggest including harder questions. These additions may improve variability and task reliably and lead to better precision estimating the mathematical vocabulary abilities of older students.

Table A1

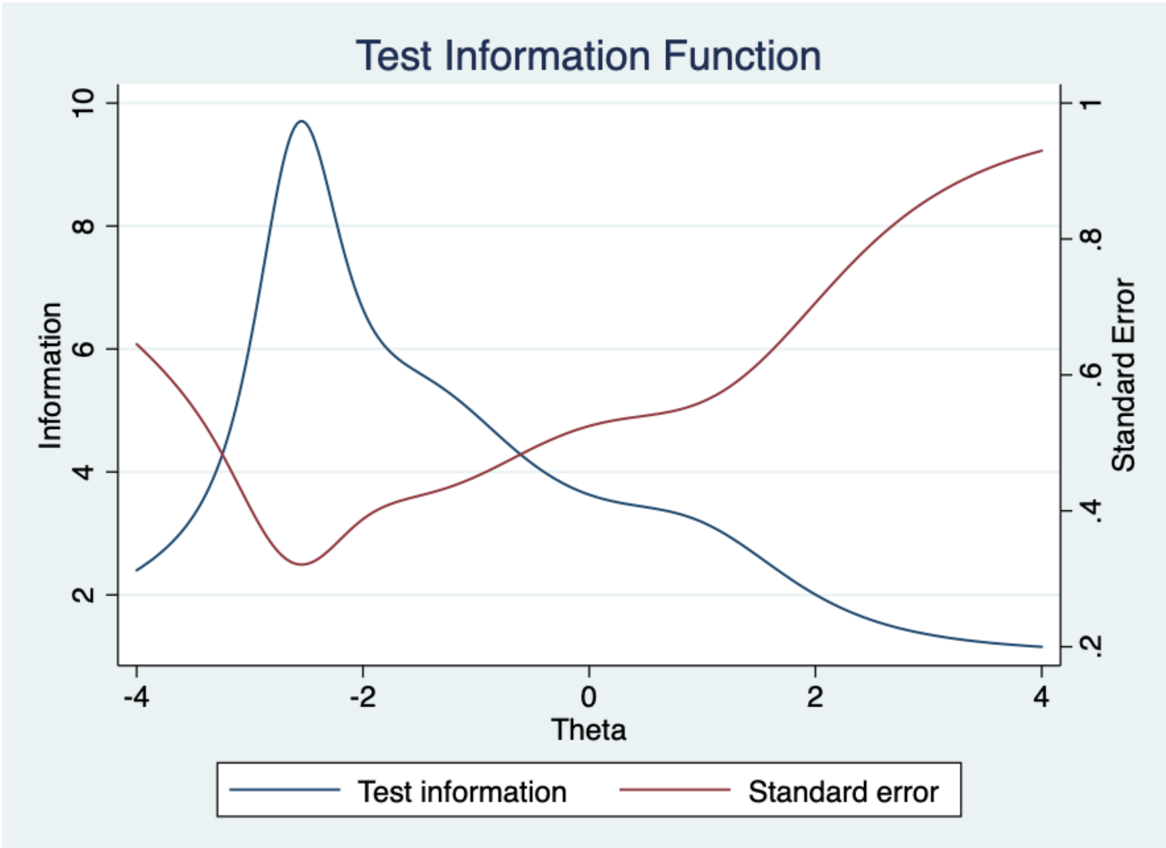
Item Response Accuracy, Discrimination, and Difficulty Parameters for the Mathematical Vocabulary Assessment (N=154)

