2022

A dimensional approach in investigating early cognitive predictors of language, reading, and mathematical achievements

Lisa Archibald

Title: Early cognitive predictors of language, reading, and mathematics outcomes in the primary grades

Introduction

Language, reading, and mathematics are foundational skills for academic success. Perhaps not surprisingly then, it is well-established that impairments in language (developmental language disorder (DLD); previously called specific language impairment), reading (dyslexia), and/or mathematics (dyscalculia) will have negative impacts on academic success, and professional and social outcomes later in life (e.g., Clegg et al., 2005; Geary et al., 2012; Ritchie & Bates, 2013). It is widely agreed that better methods of early identification are needed in order to optimize long-term outcomes for children who are struggling academically. One challenge, however, is that academic abilities can only be evaluated when children have had the opportunity to learn them, which means it takes time for the learning difficulty to be identified. This highlights a need to identify precursor skills that are less dependent on formal schooling and can be measured prior to formal instruction. Another challenge in trying to understand learning difficulties in language, reading, and mathematics is that a majority of research has focused on one domain (i.e., either language, reading, *or* mathematics) while ignoring the others (e.g., Butterworth, 2008; Morris et al., 1998) despite considerable evidence showing overlaps across these domains (Archibald et al., 2013, 2019; Peters & Ansari, 2019; Peterson et al., 2021; Shrank et al., 2014). It would follow that a student with difficulties in one domain is likely to have difficulties in other domains. The purpose of the present study was to examine how cognitive skills considered precursors of language, reading, and mathematics, measured in kindergarten children, predict academic outcomes (indexed by report card marks) over the primary grades.

The question arises regarding how early children at risk for academic struggles can be identified which requires consideration of how each skill is learned. Reading (Rupley et al., 2009) and mathematics (Doabler & Fien, 2013) require explicit, academic instruction, whereas language skills develop largely incidentally prior to school entry (Ellis, 2015). One approach—the 'wait-to-fail' approach—only intervenes *after* a student has failed to learn. However, response to intervention is better when it starts earlier. Thus, a more promising approach investigated in the present study was the predictive utility of widely recognized cognitive precursors of reading and mathematics. These cognitive precursors could be measured much earlier as they are not as dependent on formal instruction. On the other hand, language development is relatively stable by age 4 (Dale et al., 2003). At school entry, then, it is possible to use direct language measures known to be sensitive to individual differences in order to identify those with – or at risk for – language learning difficulties (Law et al., 2009). In order to identify potential cognitive precursors of reading and mathematics and discriminative measures of language, we drew from research investigating foundational skills in these areas.

Foundations of Language

Oral language, including knowledge of vocabulary and syntactic structure, is an important predictor of academic abilities (Foorman et al., 2015; Ladd et al., 2012). Indeed, successfully learning across academic subjects would depend, to a large extent, on children's language skills given that oral language is the medium for teaching and learning (Foorman et al., 2015; Ladd et al., 2012; Rubin et al., 2012). It is well-established that vocabulary knowledge uniquely predicts learning (Lee, 2008), even when age, gender, language background, and non-verbal cognitive abilities are controlled (Schuth et al., 2017).

Another task that has been widely used to index language skills is sentence recall given that this task is a multi-faceted linguistic task engaging virtually all aspects of language processing. In fact, sentence recall is considered a clinical marker of DLD across the lifespan (Archibald & Joanisse, 2009; Poll et al., 2010). Further, Klem et al. (2015) found that language skills at age 4 characterized by vocabulary knowledge, sentence recall, and grammar skills were related to later improvements in language skills at age 6.

Foundations of Reading

Several studies have investigated the cognitive foundations of early reading skills, characterized namely by word recognition or the ability to read single words. Well-established predictors of such reading skills include phonological awareness, letter name, letter sound, and rapid automatized naming (RAN) (Melby-Lervåg et al., 2012; Norton & Wolf, 2012; Hulme & Snowling, 2014). In addition, each task independently exerts some influence on reading difficulties (e.g., Bowyer-Crane et al., 2008; Wolf & Bowers, 1999).

Phonological awareness involves explicit knowledge and manipulation of phonological structures and processes, which is important for retrieving and blending sounds to recognize the word. Phonological awareness skills in kindergarten have been found to be one of the most stable and robust indicators of reading acquisition, predicting for example reading skills in first grade (Lonigan et al., 2008) and the presence of reading disability in second grade (Catts et al., 2001).

Letter knowledge is an important foundation for literacy development, specifically in alphabetic languages. This skill constitutes *letter-naming*, the ability to identify a letter by its name, and *letter-sound knowledge*, the ability to provide the sound associated with a letter. Letter-naming has consistently been found to be a powerful predictor of reading in early school years, and sometimes even the best single predictor (McBride-Chang, 1999; Catts et al., 2001).

Letter-sound knowledge in kindergarten has also been found to be a significant predictor of later reading and writing (Foulin, 2005; McBride-Chang, 1999), and kindergarteners who fail to master letter-sound correspondences are at risk for later reading difficulties (Storch & Whitehurst, 2002). The predictive power of letter knowledge likely stems from this skill being some of the first that parents teach their children before kindergarten (Ellefson et al., 2010).

Another robust predictor of reading development is RAN. RAN measures the ability to name a random sequence of objects, colours, letters, or digits as quickly as possible. In a recent meta-analysis, RAN was one of the best predictors of reading ability (Araújo et al., 2015). Although the mechanism linking RAN to reading remains unclear, the consensus is that RAN is a strong predictor of reading because it shares many processes and skills implicated in reading including automatic visual recognition, attention, and integration (Norton & Wolf, 2012).

Foundations of Mathematics

Recently, studies have begun to identify the precursors of mathematics (e.g., Nogues & Dorneles, 2021). Specifically, the concept of number sense embodies a number of math precursors utilized in the current study. Number sense refers to children's ability to understand the meaning of numbers and define different relationships among numbers (Clarke & Shinn, 2004; Commission on Standards for School Mathematics, 1989) and in fact, is one strand of the provincial mathematics curriculum (Ontario Ministry of Education, 2005). Based on a review of the literature, we focused on four tasks commonly employed to evaluate early numeracy: number line estimation, magnitude comparison, number naming, and arithmetic skills (basic calculations) (e.g., Hawes et al., 2019; Geary et al., 2012). Importantly, these early numerical concepts influence later ability to acquire important mathematical skills (Clarke & Shinn, 2004).

Numerical line estimation assesses children's ability to approximate a quantity by asking participants to estimate the position of a number on a number line. Many researchers have suggested that number line estimation is important for the development of mathematical ability (e.g., Booth & Siegler, 2006; Geary et al., 2009). However, it should be noted that performance on number line tasks may involve higher cognitive skills (e.g., spatial reasoning). As a result, young children tend to struggle with the complexities of the task rather than as a result of limited knowledge of numerical magnitudes (Hawes et al., 2019).

In magnitude comparison tasks, individuals compare symbolic (Arabic numbers) or non-symbolic representations (dot arrays) without counting. There is mixed evidence on whether both symbolic and non-symbolic tasks have the same predictive power. Most studies suggest a relationship between symbolic comparison, but not non-symbolic, and mathematics performance (Holloway & Ansari, 2009; Rousselle & Noël, 2007). Others suggest that although symbolic comparison accounts for unique variance in arithmetic performance, this task may not be the best predictor when compared to other cognitive skills such as working memory (Nosworthy et al., 2013). Still others have found that non-symbolic comparison predicts mathematic performance during the school year (Gilmore et al., 2010; but see Kolkman et al., 2013).

Number naming is the ability to identify an Arabic symbol (7) by its verbal name ("seven"). Clarke and Shinn (2004) found that number naming was one of the most reliable and valid measures that could be used for early identification. Further, number naming has been found to be predictive of complex arithmetic achievements later on (Moeller et al., 2011). Difficulties with the task have been found in children with dyscalculia (Rousselle & Noël, 2007).

Simple arithmetic calculations (addition and subtraction) served as a direct measure of numerical skills in this study. These arithmetic tasks have been found to be one of the few early numeracy skills that parents directly teach their children at home (Senechal & LeFevre, 2002). In a follow-up study, LeFevre et al. (2009) found that these types of home numeracy experiences are highly predictive of later mathematics acquisition. Moreover, arithmetic skills are central to mathematical learning in school, and hence, are frequently used for assessment (Nosworthy et al., 2013). Beyond imposing a specific load, general working memory demands may be tapped given the need to hold the verbal information in mind while executing the calculation.

Overlapping Nature of Language, Reading, and Mathematics

Despite evidence of separability, it must be acknowledged that the distinction between language, reading, and mathematics is far from clear cut. Indeed, all academic skills are correlated (Shrank et al., 2014) and there are corresponding high rates of co-occurrence between DLD and dyslexia, dyslexia and dyscalculia, and DLD and dyscalculia (Archibald et al., 2013).

