Cognitive Development 63 (2022) 101223

Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Sources of individual differences in early elementary school science achievement among multilingual and English monolingual children in the U.S.

Matthew E. Foster *, Sara A. Smith **, Trina D. Spencer ***

University of South Florida, United States

ARTICLE INFO

Keywords:
Multilingual children
Science
Math
Reading
Secondary data analysis
Science education

ABSTRACT

Using the public use file of the Early Childhood Longitudinal Study, Kindergarten class of 2010–2011, this study is the first to examine the extent to which science achievement is distinct from math and reading in kindergarten, first, and second grade; and the utility of kindergarten language proficiency and executive functioning (EF) for predicting science achievement among multilingual (n = 1022) and English monolingual (n = 12,343) children at each grade level. The results demonstrate that science is separable from math and reading achievement, though highly correlated. The results also indicate that kindergarten language proficiency and EF support science achievement similarly across elementary school for both groups. These results coupled with large correlations among science, math, and reading achievement suggest that integrated instruction that takes advantage of domain interconnectedness may reduce redundancies and help children connect with science content.

Job growth (i.e., the number of available jobs) in science, technology, engineering, and math (STEM) is outpacing job growth in all other fields (Pew Research Center, 2018). However, U.S. students underachieve in STEM compared to students in top-performing countries (Schleicher, 2018). Linguistically diverse minorities and, in particular, English learners are underrepresented in STEM fields in college and in the U.S. STEM workforce (National Academies of Sciences, Engineering, & Medicine [NASEM], 2018; Pew Research Center, 2018). To increase representational equity, more attention to early STEM achievement is needed, especially because differences in STEM achievement can be traced to kindergarten (Galindo, 2010; Greenfield et al., 2009). Indeed, students from historically marginalized groups attending U.S. schools are more likely to receive poor-quality child care and educational experiences (Flores, 2007), which manifests as achievement inequities by kindergarten (Galindo, 2010; Greenfield et al., 2009). In addition, the empirical literature focused on STEM achievement among students in early elementary school in the U.S. is sparse. Studies that have examined predictors of science, technology, and engineering (referred to as "science" herein) among multilingual children (i.e., students whose primary home language is not English; O'Neal & Ringler, 2010; Park et al., 2018) in the U.S. are almost nonexistent.

* Correspondence to: Rightpath Research and Innovation Center, University of South Florida, 13301 Bruce B. Downs Blvd., MHC 1721, Tampa, FL 33602, United States.

** Correspondence to: College of Education, University of South Florida, 4202 East Fowler Ave., EDU 302H, Tampa, FL 33620, United States.

*** Correspondence to: Rightpath Research and Innovation Center, University of South Florida, 13301 Bruce B. Downs Blvd., MHC 1719, Tampa, FL 33602, United States.

E-mail addresses: mefoster@usf.edu (M.E. Foster), sarasmith3@usf.edu (S.A. Smith), trinaspencer@usf.edu (T.D. Spencer).

https://doi.org/10.1016/j.cogdev.2022.101223
Received 21 April 2022; Received in revised form 6 July 2022; Accepted 11 July 2022
Available online 16 July 2022
0885-2014/

1. Scientific thinking and science achievement

The work of Inhelder and Piaget (1958) posited that scientific thinking—defined as the ability to generate, test, and evaluate hypotheses, theories, and data, and to reflect on this process (e.g., Bullock et al., 2009; Morris et al., 2012)—was unsystematic and characterized by incompleteness and logical errors until adolescence. However, the view has shifted; it is now readily accepted that scientific thinking in adolescence has its genesis in preschool and kindergarten (e.g., Ruffman, Perner, Olson, & Doherty, 1993; Zimmerman, 2007). Even kindergartners show emerging proficiency in basic experimentation, evaluation skills, and drawing causal inferences from experimental data (Sobel, Erb et al., 2017; van der Graaf et al., 2015). They can also interpret simple patterns of covariation data (Koerber et al., 2005; Ruffman et al., 1993) and differentiate between conclusive and inconclusive tests of a hypothesis (Piekny & Maehler, 2013; see also Sodian et al., 1991). These findings are consistent with the Next Generation Science Standards (NGSS; National Research Council, 2013), which recommend giving kindergarteners the opportunity to develop knowledge of science and thinking scientifically by learning to recognize patterns and formulating answers to questions about the world around them.

Despite increased recognition that young children can develop knowledge of science and think scientifically, few studies have investigated the associations between science and other domains of achievement or individual differences associated with science achievement compared to other domains of early achievement (i.e., math and reading). Therefore, we investigated the associations between science, math, and reading achievement separately in multilingual and English monolingual children in kindergarten through second grade. We also examined the association between cognitive and linguistic abilities in kindergarten and students' performance on an assessment of science achievement among multilingual and English monolinguals at the end of kindergarten, first, and second grade. For comparison purposes, the association between cognitive and linguistic abilities in kindergarten and students' performance on an assessment of math and reading were also examined.