Item #	Item detail (French and English)	Accuracy *M*	Item-total correlation *(r)*	Discrimi-nation (*a*)	Difficulty (*b*)
1	Which side has **fewer** dots	.90	.33**	1.24	-2.16
2	Point to the dot that is **between** the cat and the tree	.98	.10	.81	-5.21
3	Point to the picture that shows **half** of a pizza	.95	.31**	1.30	-2.86
4	How many teddy bears **altogether**	.63	.18*	.28	-1.95
5	Point to the square that is **nearest** to the tree	.99	.16*	4.52	-2.58
6	Which digit is in the **tens column**	.56	.45**	.78	-.39
7	How many **fewer** dogs does the girl have	.55	.51**	.82	-.30
8	Point to the **odd number**	.85	.47**	2.11	-1.35
9	What picture shows children **sharing** cookies **equally**	.86	.45**	1.65	-1.59
10	Touch the dot that is to the **left** of the tree	.86	.37**	.83	-2.43
11	Point to the person that is **fifth** in line	.95	.27**	1.95	-2.27
12	Point to the **cube**	.95	.29**	1.12	-3.18
13	What sign means **less than**	.39	.37**	.46	1.02
14	Which digit is in the **hundreds column**	.62	.50**	1.12	-.54
15	What operation gives you a **sum**	.43	.51**	1.09	.32
16	What number represents **one quarter**	.58	.53**	1.31	-.33
17	Point to the **cylinder**	.87	.35**	1.06	-2.14
18	What operation gives you a **difference**	.23	.56**	1.98	.99
19	Point to the set of numbers that are in **ascending order**	.53	.28**	.45	-.24
	All items	.71[1]	1.00		

Note. M accuracy on all items Range = 0.32 to 1.00; Target vocabulary words are bolded; *p<.05, **p<.01

Figure A1

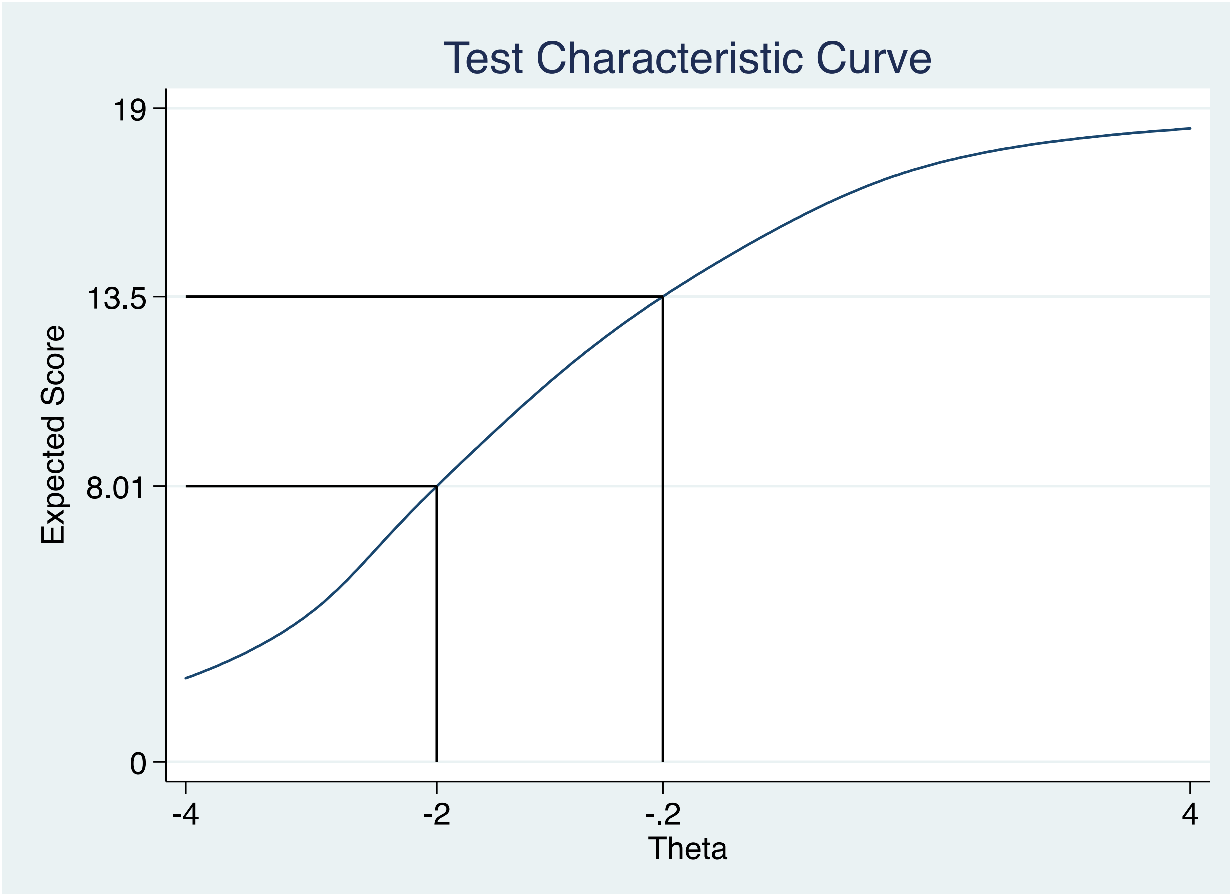
Test Information Function for the Mathematical Vocabulary Task