It is easy to see how language and reading are inextricably linked. According to the Simple View of Reading (Hoover & Gough, 1980; Ehri, 2005), one of the most influential models of reading, reading comprehension is the product of language comprehension (meaning-based skills mapping onto the language cognitive predictors) and reading decoding (code-related skills akin to the reading cognitive predictors). Perhaps not surprisingly then, limited vocabulary knowledge is not only found in children with DLD (McGregor er al., 2013; cf. Spaulding et al., 2006), but also those with poor reading comprehension (Nation & Snowling, 2004), highlighting a potentially shared cognitive risk factor. Further, sentence recall – a clinical marker of DLD – has also been used to differentiate learning profiles in children (Redmond, 2005; Archibald et al., 2013, 2019) and is predictive of risk for reading difficulty (Catts et al., 2001). Further, phonological awareness difficulties, traditionally linked to reading disorders (Kuppen &

Goswami, 2016), have also been implicated in DLD (Bishop & Snowling, 2004) as well as dyscalculia (De Smedt et al., 2010).

Even mathematical skills may be not as readily distinguishable from language and reading as previously thought. Mathematical tasks involve varying levels of language demands (Cross et al., 2019). Verbal demands are generally low in number line estimation and magnitude comparison as they rely on the visual and analogue frames, whereas number naming and arithmetic tasks are thought to have a high verbal load as they rely on verbal word frame representations. Further, the importance of reading comprehension skills cannot be overlooked in solving mathematics word problems. More broadly, math concepts inherently depend on language skills (e.g., understanding basic concepts of 'more' vs. 'less'). Given the relation between mathematics with language and literacy, this could help explain why children with co-morbid math and reading disorders struggle with number naming (Geary et al., 2000) or why children with DLD have difficulties with magnitude comparison (Donlan et al., 2007).

Taken together, language, reading, and mathematical skills overlap, albeit to varying degrees. As a result, there are mounting calls for researchers to adopt a dimensional approach to study language, reading, and mathematical rather than thinking of them as independent categories (Peters & Ansari, 2019; Peterson et al., 2021). Therefore, in the current work, we examined predictors and outcomes across domains to better understand the structure of academic skills including unique and shared variance to language, reading, and mathematics. In turn, this would help us understand the basis, identification of, and intervention for struggle learning.

The Current Study

The literature reviewed thus far provides strong evidence for overlapping relationships between language, reading, and mathematics. Nevertheless, work in this area has tended to focus on specific domains. Thus, the aim of this study was to identify early cognitive predictors *across domains* that could indicate future academic skills *across domains,* and how predictors might change over time. In this study, a large sample of children was followed over three years, from kindergarten to grade 2. In kindergarten, students completed experimental tasks in the areas of language (meaning-focused skills), literacy (code-focused skills), and numeracy. At 1- and 2-year follow up corresponding to grades 1 and 2, respectively, we used children's report card marks in language, reading, and mathematics as academic measures. First, a data-driven approach was used to explore the structure of academic skills. Then, informed by the structure of academic skills, a subsequent goal was to use kindergarten cognitive precursors to predict later academic outcomes. Having a good understanding of overlapping precursors predictive of academic skills will help us identify children with difficulties earlier in the long-term.

Methods

Participants

Our study used a multi-wave approach in which a total of 16 schools (2 rural) in Ontario, Canada participated in the study over 7 years (2013-2019), with 6-9 schools participating in new recruitment each year. Participants were followed from kindergarten to Grade 3. In Ontario, kindergarten is a 2-year program with those entering the second year of the program turning 5 years of age at some point during that September's calendar year. All children in the second year of kindergarten in participating schools received an invitation to participate. A total of 767 children (378 males; 373 females; 29 unidentified) completed at least one direct data collection session (i.e., kindergarten assessment). Data collected included our experimental screening measures and the school board's phonological awareness measure for kindergarteners and standardized tests and report card grades for Grades 1 to 3. However, due to various factors including Board-level policy changes, labor shortages and disruptions, and pandemic restrictions, not all data were available for all measures. In particular, standardized tests and Grade 3 report card grades are not analyzed in the current study due insufficient data. Thus, the current study focuses on the 563 children (388 males; 267 females; 8 unidentified) who completed our kindergarten battery in the Spring of the second year of their kindergarten program, and for whom some report card grades were reported, summarized in Table S1. We have limited information available about our sample characteristics as the demographic survey was discontinued after the first year of a study. The survey unfortunately did not ask about diagnoses or additional services, nor did the school share this information with us. As per parent report (n = 101), children were rated to be 'good' on average at: counting and recognizing numbers; letter names; number relationships; quantity concepts; understanding patterns. Children were rated to be 'satisfactory' on average at: letter sounds; meaning of written words. Our sample was largely monolingual English with a high socioeconomic status, aligning with characteristics that have been found in our previous studies with cohorts from this School Board (MASKED).

In addition, a subset of 25 kindergarten participants completed the kindergarten measures 1 month apart in Year 1 of the study in order to assess the reliability of kindergarten measures.

Procedure

All participants completed the study's *kindergarten assessment* between March and June of their year 2 kindergarten program. The screening was administered individually in a 30–40-minute session in a quiet room in the child's school by a trained research assistant. Training included viewing a video with step-by-step explanations and practicing scoring exercises for the sentence recall, letter knowledge, and number naming tasks. Each set of practice questions was reviewed for any discrepancies to ensure 100% accuracy between testers. School board data included the phonological awareness measure and report card grades for Grades 1 to 2.

Kindergarten Assessment

The kindergarten measures consisted of eight tasks: *vocabulary, sentence recall, phonological awareness, letter knowledge, rapid colour naming, number line estimation, dot and symbol magnitude comparison, number naming,* and *arithmetic skills*. All tasks except the magnitude comparison tasks were presented in a fixed order to control for order effects. The magnitude comparison tasks were counterbalanced yielding four possible orders assigned to each child at random (1) dot comparison, symbol comparison, the remaining tasks, (2) symbol comparison, dot comparison, the remaining tasks, (3) the remaining tasks, dot comparison, symbol comparison, (4) the remaining tasks, symbol comparison, dot comparison. The phonological awareness measure was administered separately at another time.

Vocabulary. The vocabulary task was from the RRST v7.5 (LDAA, 2011). The child was asked to name each of 10 coloured pictures presented in a 2 by 5 grid. The number of correct responses out of 10 was recorded. For items that had more than one correct response (e.g., broom/mop), a correct score was given for any response that was deemed acceptable from the list of correct answers. The child was encouraged to guess before indicating they did not know the name of a picture. All “I don’t know” responses were scored as zero.

Sentence Recall. The sentence recall task was adapted from Redmond (2005). It consisted of 16 sentences (9-12 words each) with an equal number of active and passive sentences. The sentences were presented via a digital audio recording of an adult female speaking over headphones. Children were asked to repeat each sentence verbatim immediately after hearing it. Sentences were scored online by the research assistant with a 2 (correct), 1 (three or fewer errors), or 0 (more than four errors or no response) for a total possible score of 32.

Phonological Awareness. The Phonological Awareness Screening Tool was a bespoke measure developed by the School Board's Speech and Language Services. This task consisted of 6 subtests requiring children to recognize rhymes, produce rhymes, combine sounds, identify sounds, segment sounds, and delete sounds and syllables. There were practice items for each subtest. Children completed all items except in the deletion task where testing was discontinued after 6 consecutive 0 scores. The number of correct items out of 33 was recorded.

Letter Sound. Children were asked to name the letter for the 26 letters of the alphabet presented in upper- and lower-case letters. The child was first presented with a card with the 26 upper- then lower-case letters of the alphabet randomly arranged in two lines on a white background and asked to name each letter. For some letters, multiple pronunciations were scored as correct (e.g., "Q" was scored as correct if pronounced as "kw", "ku" or "k"). Children were encouraged to try their best before saying "I don't know". In all cases, children were given ample time to respond, and time was not recorded. Correct responses were tallied out of 26 for each of upper case letter names and lower case letter names and the average was used for data analysis.

Letter Name. Following the letter sound task, children were presented with the same two cards, but asked to name each letter. There was again no time restrictions and the same scoring procedure was applied.

Colour Rapid Automatized Naming. We employed a nonalphanumeric RAN task (colours) designed based on our previous study (MASKED) because letter-name knowledge is not required to complete this task quickly and accurately, an important consideration for kindergarten children. A card consisting of a 5 x 10-item grid of 2 cm square coloured boxes (red, yellow, green, blue) arranged in random order was presented and the child was asked to verbally and serially name aloud the colour of the boxes in the grid as quickly and accurately as possible. Self-corrections were scored as correct. Before starting, a practice page containing four coloured blocks (red, yellow, green and blue) was presented. The time required to name all stimuli (in seconds) was recorded and used for data analysis.