2. Associations between science, math, and reading

To our best knowledge, the extent to which science is a separate domain of achievement relative to math and reading has not been studied longitudinally across early elementary school in a multilingual sample of students. However, emerging research has investigated the relation between science knowledge and reading (e.g., Hwang, 2020), as well as science and math (e.g., Barnard-Brak et al., 2017), longitudinally or concurrently in young students. Using data from the Early Childhood Longitudinal Study, Kindergarten class of 1998–1999 (ECLS-K: 1999; National Center for Education Statistics, 1999), the results from Barnard-Brak et al. (2017) provide evidence of moderate to large associations between reading (R^2 =.23 to.55) and math (R^2 =.36 to.50) measured across early elementary school and third grade science for the full sample of participating students.

The association between science, math, and reading may however differ for multilingual children, who are developing their English language skills as well as another language, relative to English monolingual children (Maerten-Rivera et al., 2010). In a study that looked at these relations separately for multilinguals and English monolinguals using data from the ELCS-K: 1999 (Hwang, 2020), the results provided evidence of moderate associations between kindergarten science and reading across early elementary school for multilinguals (R^2 =.24 to.25) and moderate to large associations among English monolinguals (R^2 =.24 to.38). In addition, associations between science, math, and reading may vary as a function of domain-general (e.g., language proficiency and executive functioning [EF]) and domain-specific learning mechanisms regardless of one's age, which we describe below.

Domain-general theories of developmental (e.g., constructivism), such as those of Piaget (1970), Vygotsky (1962), and Pascual--Leone and Johnson (2021) posit that cognitive development involves broad mental structures (e.g., language and EF) that facilitate mastery of tasks (Gelman & Brenneman, 2004). Such theories also describe children as taking an active role in their cognitive development by seeking out information and assimilating it with their prior learning. In contrast, domain-specific theories postulate that innate mental structures underpin and guide rapid learning in certain knowledge areas (Gelman & Brenneman, 2004). With respect to math, domain-specific theories describe early numerical development as a function of two core systems. The approximate magnitude system enables inexact estimation of relatively large quantities, magnitudes, or size (Geary, 2007); and the second system provides precise representation of small numbers (sets of three to four items) of individual objects (Feigenson et al., 2004). Therefore, as a construct, math may be distinct from science. However, reasoning about forces and interactions (as discussed in the NGSS) requires that children use quantitative and abstract thinking, which connects the constructs of science and math. Science and reading may also be linked via domain-general processes such as oral language. Early reading instruction emphasizes phonemic awareness and decoding, which are largely independent of the semantic-building oral language skills (e.g., vocabulary, grammar/syntax, inferencing, text knowledge) needed for comprehension (Cervetti, Pearson et al., 2020). As word reading becomes automatized, however, the contribution of oral language to reading comprehension increases (Adlof et al., 2006; Foorman et al., 2018). Empirical research has shown that, on average, struggling readers acquire fewer new word meanings (i.e., less abstract and technical vocabulary needed during science instruction) (Cunningham & Stanovich, 1997) and on average, struggling readers are likely less able to comprehend science texts and make science-related inferences compared to individuals with high literacy skills (Tate et al., 2012). Among middle school students, lower reading rates have been associated with difficulties recalling texts (Zabrucky & Ratner, 1992). Students need to comprehend what they read and the questions asked of them when taking tests, including science tests. Science and reading achievement are therefore associated with one another, although empirical studies have yet to inspect this relationship among young multilingual children closely.

3. Predictors of science achievement

Research focused on identifying malleable predictors of science achievement is in its infancy. An exploration of the cross-domain relations between science and predictive cognitive and linguistic abilities is important to explore how these skills and abilities develop simultaneously. Such exploration may also aid in understanding how cognitive and linguistic mechanisms support science learning. Among domain-general factors, language proficiency is the most necessary for learning and is associated with math and reading achievement (e.g., Foster et al., 2015; Foster et al., 2019; Cirino et al., 2018). Without language, children cannot engage in complex (e.g., scientific) reasoning (Amsel et al., 2008). In the U.S., teachers have the flexibility to offer hands-on, practical, and non-verbal science related experiences. As that may be, language is the primary medium of classroom instruction, and children need sufficient language abilities to understand science concepts and express their metaconceptual understanding of science knowledge (Koerber & Osterhaus, 2019). Whether the relationship between language proficiency and science achievement is the same for multilingual and English monolingual children within the U.S. educational context, however, is unknown.

In addition to language proficiency, EF is a domain-general cognitive ability consistently related to several types of problem solving, including children's math and reading achievement (e.g., Bull et al., 2008; Cirino et al., 2018; Clark et al., 2010; Willoughby et al., 2019). While EF structure and development is still a controversial field (e.g., Morra et al., 2018; Panesi & Morra, 2020), it is generally recognized that EF describes a constellation of cognitive abilities—including inhibitory control, working memory, and cognitive flexibility—needed to manage and allocate cognitive resources during cognitively challenging activities, such as switching between rules or tasks, controlling and focusing attention, ignoring distractions, and inhibiting impulses (Diamond, 2013; Pascual-Leone & Johnson, 2021). Few studies have investigated the association between EF and science achievement in the early elementary school years. As that may be, EF is fundamental to sustained attention, problem solving, planning, critical thinking, and goal-oriented behavior (Diamond, 2013; Gathercole, Pickering, Ambridge et al., 2004; Hofmann et al., 2012; Pascual-Leone & Johnson, 2021), suggesting that science achievement may improve as EF ability increases.