Note. Theta is the latent trait continuum representing mathematical vocabulary skills. A higher theta indicates more skilled performance.

Figure A2

Test Characteristic Curve for the Mathematical Vocabulary Task



Note. Predicted scores corresponding with latent trait values θ = -2 and θ = -.2 are 8 and 13.5 respectively. Maximum score is 19.

Appendix B

Evaluation of Mathematical Orthography Measure

The mathematical orthography task was modelled on the Grades 7 and 8 symbol decision task developed by Headley (2016). To create the children's version of the task, we needed to select age-appropriate math symbols and transform the symbols to create non-conventional stimuli. To select stimuli, we reviewed the Ontario curriculum to identify when the symbols are first introduced (see Table B1). We designed the task so stimuli would be age-appropriate, range in familiarity, and reflect the different types of symbols students work with in school.

In the children's mathematical symbol decision task, students judge whether a string of mathematical symbols is conventional/readable (e.g., 3 – 1 = 2) or non-conventional/non-readable (e.g. 3 _ 1 = 2). Half the stimuli are orthographically correct (i.e., conventional) and half are not (i.e., non-conventional). There are 44 randomly presented trials, 22 conventional and 22 non-conventional. See Table B2 for stimuli and item accuracy for the present study.

The task was piloted with 417 students in grades 1-3. Internal consistency of responses to conventional and non-conventional stimuli were measured separately as this approach aligns most closely with task scoring. Internal consistency was good across all conventional items (Cronbach's α = 0.81) and acceptable across all non-conventional items (Cronbach's α = 0.74). Items represented a range of mathematical symbols and types of transformations, and thus, were quite variable. Nonetheless, the reliability findings indicate that students responded consistently to both the conventional items and the non-conventional items. To determine if there was an effect of language on scoring, we considered scores of second grade students. Using a series of two-group (Instructional language: English, French) ANOVAs we found that dPrime scores did not differ amongst students learning mathematics in French compared to students learning mathematics in English, $F(1,264) = .59$ $p = 0.44$. Similarly, no differences were observed in conventional (p = .43) and non-conventional accuracy scores (p = .94).

To see if students responded similarly to the different types of symbols, a principal component analysis with a varimax rotation was conducted to see which conventional items grouped together. If the task is well-designed, students should respond consistently to similar types of symbol stimuli. That is, factors should reflect the types of symbols. As shown in Table B3, the analysis yielded a 7-factor solution. Stimuli involving fractions grouped together, relational symbols grouped together, and time, temperature, and money abbreviations grouped together. Interestingly, three of the factors related to equations: familiar addition, less familiar multiplication and division, and least familiar equation formats (i.e., 3 = 2 + 1 versus 2 + 1 = 3). The one subtraction item (3 – 1 = 2) did not load strongly on any factor suggesting that students may be responding differently to subtraction equations than either addition or multiplication. However, with only one subtraction question, it is difficult to draw any meaningful conclusions about how familiar students are with subtraction equations. Importantly, the principal component analysis suggests our stimuli capture a range of symbol exposure and further, student responses are consistent within symbol classes thus further validating the task.