Number Line Estimation. This task is described more fully in Hawes et al. (2019). Briefly, the child was asked to estimate the spatial position of an Arabic digit on a physical number line when presented with a 25 cm number line with the Arabic digits 0 and 10 at respective ends of the line. A new number line on a separate card was presented for each of the target numbers 1-9, completed in a fixed random order. Children were asked to make a mark on the line to indicate where the number would go on the line. The first digit, 5, served as a sample item to be sure they understood the task. For the sample item only, feedback was provided if it was clear that the child did not understand the task. The average proportion of estimation error was calculated across all nine trials for each child. Specifically, the distance from the starting point of the number line to the child's mark (observed value) was compared to the distance from the starting point of the number line to the correct point on the number line (expected value). The proportion score was based on the discrepancy between these values according to the following formula: |(expected value – observed value) / 10|. This formula is similar to percent absolute error except that we did not multiple the proportion by 100.

Magnitude Comparison. The magnitude comparison tasks constitute the *Numeracy Screener* (e.g., Hawes et al., 2019; Nosworthy et al., 2013) and available for free download at numeracyscreener.org. The child was required to compare pairs of magnitudes ranging from 1 to 9 presented either in symbolic (56 digit pairs) or non-symbolic (56 pairs of dot arrays) formats.

The child was instructed to cross out the larger of the two magnitudes and were given two minutes to complete each condition (symbolic and non-symbolic). For practice, each child completed three sample items with the examiner and then nine practice items on their own for each format. During the instructions for the non-symbolic condition, participants were told not to count the dots. Examiners were again able to emphasize this instruction during the participants' completion of the practice items. The child was told to work as quickly and accurately as possible; self-corrections were permitted. Each task had a limit of two minutes; the researcher recorded the finish time if completed earlier. In our analysis, we used proportion correct based on total number of correct responses relative to number of items answered.

Number Naming. The child was asked to name the 10 digits from 0 to 9 out loud. A card with the digits 0 though 9 listed in random order on a white background was presented. There was no time limit and time was not recorded. The child was encouraged to try their best before saying "I don't know". Correct responses were tallied out of 10.

Arithmetic Calculations. A simple, non-standardized, paper-and-pencil arithmetic measure was attempted by each child (see also Hawes et al., 2019). The child was given 5 single-digit addition then 5 single-digit subtraction problems, and asked to complete any they could with no time constraints. Before the experimental task, one practice question in the respective domain was completed. Guidance was provided if needed on the practice trial only. The number of correct responses out of 10 was tallied.

School Grades

Children's report card marks from the Ontario primary curriculum were obtained. We received alphanumeric grades from the schools, ranging from D- (lowest) to A+ (highest). Grades were from the Language curriculum: reading, writing, oral, media literacy, and speaking.

And Math curriculum: number sense, measurement, geometry, patterns, and data. A description of the curriculum can be found in the online supplement. Although grades are ordinal in nature, we could not presume equal intervals, and hence, treated this outcome measure as ‘categorical’ in our analysis. However, the Mplus software that was used for analysis only allows a maximum of 10 categories for categorical variables, so we combined similar letter grade categories to reduce the number of categories from 12 to 4 categories corresponding to A, B, C, and D. In following analyses, the 4 categories were used unless otherwise stated.

Statistical Analysis

Missing data. Table 1 shows descriptive statistics and missing data rates for the original data set for each measure. Measures with missing data above 50% were removed from the remaining analysis including Grade 1 (G1) English Speaking, G1 Math Measurement, and Grade 2 (G2) Math Measurement. An empirical test revealed that the data were not missing completely at random (Little, 1988): $\chi^2(2532) = 3427$, $p < .001$. The approach we used to deal with missing data is data imputation via the K Nearest Neighbour (KNN) imputation method (Troyanskaya et al., 2001) using the DMwR2 package in R software with k set to 30. The KNN method imputes missing data by making estimates based on ‘k’ samples in the dataset that are similar or close in the space. Each sample’s missing values are imputed using the weighted mean value of the ‘k’-neighbors found in the dataset. The imputed dataset was used in the remaining analyses.

Exploratory factor analysis. Given the need to perform both an exploratory analysis and, within growth curve modelling, a confirmatory analysis to address the goals of the study, the dataset was randomly split in half (n = 383 and n = 384, respectively) to avoid overfitting.

To address the first goal with respect to the structure of academic skills, exploratory factor analyses were used to identify the common underlying latent factors associated with the various measures collected in kindergarten and academic grades. A significant Bartlett's test of sphericity and a Kaiser-Meyer-Olkin (KMO) value greater than 0.8 would indicate that the data were suited for factor analysis. Parallel analysis would then be used to determine the number of factors extracted. A maximum likelihood factor analysis procedure was applied and then rotated to a final solution with an oblimin rotation. Eigenvalue greater than 0.70 were retained (Jolliffe, 1972) and factor loadings less than .40 were not retained (Stevens, 2009).

Growth curve model. To address the second goal of investigating the predictive relationships between kindergarten cognitive factors and later academic grades, results from the exploratory factor analysis informed the structural equation models (SEM). We first used confirmatory factor analyses to confirm the fit of the yielded models from the exploratory factor analyses, if needed. Then growth SEM were used to assess the relation between kindergarten predictors and later academic outcomes. Notably, although different factors might emerge within grade, models would be fitted jointly to better understand shared and unique influences. Growth models were fitted using robust maximum likelihood estimation to deal with non-normality using Mplus version 8.7 (Muthén & Muthén, 2006-2017) and type is complex option to account for participants being nested within 16 schools. In these growth models, the "intercept" corresponds to the initial status in grade 1, whereas the "slope" corresponds to the difference (e.g., a gain) from grade 1 to 2. Given that our model involved a combination of continuous latent variables (i.e., kindergarten predictors) and categorical observed variables (i.e., academic outcomes), numerical integration is needed. As a result, typical fit statistics are not available with numerical integration (Muthen, 2014). Instead, the fit of the different models was compared by means of the log-likelihood ratio test (Muthen, 2014). The value of twice the log-likelihood difference between the two models (i.e., 2(log-likelihood of complex model – log-likelihood of simple model)) follows a chi-square (χ^2) distribution with the degrees of freedom (*df*) equal to the difference in numbers of estimated parameters. A non-significant finding ($p > .05$) indicates that the simple model was not significantly worse than the fit of the more complex model, and hence, would be kept as the most parsimonious and best-fitting model. Relative quality of the models was also evaluated by comparing Akaike information criterion (AIC) the Bayesian information criterion (BIC) across models, with lowest values being optimal (Burnham & Anderson, 2004).

Table 1.
Descriptive statistics and missing data analysis for the original data set

Measure	N	Mean	SD	Maximum score	Missingness		Test-retest	ICC
					Count	Percent	reliability	
Kindergarten measure								
VT	557	8.91	1.095	10	6	1.1	0.73	.007
SR	558	17.27	7.93	32	5	.9	0.95	.048
PA	448	27.22	7.62	33	115	20.4	0.74	.060
LN	558	24.094	4.55	26	5	.90	0.90	.044
LS	559	22.87	3.86	26	2	.36	0.92	.067
CR	555	66.40	22.23	Range = 33 – 257 s	8	1.4	0.89	.007
NL	562	.31	.21	1	1	.2	0.72	.12
NC	550	.92	.073	1	13	2.3	0.90	.011
SC	548	.95	.097	1	15	2.7	0.61	.011
NN	557	9.74	1.024	10	6	1.1	0.93	.005
AC	561	4.05	2.91	10	2	.4	0.80	.11
Grade 1 report card subject								
ER1	442	7.54	2.55	12	121	21.5		.098
EW1	442	7.10	2.25	12	121	21.5		.091
EO1	442	7.96	1.62	12	121	21.5		.14
EL1	393	8.20	1.40	12	170	30.2		.074
ES1	119	8.07	1.69	12	444	78.9		
MN1	442	7.98	2.011	12	121	21.5		.056
MM1	258	8.29	1.73	12	305	54.2		
MG1	430	8.32	1.68	12	133	23.6		.089
MP1	434	8.26	1.79	12	129	22.9		.047
MD1	413	8.06	1.81	12	150	26.6		.083
Grade 2 report card subject								
ER2	378	7.56	2.62	12	185	32.9		.056
EW2	378	7.03	2.29	12	185	32.9		.055
EO2	381	8.11	1.82	12	182	32.3		.036
EL2	334	8.28	1.69	12	229	40.7		.037
MN2	381	8.04	2.22	12	182	32.3		.058
MM2	171	8.23	1.97	12	392	69.6		
MG2	319	8.47	1.90	12	244	43.3		.15
MP2	363	8.26	1.95	12	200	35.5		.14
MD2	360	8.22	1.84	12	203	36.1		.047

Note.
Kindergarten measure: VT = Vocabulary test. SR = Sentence recall. PA = Phonological awareness. LK = Letter knowledge. CR = Colour Rapid Automatized Naming. NL = Number line estimation. NC = Non-symbolic magnitude comparison. SC = Symbolic magnitude comparison. NN = Number naming. AC = Arithmetic calculation.

Report card subject: ER = English reading. EW = English writing. EO = English oral. EL = English literacy. ES = English speaking. MN = Math number sense. MM = Math measurement. MG = math geometry. MP = Math patterns. MD = Math data. 1 = Grade 1. 2 = Grade 2.

Results

Kindergarten test-retest reliability.

Test-retest correlations were carried out on a separate sample of children (n = 25). Almost all measures met the recommended test-retest cut-off of 0.65 (Cohen & Swerdlik, 2009), except for symbolic comparison which was slightly under the cut-off. For most measures, test-retest reliability was excellent, ranging from 0.72 to 0.85 (Table 1). Test-retest correlation for symbolic comparison was 0.61 for our sample (but was 0.72 in Hawes et al., 2019).

Correlations

Kindergarten. Pearson correlations with false discovery rate correction for multiple comparisons for the kindergarten assessment battery are displayed in Table 2. Most variables correlated significantly with each other except number line estimation. The strongest correlations were between letter name and sound (r = .90, p < .001), letter sound and phonological awareness (r = 0.65, p < .001), and letter name and number name (r = 0.63, p < .001).

School grades. Table 3 displays the correlations with false discovery rate correction for multiple comparisons for grade 1 and 2 report card marks across the curriculum using the 12-point scale. All correlations were significant at p < .001. Nominally, all school grades slightly increased from grade 1 to 2. Correlational analyses substantiate this observation by showing high levels of stability for the same subject over time, correlations ranging from r = .53 to .67.

Table 2.

Correlations for the kindergarten measures.

Test	SR	AC	LN	LS	PA	NN	NC	SC	CR	NL
VT	0.39***	0.20***	0.22***	0.25***	0.23***	0.19***	0.11**	0.22***	-0.12**	-0.06
SR		0.40***	0.42***	0.44***	0.43***	0.26***	0.19***	0.32***	-0.34***	-0.19***
AC			0.41***	0.46***	0.41***	0.26***	0.19***	0.28***	-0.33***	-0.25***
LN				0.90***	0.61***	0.63***	0.16**	0.58***	-0.32***	-0.16**
LS					0.65***	0.51***	0.13*	0.53***	-0.31***	-0.19**
PA						0.47***	0.25***	0.47***	-0.34***	-0.09*
NN							0.29***	0.58***	-0.27***	-0.08
NC								0.48***	-0.13**	-0.05
SC									-0.38***	-0.15***
CR										0.13***

Note. * = *p* < .05. ** = *p* < .01. *** = *p* < .001.
VT = Vocabulary test. SR = Sentence recall. AC = Arithmetic calculations. LN = Letter name. LS = Letter sound. PA = Phonological awareness. NN = Number naming. NC = Non-symbolic magnitude comparison. SC = Symbolic magnitude comparison. CR = Colour Rapid Automatized Naming. NL = Number line estimation.

Table 3.

Correlations for the Grade measures. All correlations are significant at p < .001.

Grade	EO2	EL1	EL2	EW1	EW2	ER1	ER2	MN1	MN2	MG1	MG2	MP1	MP2	MD1	MD2
EO1	0.58	0.72	0.60	0.75	0.55	0.67	0.52	0.71	0.50	0.65	0.55	0.70	0.50	0.71	0.52
EO2		0.52	0.70	0.62	0.76	0.56	0.70	0.60	0.70	0.51	0.71	0.55	0.69	0.56	0.71
EL1			0.57	0.71	0.51	0.62	0.47	0.68	0.45	0.66	0.49	0.70	0.44	0.71	0.45
EL2				0.65	0.73	0.59	0.64	0.62	0.67	0.56	0.70	0.62	0.69	0.60	0.66
EW1					0.70	0.83	0.66	0.78	0.60	0.71	0.61	0.73	0.59	0.74	0.60
EW2						0.67	0.83	0.63	0.76	0.52	0.75	0.58	0.76	0.57	0.77
ER1							0.75	0.78	0.60	0.70	0.56	0.69	0.59	0.69	0.58
ER2								0.64	0.69	0.53	0.67	0.54	0.70	0.55	0.68

MN1	0.67	0.82	0.63	0.84	0.64	0.82	0.62
MN2		0.52	0.79	0.59	0.87	0.59	0.82
MG1			0.53	0.77	0.49	0.79	0.49
MG2				0.60	0.78	0.60	0.80
MP1					0.57	0.80	0.59
MP2						0.57	0.83
MD1							0.55

Note. EO = English oral. EL = English literacy. EW = English writing. ER = English reading. MN = Math number sense. MG = Math geometry. MP = Math patterns. MD = Math data. 1 = Grade 1. 2 = Grade 2.

Models of Learning Domains

Separate factor analyses for kindergarten, grade 1, and grade 2 were conducted to examine the factor structure across domains and group the variables into factors that measure like constructs. We then used these factor structure in our structural equation models.

Kindergarten predictors. Both a significant Bartlett's test of sphericity ($\chi^2(55) = 1210.44$, $p < .001$) and the KMO measure of 0.80 indicated that the implementation of the factor analysis was appropriate. Parallel analysis suggested 2 or 3 factors be extracted and we found that the three-factor model was more interpretable. The three factors explained 39%, 12%, and 10% of the variance respectively (the respective eigenvalue being 4.30, 1.33, and 1.15).

According to the three-factor model, letter sound, letter naming, number naming, and phonological awareness loaded on factor 1, deemed the LITERACY factor related to code-based skills; symbolic and non-symbolic magnitude comparison loaded on factor 2, deemed the NUMERACY factor; and finally, vocabulary, sentence recall, and arithmetic calculations, loaded on factor 3, which was deemed the EARLY LANGUAGE factor related to meaning-based skills. Number line estimation and colour RAN did not load on any factors. These three factors will be used as independent predictors in the subsequent SEM analyses.

Grade 1. The factor analysis for grade 1 was appropriate given the significant Bartlett's test, $\chi^2(28) = 1407.64$, $p < .001$, and KMO value of 0.92. Parallel analysis suggested a one-factor model of academic grades. This unidimensional model explained 64% of the variance with an eigenvalue of 5.12.

Grade 2. The grade 2 data were also suitable for factor analysis, $\chi^2(28) = 1627.73$, $p < .001$, and KMO = 0.90. Parallel analysis suggested a 1- or 2-factor solution. The one-factor model of academic grades explained 67% of the variance with an eigenvalue of 5.38. The two-factor model, revealing a clear distinction between Math and Language grades, explained 69% and 7% of the variance respectively (the respective eigenvalue being 5.30 and 0.74). In this multidimensional model, all the Math-related subjects loaded onto factor 1, this was deemed the MATH factor, whereas all the Language-related subjects loaded onto factor 2, this was deemed the LANGUAGE factor. Interestingly, media literacy had a low cross-loading of 0.35 on the math factor but was not retained given our cut-off. Confirmatory factor analysis using the remaining subset of our data indicated a better fit for the two-factor model than one-factor model of academic grades. Results for the two-factor model: CFI = .96, TLI = .95, RMSEA = .097; SRMR = .037. Results for the one-factor model: CFI = .94, TLI = .92, RMSEA = .12; SRMR = .043.

Models Predicting Academic Achievements

We were guided by the results of our factor analyses when constructing growth SEMs, using early cognitive skills to predict later academic skills. From our kindergarten measures, we had three predictors: EARLY LANGUAGE (sentence recall, vocabulary, arithmetic), LITERACY (letter knowledge, phonological awareness, number naming), and NUMERACY (symbolic and non-symbolic comparisons). Based on the results of the grade 1 and 2 marks, we evaluated two models: i) a single-factor model of academic grades based on the grade 1 factor analysis and ii) a two-factor model distinguishing between math and language based on the grade 2 factor analysis.

For each model, growth curves were fitted to predict the initial status and difference in academic grades from the three latent predictable variables in kindergarten (language, reading, numeracy)

Students nested within schools. The possibility that students assessed by the same school would be in better agreement from grade 1 to grade 2 was considered (cluster effect). However, calculations of intra-school (intra-class) correlation coefficients using multilevel modeling with the cluster procedure in Mplus indicated that the cluster effect was negligible (Table 1). ICC ranged from 0.005 to 0.109 for the kindergarten predictors (independent measures) and ranged from 0.036 to 0.136 for the grade outcomes (dependent measures). Nevertheless, we accounted for the number of clusters (16 schools) in the SEM multilevel models.