Finally, to identify any problematic test items, item theory was used to examine how each individual test item related to the corresponding trait or skill (i.e., conventional orthography, non-conventional orthography) and how the group of test items together related to the skill. Two-parameter (2-PL) item response models (IRT) were run using Stata/SE 15.1 (StataCorp, 2017b). Two conventional test items were identified (2 = 3 - 1, 3 = 2 + 1) that posed a problem. In spite of the fact that scale performance positively correlated with item performance for the addition item, the discrimination values for both items were negative (see Table B4). The probability of getting these items correct decreased with the student's skill. We therefore excluded these items and their corresponding non-conventional items from the final scoring. As anticipated all items were relatively easy with *b* values either below zero or close to zero. We did not have to exclude any items based on item difficulty.

Table B1

Mathematical Symbols Organized by Type, Grade Introduced, and Applications Across the Grades

Symbol	Classification	Symbol Introduction	Symbol use across the grades		
			Grade 1	Grade 2	Grade 3
+	operation	Grade 1	Single digit addition	Two-digit addition	Three-digit addition
–	operation	Grade 1	Single digit subtraction	Two-digit subtraction	Three-digit subtraction
×	operation	Grade 2 (limited)		Multiplication as repeated addition (e.g. 2+2+2 = 3X2)	Multiply one-digit numbers
÷	operation	Grade 2 (limited)		Division as equal sharing	Divide two-digit numbers by single digit (e.g. 49÷7=7)
=	relation	Grade 1 in equations	Equality with concrete items	Equality with symbols (e.g., 7+3=4+ __)	Equality with symbols (e.g., 25-4=15+__)
< and >	relation	See note Grade 2/3	Use words to relate numbers to anchors of 5 and 10	Appears in few texts	More widely used
: (time)	abbreviation	Grade 1	Write to the hour and half-hour	Write to the quarter hour	Write 5-minute intervals
¢	abbreviation	Grade 2	Describe coins, state their value	Represent money amounts to 100¢, Use the ¢ symbol	
$	abbreviation	Grade 3	Describe coins, state their value		Represent money to $10, use the $ symbol
°C	abbreviation	Grade 2	Use non-standard units to describe temperature	Record temp using °C,	Record temp using °C, identify temp. benchmarks
3-digit number	Concept	Grade 2	Whole numbers <50	Whole numbers to 100	Whole numbers to 1000numbers
Fractions	Concept	Grade 4	Fractional names	Fractional names	Fractional names

Note. The < and > are not described explicitly in the curriculum document. However, comparing quantities is a curriculum expectation from grade 1 on and these symbols are included in workbooks from grades 2 and 3 (*Jump Math Student Essentials C*, 2012; Morrow & Connell, 2010).

Table B2

Item Accuracy Across All Participants (N=151) on the Symbol Decision Task

Item	Conventional	*M.*	*S.D.*	Item	Non-conventional	*M.*	*S.D.*
1	430	.77	.43	2	026	.59	.49
3	$\frac{1}{2}$	.59	.49	4	$\begin{array}{c}1\\2\\\hline\end{array}$	.64	.48
5	$\frac{1}{3}$	.60	.49	6	$\overset{2}{\underset{3}{\equiv}}$	.77	.42
7	$3.00	.61	.49	8	$2;00	.82	.39
9	4>2	.67	.47	10	3 ∨ 2	.89	.31
11	2 + 2 = 4	.96	.19	12	2 + 2 ‖ 4	.77	.43
13	3<5	.67	.47	14	4<<5	.87	.34
15	2<3	.65	.48	16	1∧3	.90	.30
17	3 + 2 = 5	.94	.25	18	3 + 2 ≡ 5	.44	.50
19	2 + 1 = 3	.88	.32	20	+ 2 1 = 3	.80	.40
21	3 – 1 = 2	.81	.39	22	4_1 = 3	.79	.41
23	2 x 3 = 6	.75	.44	24	2 *x* 4 = 8	.69	.46
25	3 x 2 = 6	.77	.43	26	4 ~ 2 = 8	.89	.31
27	3 = 2 + 1	.51	.50	28	4 = 3 ⊕ 1	.79	.41
29	8 ÷ 2 = 4	.64	.48	30	8·ı·2 = 4	.75	.44
31	7 + 2 =9	.90	.30	32	$1^{+}2 = 3$	.66	.48
33	2 = 3 - 1	.49	.50	34	2 = 3 ı 1	.82	.38
35	35¢	.69	.46	36*	3¢5	.81	.40
37	15¢	.75	.43	38	20©	.76	.43
39	2:30	.77	.42	40	13:0	.58	.50
41	1:00	.78	.42	42	3..00	.82	.39
43	15^{o}C	.67	.47	44	2^{o}0C	.66	.47