1. Unidimensional model of academic grades

The first structural analysis tested the relation between kindergarten precursors in predicting the academic grades as a single factor. Figure 1 shows the model with standardized parameter estimates displayed (unstandardized results are reported in the supplementary materials, Table S2). Kindergarten language and literacy predicted the intercept (0.56, $p < .001$ and 0.23, $p = .001$, respectively), accounting for 59% of the variance in initial grades. The positive correlation suggests that children with higher kindergarten language and literacy scores also had higher academic report card marks in grade 1. Further, it could also be observed that kindergarten literacy (.56) had a higher impact on the intercept than early language (.23). Conversely, only kindergarten literacy predicted the slope, and this relationship was negative (-.27, $p = .013$, 6% of the variance explained), suggesting that children with lower literacy scores had a larger gain from grade 1 to 2. Kindergarten numeracy was not a significant predictor of initial grades or a difference in grades. Finally, the association between intercept and gain of academic grades as a single-factor was a strong and negative relationship (-.42, $p < .001$), indicating that children with poorer grade 1 marks showed a faster gain from grade 1 to grade 2 (e.g., they are catching up to students who had higher marks at the beginning).

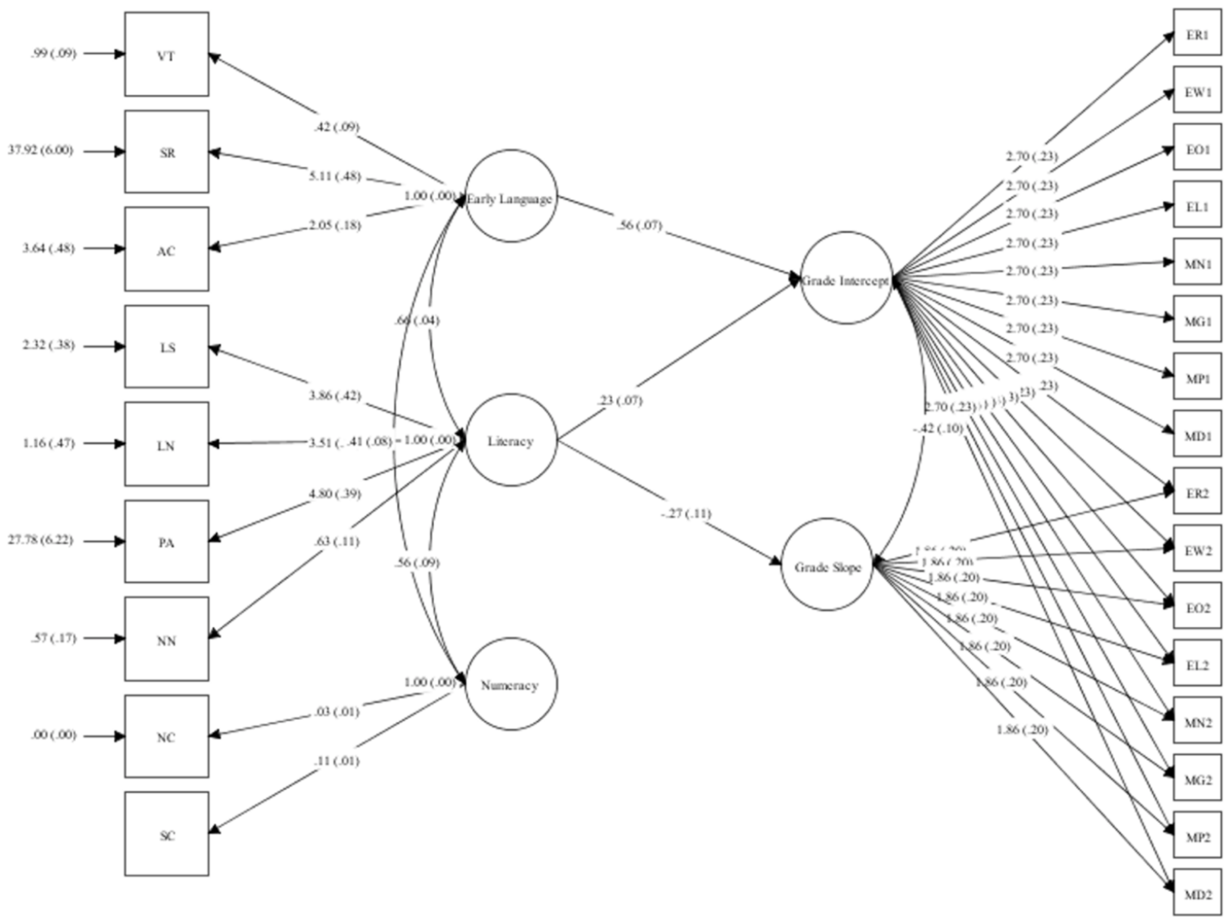

Figure 1. *Unidimensional model of academic skills predicted early language, literacy, and numeracy constructs measured in kindergarten (using imputed data).*
Note. Only significant standardized coefficient paths are displayed.
NN = Number naming. PA = Phonological awareness. LK = Letter knowledge. AC = Arithmetic calculation. SR = Sentence recall. VT = Vocabulary test. SC = Symbolic magnitude comparison. NC = Non-symbolic magnitude comparison. KRead = Kindergarten literacy factor. Klang = Kindergarten language factor. Knum = Kindergarten numeracy factor. EO = English oral. EL = English literacy. MN = Math number sense. MD = Math data. MG = Math geometry. MP = Math patterns. 1 = Grade 1. 2 = Grade 1.

2. Multidimensional model distinguishing math and language grades

The second structural analysis tested relation between kindergarten predictor variables in predicting grades related to math and language subjects. Standardized path coefficients for the two-factor predictive model of academic grades are shown in Figure 2 (unstandardized results are reported in the supplementary materials, Table S3). In this model, the language intercept was significantly predicted by early language (0.36, $p < .001$) and literacy (0.33, $p = .001$). Both predictors accounted for 48% of the variance and had about equal impact. In contrast, only early language was also associated with the gain from grade 1 to 2 (.43, $p = .012$), explaining 11% of the variance. With respect to the math intercept, kindergarten language and literacy again predicted grade 1 math marks (.45, $p < .001$ and .31, $p < .001$, respectively), explaining 57% of the variance. For the gain in math grades from grade 1 to 2, only the literacy path was significant and in a negative direction (-.29, $p = .021$, 13% of the variance). Interestingly, numeracy was not a predictor of later academic grades including math grades. Similar to the one-factor model, initial status in language was negatively related to gain (-.43, $p = .001$) as was the relationship between math intercept and gain (-.49, $p < .001$) indicating that students with poor grades in grade 1 in respective domains make gain faster compared to those with more higher grades in the beginning.

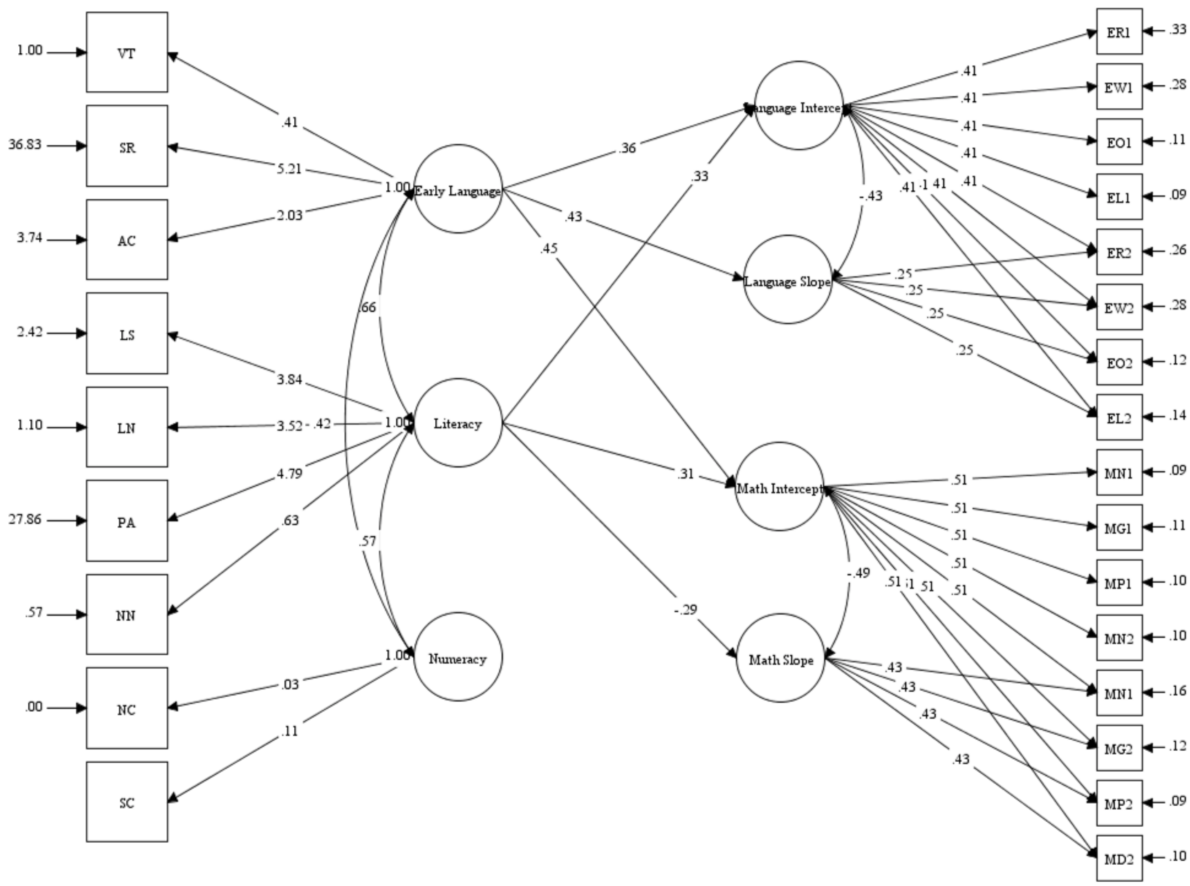

Figure 2.
Multidimensional model of academic skills predicted early language, literacy, and numeracy constructs measured in kindergarten (using imputed data).
Note. Only significant standardized coefficient paths are displayed.
NN = Number naming. PA = Phonological awareness. LK = Letter knowledge. AC = Arithmetic calculation. SR = Sentence recall. VT = Vocabulary test. SC = Symbolic magnitude comparison. NC = Non-symbolic magnitude comparison. Kread = Kindergarten literacy factor. Klang = Kindergarten language factor. Knum = Kindergarten numeracy factor. EO = English oral. EL = English literacy. MN = Math number sense. MD = Math data. MG = Math geometry. MP = Math patterns. 1 = Grade 1. 2 = Grade 1.

Model comparison. Given that numerical integration was required, typical fit indices were not available. Instead, we compared the two-factor model (complex model) to one-factor model (simple model) using a likelihood ratio test. The value of twice the log-likelihood difference was: 2(-6503.283-(-6630.502)) = 254.44 with *df* = 15 resulting in p < .001. Given that the p-value was significant, the two-factor model of predicting academic grades was deemed the best-fitting model. Evaluating the model fit based on AIC (simple model = 13461 vs. complex model = 13150) and BIC metrics (simple model = 13618 vs. complex model = 13412) also showed that two-factor model was a better fit than the simple model. Therefore, we can conclude that the two-factor model showing that early language and literacy skills were individually predictive of later language and math grades is the best fit for the data.

Discussion

In this large-scale study over three years, we examined how a range of kindergarten skills can predict early academic outcomes in language, reading, and mathematics. The first goal was to determine the dimensionality of these theoretically separate domains. As a result, no pre-existing assumptions were made about what tasks/subjects were measuring leading us to find that the structure of academic domains varied from kindergarten to grade 2, with overlaps between domains. The second goal was to assess the predictive value of early cognitive predictors. The best-fitting model was one in which early cognitive predictors were independently predictive of language or math subjects in expected and non-obvious ways (e.g., verbal domains predicted language grades but also mathematics). Finally, the relationship between beginning language or math skills (intercept) and later gains (slope) was negative, indicating that students with initially low grade 1 grades make positive gains to narrow the gap by grade 2. Overall, the novel finding of overlapping cognitive factors being predictive of skills across learning domains is consistent with recent work moving away from the categorical approach.

Kindergarten Precursors

To our knowledge, our study is the first to use a range of measures to identify key cognitive predictors of academic skills in the first few years of school. Specifically, we focused on measures that have been shown to be robust predictors in their respected domains. However, unlike prior work we did not group these predictors a priori. Instead, we took a data-driven approach to examine underlying shared and unique influences across domains using a factor analysis. Results showed that while some measures clustered together as expected, others overlapped in non-obvious ways. As expected, measures that predict meaning-based skills formed the EARLY LANGUAGE factor, measures that predict code-related skills formed the LITERACY factor, and the two magnitude comparison tasks loaded together to form the NUMERACY factor. Interestingly, the remaining numeracy tasks measured in kindergarten seemed to have diverged and loaded on different factors depending on the extent to which language demands are involved. Converging evidence supporting the divergence among mathematics tasks comes from Cross et al. (2019) who suggested that magnitude comparisons rely on the nonverbal route within the triple-code model (Dehaene & Cohen, 1995) and hence has low verbal demands, whereas arithmetic and number naming may be language-based, relying on verbal number representations. Indeed, the arithmetic task loaded with the EARLY LANGUAGE factor, highlighting a shared underlying verbal load as well as these direct measures requiring rich home-environments to teach language and numeracy skills. Number naming was related to the LITERACY factor, possibly reflecting underlying symbolic knowledge required by this factor. Indeed, number naming and letter naming was highly correlated ($r = 0.63$). Identifying verbal numerals of Arabic numbers is a similar process to recognizing the name of the letters in the alphabet as well as both tasks imposing a phonological word form load.

We found that colour RAN and number line estimation did not load with any factors. The finding that colour RAN did not load with the kindergarten reading predictors might seem to conflict with robust findings of RAN being a predictor of reading development (Araujo et al., 2015; Norton & Wolf, 2012). However, we did not evaluate the full breadth of reading skills involved such as comprehension and reading fluency, for example. Alternatively, it could be that variance that colour RAN shares with other skills is being removed and thus minimizing the independent contribution from RAN itself. For instance, RAN is considered to tap recall automaticity and controlled attention which may be captured by the high loadings of letter knowledge (automaticity with recognizing letters) on the LITERACY factor and/or of sentence recall (attend to incoming information that will need to be recalled) on the EARLY LANGUAGE factor. Performance on number line estimation also relies on other cognitive skills such as spatial reasoning and proportional reasoning that were not evaluated in the current study. Further, young children might be struggling with such complex processing requirements rather than numerical knowledge itself (Hawes et al., 2019). Future work should consider the inclusion of predictors that would reflect the full breadth of reading and mathematics as well as predictors of other cognitive skills.

Predicting Academic Skills

The factor analysis also highlighted a developmental progression of skills over time. In grade 1, academic grades were explained by a one-factor model and by grade 2, a differentiation between mathematics and language-related subjects (i.e., language and reading) emerged. The coupling of mathematics and language in grade 1 can help us understand how early mathematics skills develop. Language and early mathematical skills may be linked early on due to the nature of the mathematics curriculum drawing on language skills; students need language to understand mathematical concepts, are encouraged to talk about and through math problems, and need reading comprehensions skills for solving math problems (Ministry of Education, 2016). It would follow that early mathematics development necessitates good language and reading skills. Once this foundation is laid, more advanced and abstract arithmetic strategies can be understood, making mathematics more modularized by grade 2. Nominally, the cross-loading of media literacy on both language and mathematics in grade 2 suggest continual involvement of verbal skills to some extent. The pattern of increasing differentiation between domains as children get older is consistent with the view of development as a process of modularization, that is, skills become more specific over time (Karmiloff-Smith, 1998; Tomblin & Zhang, 2006).

Nonetheless, when we examined the role of multiple cognitive factors (language, literacy, and numeracy) as predictors of academic grades, the multidimensional model contrasting language and math outcomes fitted better than the unidimensional model. Interestingly, results showed that while some cognitive factors overlapped, others uniquely predicted later language and mathematic skills in unexpected ways. With respect to predicting language grades, not surprisingly, the verbal domains of early language and literacy were independent predictors of grade 1 language skills but only early language remained a robust predictor of gains by grade 2. It is likely that early on, students are in the ‘learning to read’ phase, and hence, language learning and development consists primarily of decoding skills (indexed by literacy) as well as starting to make meaning (indexed by early language). But, once these word recognition skills are mastered, language learning are less dependent on code-related skills. Thus, later language grades likely reflect the ‘reading to learn’ phase and become more dependent on other indicators of advanced development such as direct measures of early language.