Note. Conventional and non-conventional stimuli are shaded for easy identification.

Table B3

Exploratory Factor Analysis of Conventional Stimuli for the Symbol Decision Task (N=417)

Notes.

#	Stimuli	Primary factor loadings for symbol classes						
		Time, temp., money	> and <	× and ÷	Fraction	+	=	
7	\$3.00	**.63**	.00	.25	.08	.09	-.05	-.27
35	35¢	**.67**	.06	.10	.09	.09	-.09	.01
37	15¢	**.72**	.14	.05	-.00	.11	.08	.00
39	2:30	**.73**	-.03	.12	-.03	-.10	-.05	.21
41	1:00	**.75**	.00	.06	-.04	-.16	-.05	.14
43	15ºC	**.56**	.31	.12	.09	.16	.06	-.25
9	4>2	.09	**.79**	.05	.06	.06	-.05	-.01
13	3<5	.03	**.77**	.12	.14	.07	.02	.03
15	2<3	.09	**.81**	.05	.00	-.02	-.03	.13
23	2 x 3 = 6	.09	.07	**.73**	.04	.20	-.03	.00
25	3 x 2 = 6	.14	.06	**.76**	.05	.06	-.01	.09
29	8 ÷ 2 = 4	.20	.10	**.68**	.02	-.02	-.00	.09
3	$\frac{1}{2}$	.07	.06	.00	**.89**	-.03	.05	.02
5	$\frac{1}{3}$	.07	.14	.11	**.85**	.06	.01	.05
11	2 + 2 = 4	-.04	-.06	.07	.08	**.62**	-.04	.03
19	2 + 1 = 3	.09	.07	.04	-.14	**.57**	.32	-.01
31	7 + 2 = 9	.04	.11	.10	.01	**.57**	-.05	.19
27	3 = 2 + 1	-.05	-.02	.03	.13	-.13	**.74**	.34
33	2 = 3 - 1	-.09	-.05	-.07	-.02	.13	**.77**	-.16
17	3 + 2 = 5	.08	.08	-.01	.20	.40	-.08	**.56**
21	3 – 1 = 2	-.00	.10	.23	-.09	.13	.17	**.61**
1	430	.44	-.01	.01	.15	.01	-.09	.36
% of variance explained		13.9	9.3	8.1	7.7	7.0	6.0	5.7

1. Factor loadings greater than 0.50 are bolded.
2. Factors were extracted using principal component analysis and a Varimax rotation. Factors converged in 6 iterations.
3. 57.2% of total variance is explained in the 7-factor solution.