The relation between early cognitive predictors and mathematics was interesting. We once again found that early language and literacy skills were individually predictive of mathematics in the beginning, but that literacy skills remain negatively associated with gains. As previously discussed, the link between verbal skills and early mathematics skills might be attributed to the verbal demands necessary for mathematical learning and development. Children who enter formal schooling with good language and literacy foundations could more readily use their verbal skills to scaffold mathematical learning. However, given that the literacy skills were negatively related to gains in math by grade 2, this relation would suggest a compensatory pattern. That is, students who enter kindergarten with advanced literacy skills might have developed mathematical skills to a greater degree in grade 1 than those with poor literacy skills resulting in smaller gains by grade 2. On the other hand, those with poor literacy skills initially had more room for improvement as their literacy skills improved.

Interestingly, numeracy skills did not predict any models in the study when several other predictors were included. At first glance, this finding might seem to conflict with previous work showing that early mathematical skills predict later mathematical abilities and academics in general (e.g., Geary, 2012; Clarke & Shinn, 2004). However, one possible limitation is that the numeracy factor in the current study was characterized by two very similar tasks related to magnitude comparison, and hence, did not capture the breadth of numeracy nor mathematical skills. The research on numeracy precursors and their psychometric properties too is only in its infancy compared to the long history of research on language and reading predictors. Nevertheless, the numeracy predictors that loaded on other domains did indeed predict later achievements, suggesting that early mathematics curriculum may reflect other domain general aspects of math primarily such as language and symbolic knowledge. As we have alluded to, early mathematics curriculum might depend on children having strong verbal skills to engage in 'math talk'. It would follow that children who have good language and literacy skills at school entry will have greater access to content across the curriculum. Nevertheless, a balanced mathematics composite should be considered and investigated further in future research.

Implications

The different trajectory of predictors has implications for practice and research. We found that early language and literacy skills can predict academic performance across the curriculum and over time. It would follow that universal, early screening of language skills is a good practice for identifying children at risk for general academic difficulties immediately, but especially language learning and development. Screening for children with poor reading skills can also index immediate academic difficulties as well as mathematical difficulties in later grades. Mathematical abilities that overlapped with other cognitive factors (i.e., arithmetic with language and number naming with reading) also served as predictors for academic performance. Thus, we advocate for early screening using cognitive predictors across domains. Given the overlap with both the predictors (e.g., math tasks loading with language or literacy factors) and how predictors relate to outcomes (e.g., literacy predicting later mathematical grades), it would be valuable for future work to consider how a range of cognitive predictors influence outcomes across academic domains instead of investigating specific and separable domains. Children who struggle in one domain are likely to be at risk for academic difficulties more generally.

Limitations

Limitations have been mentioned throughout the discussion and there are other considerations that should be noted. A major limitation was that we used report card marks as a measure of academic performance in the early school years. Teacher-assigned grades are not standardized and have been criticized to be subjective and unreliable. However, in a comprehensive review of the literature, Bowers (2019) reported that grades are not as subjective and unreliable as previously suggested. Instead, grades consistently and moderately correlate at about 0.50 to standardized measures. In fact, a subset of our sample completed subtests from the Woodcock-Johnson III (WJ III) Tests of Achievement in grade 1 ($n = 218$) and grade 2 ($n = 124$) and preliminary results show decisive evidence for a relationship between grades and standard scores, correlations ranging from 0.36 to 0.77, $BF_{10} > 100$ in all cases (Pham et al., in prep). In the current study, we also mitigated this limitation by treating grades as a categorical outcome without assuming equal intervals. Nevertheless, the interpretations of our results may be limited to the use of academic grades as outcomes, calling for future work to substantiate our results with academic outcomes indexed by more standardized tests.

Another crucial limitation is the lack of having detailed description of the participants with respect to diagnosis or additional services to support the implication of early screening with predictors across several domains for struggling students. However, overall, children in the sample seemed to demonstrate typically developing skills across domains. For example, all report card scores slightly increased from grade 1 to grade 2, with high levels of stability for each subject. Further, based on WJ III for a subset of the sample, children scored within normal levels with standard scores ranging from 80 to 106 ($SD = 13$ to 40). Standard scores are scaled to a mean of 100 ($SD = 15$) relative to a larger sample of children with typical development. Nevertheless, future work should consider including information about diagnosis and services in growth models to understand how results may differ for students who are revealed as at-risk versus on track for academic success.

Another limitation was that our model tested the predictive value of cognitive factors over a short time frame, up to grade 2. Grade 3 data could not be included given insufficient data for various reasons. Having more timepoints would improve statistical power and contribute to the understanding of cognitive factors overtime. Lastly, in terms of our kindergarten test battery, the full breadth of reading and math skills was not represented. We had to find a balance between forming a comprehensive screening tool and the feasibility of administering the assessment, thus, we were limited by the number of measures that could be included.

Conclusion

In this study, we examined how cognitive predictors *across* domains could predict future academic skills *across* domains. Early language (language measures and arithmetic calculations) and literacy predictors (reading measures and number naming) were individually predictive of language and math skills in grade 1. Language continued to predict later language grades in grade 2, whereas literacy was related to later math skills in grade 2. Further, numeracy precursors that overlapped with language or reading precursors played a role on later grades across domains. Finally, the achievement gap narrowed over time as children who initially scored lower made more positive gains faster compared to higher scorers. Overall, these findings highlight the importance of considering how learning domains is likely affected by early language, reading, and mathematic skills in overlapping ways, instead of focusing on specific domains.

Acknowledgements. This study was made possible by funding from the Natural Sciences and Engineering Research Council of Canada to MASKED.

Funding. The work of many research assistants is gratefully acknowledged. We would also like to thank the children and their families as well as the school districts who partnered with us.

Conflict of interest. The authors declare that they have no conflict of interest.

Data availability statement. The data necessary to reproduce the analyses presented here are not publicly accessible.

Analytic code. The analytic code necessary to reproduce the analyses presented here is not publicly accessible.

Materials. Some assessment materials for which we own the copyright are publicly accessible on https://osf.io/rj854/?view_only=2e4b6fa906464bcd9954430b513a27aa.

Preregistration. The study and analyses presented here were not preregistered.

Ethics statement. Ethics approval was obtained from the local university and school board in which the study was completed.

Informed consent. Informed consent was obtained from all participants.

References

Araújo, Reis, A., Petersson, K. M., & Faísca, L. (2015). Rapid automatized naming and reading performance: A meta-analysis. *Journal of Educational Psychology, 107(3),* 868–883. https://doi.org/10.1037/edu0000006

Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of non-word repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research, 52,* 899–914. https://doi.org/10.1044/1092-4388(2009/08-0099)

Archibald, L. M., Cardy, J. O., Ansari, D., Olino, T., & Joanisse, M. F. (2019). The consistency and cognitive predictors of children's oral language, reading, and math learning profiles. *Learning and Individual Differences, 70,* 130-141. https://doi.org/10.1016/j.lindif.2019.02.003

Archibald, L. M., Oram Cardy, J., Joanisse, M. F., & Ansari, D. (2013). Language, reading, and math learning profiles in an epidemiological sample of school age children. *PLOS One, 8(10),* e77463. https://doi.org/10.1371/journal.pone.0077463

Bishop, D. V. M., & Snowling, M. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130,* 858–886. https://doi.org/10.1037/0033-2909.130.6.858

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*(1), 189-201. https://doi.org/10.1037/0012-1649.41.6.189

Bowers, A. J. (2019) Report card grades and educational outcomes. In Guskey, T., Brookhart, S. (Eds.) *What we know about grading: What works, what doesn't, and what's next,* (p.32-56). Alexandria, VA: Association for Supervision and Curriculum Development.

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry, 49,* 422–432. https://doi.org/10.1111/j.1469-7610.2007.01849.x

Butterworth, B. (2008). Developmental dyscalculia. In J. Reed, J. Warner-Rogers (Eds.). *Child Neuropsychology: Concepts, Theory, and Practice*. Wiley-Blackwell

Catts, H.W., Fey, M.E., Zhang, X. &Tomblin, J.B. (2001). Estimating the risk of future reading difficulties in kindergarten children. *Language, Speech, and Hearing Services in Schools, 32,* 38–50. https://doi.org/10.1044/0161-1461(2001/004)

Clarke, B., & Shinn, M. R. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33(2),* 234-248. https://doi.org/10.1080/02796015.2004.12086245

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders – A follow-up in later adult life. Cognitive, language, and psychosocial outcomes. *Journal of Child Psychology and Psychiatry, 46,* 128–149. https://doi.org/10.1111/j.1469-7610.2004.00342.x

Commission on Standards for School Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics

Cross, A. M., Joanisse, M. F., & Archibald, L. M. (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

Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research, 46(3),* 544–560. https://doi.org/10.1044/1092-4388(2003/044)

De Smedt, B., & Boets, B. (2010). Phonological processing and arithmetic fact retrieval: Evidence from developmental dyslexia. *Neuropsychologia, 48,* 3973–3981. https://doi.org/10.1016/j.neuropsychologia.2010.10.018

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1(1),* 83–120.