Table B4

Item Response Parameters for Conventional and Non-conventional Test Items

Conventional test items			Non-conventional test items		
Stimuli	Discrimination Value *(a)*	Difficulty value *(b)*	Stimuli	Discrimination Value *(a)*	Difficulty value *(b)*
2 = 3 - 1	-.28	-.83	2 = 3 ı 1	1.57	-1.35
3 = 2 + 1	-.06	-3.92	4 = 3 ⊕ 1	.91	-1.90
2 + 2 = 4	.35	-8.06	2 + 2 ıı 4	1.58	-.93
3 – 1 = 2	.45	-3.30	4_1 = 3	1.03	-1.56
2 + 1 = 3	.47	-4.94	+ 2 1 = 3	1.33	-1.25
$\frac{1}{2}$	.49	-1.07	$\begin{matrix}1\\ \underline{2}\end{matrix}$	.88	-.67
7 + 2 =9	.60	-3.45	$1^{+}2 = 3$	1.70	-.59
$\frac{1}{3}$	.65	-.68	$\begin{matrix}2\\ =\\ 3\end{matrix}$	1.45	-1.17
4>2	.85	-.51	3 ∨ 2	1.33	-1.82
3<5	.80	-.46	4<<5	1.30	-1.73
2<3	.82	-.43	1∧3	1.61	-1.66
3 + 2 = 5	.86	-3.21	3 + 2 ≡ 5	.57	.50
430	.98	-1.26	026	.79	-.67
2 x 3 = 6	1.00	-1.28	2 *x* 4 = 8	.51	-1.41
3 x 2 = 6	1.15	-1.21	4 ~ 2 = 8	1.05	-2.23
8 ÷ 2 = 4	1.05	-.64	8·ı·2 = 4	.31	-2.10
$3.00	1.60	-.59	$2;00	1.38	-1.11
15^{o}C	1.66	-.46	2^{o}0C	.62	-1.16
35¢	2.01	-.76	3¢5	1.38	-1.37
15¢	2.30	-.88	20©	.84	-1.48
2:30	2.12	-.94	13:0	.72	-.82
1:00	2.17	-1.17	3..00	1.59	-1.03

Note. Higher ‘a’ values show greater discrimination between high- and low-scoring students.

Appendix C

Table C1

Performance on Measures (Before Standardization) for Students who Received Instruction in English vs. French

	English instruction				French instruction			
	Mean(*SD*)	Min, Max	$Z_{skewness}$	$Z_{kurtosis}$	Mean(*SD*)	Min, Max	$Z_{skewness}$	$Z_{kurtosis}$
Receptive vocabulary	43.56(12.80)	2, 60	-4.76	2.45	27.58(4.78)	11, 35	-4.65	3.12
Working memory								
Digit forward	8.20(2.22)	1, 14	.31	1.36	7.40(2.09)	2, 12	.80	-.38
Digit backward	5.80(1.44)	1, 9	-1.26	.60	5.47(1.20)	4, 9	2.48	.16
Spatial span	9.67(3.78)	1, 23	.46	2.47	8.95(3.04)	1, 20	.24	3.07
Quantitative skills								
Number comparison	.97(.20)	.37, 1.38	-.82	.25	.86(.16)	.43, 1.32	.75	1.49
Subitizing	1.66(.31)	1, 2.78	2.25	1.84	1.48(.30)	.73, 2.14	-.41	.13
Mathematical vocabulary	.69(.15)	.32, 1.00	-.54	-1.06	.74(.15)	.42, 1.00	-.87	-.91
Mathematical orthography	1.75(.77)	.16, 3.69	.49	-.13	1.17(.63)	-.35, 2.79	1.00	.57
Word-problem solving	5.09(2.55)	0, 10	-.02	-1.77	4.97(2.26)	1, 10	1.07	-1.35
Arithmetic fluency	19.10(9.11)	1, 44	1.90	-.34	15.49(9.19)	2, 46	4.15	2.07
Word reading	29.32(8.21)	5, 40	-3.55	.45	22.81(2.25)	15, 26	3.60	3.05

Note. Scores reflect performance on measures before standardization. There were 79 students who received instruction in English (21 second graders and 58 third graders) and 81 students (second graders) who received instruction in French

Appendix D

Table D1

Hierarchical Regression of Word-problem Solving Performance Without Imputation

	First-Language Learners [a]				Second-Language Learners [b]			
	Beta	*t*	*p*	*Unique r^2*	*Beta*	*t*	*p*	*Unique r^2*
Model 1								
Receptive Vocabulary	**.374*** **	**4.149*** **	**<.001**	**.136**	**.491*** **	**4.210*** **	**<.001**	**.200**
Working Memory	**.249** **	**2.708** **	**.008**	**.058**	.249	1.971	.055	.044
Quantitative Skill	**.278** **	**3.046** **	**.003**	**.073**	.163	1.283	.206	.018
ΔR^2	**.352*** **				**.515*** **			
Model 2								
Receptive Vocabulary	**.196***	**2.160***	**.043**	**.030**	**.415** **	**2.847** **	**.007**	**.094** **
Working Memory	.168	1.963	.053	.025	.248	1.938	.059	.043
Quantitative Skill	.173	1.974	.052	.025	.152	1.131	.265	.015
Mathematical Vocabulary	**.287** **	**2.862** **	**.005**	**.053**	.059	.406	.687	.002
Mathematical Orthography	**.229***	**2.493***	**.015**	**.040**	.096	.785	.362	.007
ΔR^2	**.130*** **				.011			

Note: Squared semi-partial correlations indicate unique r^2 within that specific model tested. $p < .05$*, $p < .01$**, $p < .001$***.

[a] In Model 1, $R^2 = 35.2\%$, $F(3, 82) = 14.87$, $p < .001$; in Model 2, $R^2 = 48.2\%$, $F(5, 80) = 14.92$, $p < .001$ for first-language learners.

[b] In Model 1, $R^2 = 51.5\%$, $F(3, 43) = 15.25$, $p < .001$; in Model 2, $R^2 = 52.6\%$, $F(5, 41) = 9.11$, $p < .001$ for second-language learners.

Table D2

Hierarchical Regression of Arithmetic Fluency Without Imputation

	First-Language Learners [a]				Second-Language Learners [b]			
	Beta	*t*	*p*	*Unique r^2*	*Beta*	*t*	*p*	*Unique r^2*
Model 1								
Working Memory	**.270****	**3.099****	**.003**	**.070**	-.161	-1.190	.241	.019
Quantitative Skill	**.514*****	**5.887*****	**<.001**	**.251**	**.703*****	**5.210*****	**<.001**	**.367**
ΔR^2	**.398*****				**.406*****			
Model 2								
Working Memory	**.210***	**2.409***	**.018**	**.039**	-.177	-1.382	.174	.023
Quantitative Skill	**.437*****	**4.912*****	**<.001**	**.162**	**.564*****	**4.113*****	**<.001**	**.203**
Mathematical Vocabulary	**.220***	**2.270***	**.026**	**.035**	**.356****	**2.712****	**.010**	**.088**
Mathematical Orthography	.076	.833	.407	.005	-.070	-.586	.554	.004
ΔR^2	**.057***				**.090***			

Note: Squared semi-partial correlations indicate unique r^2 within that specific model tested. $p < .05$*, $p < .01$**, $p < .001$***.

[a] In Model 1, $R^2 = 39.8\%$, $F(2, 83) = 27.44$, $p < .001$; in Model 2, $R^2 = 45.5\%$, $F(4, 81) = 16.91$, $p < .001$ for first-language learners.

[b] In Model 1, $R^2 = 40.6\%$, $F(2, 44) = 15.01$, $p < .001$; in Model 2, $R^2 = 49.6\%$, $F(4, 42) = 10.34$, $p < .001$ for second-language learners.

Table D3

Hierarchical Regression of Word Reading Without Imputation

	First-Language Learners [a]				Second-Language Learners [b]			
	Beta	*t*	*p*	*Unique r^2*	*Beta*	*t*	*p*	*Unique r^2*
Model 1								
Receptive Vocabulary	**.252***	**2.515***	**.014**	**.062**	**.396***	**2.631***	**.012**	**.121**
Working Memory	**.235***	**2.294***	**.024**	**.052**	.098	.623	.537	.007
Quantitative Skill	.184	1.812	.074	.032	.064	.427	.672	.003
ΔR^2	**.197****				**.228****			
Model 2								
Receptive Vocabulary	.146	1.335	.186	.017	.233	1.254	.217	.027
Working Memory	.184	1.783	.078	.030	.096	.621	.538	.007
Quantitative Skill	.131	1.240	.219	.014	.052	.341	.734	.002
Mathematical Vocabulary	.130	1.077	.285	.011	.094	.522	.604	.005
Mathematical Orthography	.184	1.655	.102	.026	.239	1.609	.115	.044
ΔR^2	.050				.055			

Note: Squared semi-partial correlations indicate unique r^2 within that specific model tested. $p < .05$*, $p < .01$**.