Donlan, C., Cowan, R., Newton, E. J., & Lloyd, D. (2007). The role of language in mathematical development: Evidence from children with specific language impairments. *Cognition, 103,* 23–33. https://doi.org/10.1016/j.cognition.2006.02.007

Ehri, L. (2005). Learning to read words: Theory, findings and issues. *Scientific Studies of Reading, 9,* 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Ellefson, M. R., Treiman, R., & Kessler, B. (2009). Learning to label letters by sounds or names: A comparison of England and the United States. *Journal of Experimental Child Psychology, 102(3),* 323-341. https://doi.org/10.1016/j.jecp.2008.05.008

Ellis, R. (2015). *Understanding second language acquisition (2nd Edition)*. Oxford University Press.

Foorman, B. R., Koon, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). Examining general and specific factors in the dimensionality of oral language and reading in 4th–10th grades. *Journal of Educational Psychology, 107(3),* 884–899. https://doi.org/10.1037/edu0000026

Foulin, J. N. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing: An Interdisciplinary Journal, 18*(2), 129–155. https://doi.org/10.1007/s11145-004-5892-2

Geary, D. C., Bailey, D. H., & Hoard, M. K. (2009). Predicting mathematical achievement and mathematical learning disability with a simple screening tool: The number sets test. *Journal of Psychoeducational Assessment, 27(3),* 265-279. https://doi.org/10.1177/0734282908330592

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77(3),* 236-263. https://doi.org/10.1006/jecp.2000.2561

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology*. 104(1), 206–223. https://doi.org/10.1037/a0025398

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition*, *115*(3), 394-406. https://doi.org/10.1016/j.cognition.2010.02.002

Hawes, Z., Nosworthy, N., Archibald, L., and Ansari, D. (2019). Kindergarten children's symbolic number comparison skills predict 1st grade mathematics achievement: evidence from a two-minute paper-and-pencil test. *Learning and Instruction*. *59,* 21–33. https://doi.org/10.1016/j.learninstruc.2018.09.004

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*(1), 17-29. https://doi.org/10.1016/j.jecp.2008.04.001

Hoover, W. A., & Gough, P. B. (1990). The Simple View of Reading. *Reading and Writing*, *2(2),* 127–160. https://doi.org/10.1007/BF00401799

Howe, A. E., Arnell, K. M., Klein, R. M., Joanisse, M. F., & Tannock, R. (2006). The ABCs of computerized naming: equivalency, reliability, and predictive validity of a computerized rapid automatized naming (RAN) task. *Journal of Neuroscience Methods, 151(1),* 30-37. https://doi.org/10.1016/j.jneumeth.2005.07.014

Hulme, C., & Snowling, M. J. (2014). The interface between spoken and written language: Developmental disorders. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 369(1634),* 20120395. https://doi.org/10.1098/ rstb.2012.0395

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2,* 389–398. http://doi.org/10.1016/s1364-6613(98)01230-3

Klem, M., Melby-Lervåg,M., Hagtvet, B.,Lyster, S. A. H., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science, 18(1),* 146–154. https://doi.org/10.1111/desc.12202

Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction*, *25*, 95-103. https://doi.org/10.1016/j.learninstruc.2012.12.001

Kuppen, S. E. A., & Goswami, U. (2016). Developmental trajectories for children with dyslexia and low IQ poor readers. *Developmental Psychology, 52,* 717–734. https://doi.org/10.1037/a0040207

Ladd, G., Kochenderfer-Ladd, B., Visconti, K. J., & Ettekal, I. (2012). Classroom peer relations and children's social and scholastic development: Risk factors and resources. In *Peer Relationships and Adjustment at School* (pp. 11–49). Information Age Publishing.

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling developmental language difficulties from school entry into adulthood: Literacy, mental health, and employment outcomes. *Journal of Speech, Language, and Hearing Research, 52(6),* 1401–1416. https://doi.org/10.1044/1092-4388(2009/08-0142)

Learning Disability Association of Alberta. (2011). Reading readiness screening tool (v7.5)

Lee, J. (2011). Size matters: Early vocabulary as a preadictor of language and literacy competence. *Applied Psycholinguistics, 32(1),* 69-92. https://doi.org/10.1017/S0142716410000299

LeFevre, J. A., Skwarchuk, S. L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science*, *41*(2), 55. https://doi.org/10.1037/a0014532

Lonigan, C.J., Schatschneider C, & Westberg, L. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. In *Developing early literacy: Report of the National Early Literacy Panel* (pp. 55–106). Washington, DC: National Institute for Literacy.

McBride-Chang, C. (1999). The ABCs of the ABCs: The development of letter-name and letter-sound knowledge. *Merrill-Palmer Quarterly (1982-)*, 285-308.

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language & Communication Disorders, 48(3),* 307–319. https://doi.org/10.1111/1460-6984.12008

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*(2), 322–352. https://doi.org/10.1037/a0026744

Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H. C. (2011). Early place-value understanding as a precursor for later arithmetic performance—A longitudinal study on numerical development. *Research in Developmental Disabilities, 32(5),* 1837-1851. https://doi.org/10.1016/j.ridd.2011.03.012

Morris, R.D., Stuebing, K.K., Fletcher, J.M., Shaywitz, S.E., Lyon, G.R., Shankweiler, D.P., Katz, L., Francis, D.J,. & Shaywitz, B.A. (1998). Subtypes of reading disability: variability around a phonological core. *Journal of Educational Psychology, 90,* 347-373. https://doi.org/10.1037/0022-0663.90.3.347

Muthén, L. K., & Muthén, B. O. (1998–2017). Mplus User's Guide. Eight Edition. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K. (2010, April 25). Mplus not reporting model fit index. Message posted to http://www.statmodel.com/discussion/messages/11/19585.html?1408576087

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342–356. https://doi.org/10.1111/j.1467-9817.2004.00238.x

Nogues, C. P., & Dorneles, B. V. (2021). Systematic review on the precursors of initial mathematical performance. *International Journal of Educational Research Open, 2,* 100035. https://doi.org/10.1016/j.ijedro.2021.100035

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, *63*(1), 427-452. https://doi.org/10.1146/annurev-psych-120710-100431

Nosworthy, N., Bugden, S., Archibald, L., Evans, B., & Ansari, D. (2013). A two-minute paper-and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS ONE, 8(7),* e67918. https://doi.org/10.1371/journal.pone.0067918

Ontario Ministry of Education. (2005). *The Ontario curriculum grades 1-8: Mathematics*. Retrieved from http://www.edu.gov.on.ca/eng/curriculum/elementary/math18curr.pdf

Peters, L., & Ansari, D. (2019). Are specific learning disorders truly specific, and are they disorders?. *Trends in Neuroscience and Education, 17,* 100115. https://doi.org/10.1016/j.tine.2019.100115

Peterson, R. L., McGrath, L. M., Willcutt, E. G., Keenan, J. M., Olson, R. K., & Pennington, B. F. (2021). How specific are learning disabilities? *Journal of Learning Disabilities*, *54*(6), 466–483. https://doi.org/10.1177/0022219420982981

Poll, G. H., Betz, S. K., & Miller, C. A. (2010). Identification of clinical markers of specific language impairment in adults. *Journal of Speech, Language, and Hearing Research, 53(2),* 414–429. https://doi.org/10.1044/1092-4388(2009/08-0016)

Redmond, S. (2005). Differentiating SLI from ADHD using children's sentence recall and production of past tense morphology. *Clinical Linguistics & Phonetics, 19(2),* 109–127. https://doi.org/10.1080/02699200410001669870

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

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

Rupley, W. H., Blair, T. R., & Nichols, W. D. (2009). Effective reading instruction for struggling readers: The role of direct/explicit teaching. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 25(2-3),* 125–138. https://doi.org/10.1080/10573560802683523

Schuth, E., Köhne, J., & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning and Instruction, 49,* 157–165. https://doi.org/10.1016/j.learninstruc.2017.01.005

Senechal, M. and LeFevre, J. (2002) Parental involvement in the development of children's reading skill: A five year longitudinal study. *Child Development, 73,* 445-460. http://dx.doi.org/10.1111/1467-8624.00417

Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Eligibility criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech & Hearing Services in Schools, 37(1),* 61–72. https://doi.org/10.1044/0161-1461(2006/007)

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). Hillsdale: Erlbaum.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38(6),* 934–947. https://doi.org/10.1037/0012-1649.38.6.934

Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research*, *49(6)*, 1193–1208. https://doi.org/10.1044/1092-4388(2006/086)

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., & Altman, R. B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics (Oxford, England)*, *17*(6), 520–525. https://doi.org/10.1093/bioinformatics/17.6.520

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91(3),* 415. https://doi.org/10.1037/0022-0663.91.3.415