[a] In Model 1, $R^2 = 19.7\%$, $F(3, 82) = 6.70$, $p < .001$; in Model 2, $R^2 = 24.7\%$, $F(5, 80) = 5.25$, $p < .001$ for first-language learners.

[b] In Model 1, $R^2 = 22.8\%$, $F(3, 44) = 4.33$, $p = .009$; in Model 2, $R^2 = 28.3\%$, $F(5, 42) = 3.32$, $p = .013$ for second-language learners

Appendix E

Table E1

Hierarchical Regression of Arithmetic Fluency Showing Suppressor Patterns

	First-Language Learners [a]				Second-Language Learners [b]			
	Beta	*t*	*p*	*Unique r^2*	*Beta*	*t*	*p*	*Unique r^2*
Model 1								
Receptive Vocabulary	.136	1.720	.086	.018	-.091	-.731	.465	.006
Working Memory	**.316*** **	**3.820*** **	**<.001**	**.092*** **	.100	.723	.470	.007
Quantitative Skills	**.456*** **	**5.614*** **	**<.001**	**.196*** **	**.590*** **	**4.460*** **	**<.001**	**.269*** **
Averaged ΔR^2 across imputations	**.415*** **				**.384*** **			
Model 2								
Receptive Vocabulary	.020	.245	.807	.000	-.369**	-2.889**	.004	.076**
Working Memory	**.254** **	**3.079** **	**.002**	**.055** **	-.041	.326	.745	.001
Quantitative Skills	**.391*** **	**4.835*** **	**<.001**	**.132*** **	**.466*** **	**3.995*** **	**<.001**	**.156*** **
Mathematical Vocabulary	**.224* **	**2.406* **	**.016**	**.033* **	**.551*** **	**4.390*** **	**<.001**	**.165*** **
Mathematical Orthography	.113	1.264	.207	.009	.031	.256	.799	.001
Averaged ΔR^2 across imputations	**.059** **				**.183*** **			

Note. Squared semi-partial correlations indicate unique r^2 within that specific model tested. $p < .05$*, $p < .01$**, $p < .001$***.

[a] Model R^2s across imputations ranged from 40.3% to 42.7% in Model 1, $F(3, 99)$ = from 22.26 to 24.60, *p*s < .001, and in Model 2, from 45.3% to 49.1%, $M = 45.2\%$, $F(5, 97)$ = from 16.09 to 18.69, *p*s < .001 for first-language learners.

[b] Model R^2s across imputations ranged from 36.2% to 41.4% in Model 1, $F(3,53)$ = from 10.01 to 12.46, *p*s < .001, and in Model 2, from 54.9% to 61.1%, $M = 56.7\%$, $F(5, 51)$ = from 12.44 to 16.01, *p*s < .001 for second-language learners.

References

Headley, M. G. (2016). *What is Symbolic Mathematics Language Literacy? A Multilevel Mixed Methods Study of Adolescents in a Middle School* [Electronic Dissertation, University of Cincinnati, Education]. http://rave.ohiolink.edu/etdc/view?acc_num=ucin1470045155

Jump Math Student Essentials C. (2012). Jump Math.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, *65*(23), 2276–2284. https://doi.org/10.2146/ajhp070364

Morrow, P., & Connell, M. M. (2010). *Math Makes Sense Practice and Homework Book 3* (M. Czukar, Ed.). Pearson Canada Inc.

Purpura, D. J., & Reid, E. E. (2016). Mathematics and language: Individual and group differences in mathematical language skills in young children. *Early Childhood Research Quarterly*, *36*, 1–10. https://doi.org/10.1016/j.ecresq.2015.12.020

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, *114*(3), 552–566. https://doi.org/10.1037/0033-2909.114.3.552

StataCorp. (2017a). *STATA Item Response Theory Reference Manual, Release 15*.

StataCorp. (2017b). *Stata Statistical Software: Release 15.* StataCorp LLC.