As children gain control over their executive reasoning, they learn how to think analytically, focus their attention on relevant elements of a problem, discard irrelevant information, generate predictions, and interpret observations (Amsel et al., 2008; Zimmerman, 2007), as well as integrate new information and revise their conclusions accordingly, all of which demand EF. In addition, the NGSS requires students in early elementary school (i.e., kindergarten through second grade) to "[analyze] data to determine if a design solution works as intended to change the speed or direction of an object with a push or a pull" (K-PS2–2; National Research Council, 2013). To achieve this standard, students can test the use of different solutions (e.g., increasing the ramp incline or length) to increase the speed of an object (e.g., marble). In this task, students are asked to observe, record data, graph and interpret data, compare results to their hypothesis, draw conclusions, and update their knowledge of forces and interactions. Greater capacity to identify a goal, create a plan, ignore distractions, focus attention on relevant observations, separate probable from improbable influences, and use their equipment properly should be associated with higher performance on these cognitively challenging science tasks. Compared to studies focused on math and reading achievement, few studies in the U.S. have explored the prediction of scientific achievement (or scientific thinking) from cognitive and linguistic abilities, though research in other parts of the world have examined these connections. For example, in a study of German fourth-graders, Mayer, Sodian, Koerber, and Schwippert (2014) explored the prediction of scientific thinking from concurrently measured inhibitory control, spatial abilities, and problem-solving skills. Their results suggested that spatial abilities and EF as indexed by a measure that captured children's planning, problem solving, and strategic decision-making abilities were statistically significant predictors of scientific thinking, when controlling for reading comprehension, whereas inhibition and IQ were not. In a study of German kindergartners, Koerber and Osterhaus (2019) demonstrated that IQ and language proficiency were statistically significant predictors of concurrently measured scientific thinking. In a study of first-year Dutch kindergartners, van der Graaf, Segers, and Verhoeven (2016) showed that visuospatial and verbal working memory, inhibition, cognitive flexibility, vocabulary, grammatical ability, and spatial visualization were related to concurrently measured scientific thinking. Finally, in a longitudinal study of German preschoolers, neither language proficiency, EF, working memory, nor IQ when measured at roughly five years of age (i.e., the beginning of the school year) predicted experimentation skills at the end of the school year (Piekny, Grube et al., 2013). While the results of the previous international studies are mixed, they provide initial evidence to suggest that language proficiency and EF may be related to scientific thinking as early as kindergarten.

Research in the U.S. exploring relationships between language proficiency, EF, and scientific thinking/achievement has largely excluded multilingual children. While scientific thinking has been investigated in several studies of young children, with the exception of Morgan, Farkas, Hillemeier, and Maczuga (2016), all of the studies we identified were conducted with children outside of the U.S. Further, Morgan et al. (2016) focused on identifying science achievement gaps from third to eighth grade in children in the U.S. Morgan et al. (2016) did not include language proficiency or EF as predictors of science achievement or report the results of predictive relations for multilingual and English monolingual children separately. Thus, research on individual differences related to science achievement in multilingual children in the U.S. seems to be nonexistent.

Studies in the U.S. have examined math achievement with multilingual (Foster et al., 2019) and English monolinguals (Foster et al., 2015) and suggest that language proficiency is more strongly related to math achievement in multilingual than in English monolingual children. When accounting for the effects of prior numeracy and several cognitive and linguistic abilities measured in English *and* Spanish among Spanish-English multilingual children in Foster et al. (2019), Spanish vocabulary proficiency continued to significantly predict Spanish numeracy and Spanish applied problem solving in kindergarten. In contrast, after accounting for the effects of English numeracy, English vocabulary proficiency was no longer predictive of English applied problem solving in a parallel study with English monolinguals (Foster et al., 2015). Given the complexities of making inferences from between group comparisons across studies, an important next step is to investigate such a relationship with science directly within a single study in the U.S.

In the U.S., multilingual children are one of the fastest growing subgroups of K–12 students (Park et al., 2018), making up approximately 22 % of children between 5 and 17 years of age (U.S. Department of Health & Human Services & U.S. Department of Education, 2017). Excluding such a large group from the extant body of research has the potential to obscure patterns, produce incomplete knowledge, and exacerbate rather than ameliorate inequities. If, for example, the relationships between language proficiency, EF, and scientific achievement are stronger for multilingual than English monolingual children, these relationships will have important implications for language use during instruction. For example, multilingual instruction may accelerate academic achievement trajectories among the former group. Indeed, some research suggests that after students reach a threshold of linguistic competence, the cognitive demands of bilingualism (i.e., switching, inhibiting, and attending) may support multilingual children's achievement (e.g., Ardasheva et al., 2012; Choi et al., 2018; Espinosa, 2015; Hartanto et al., 2018; Kempert et al., 2011; Swanson et al., 2020). Alternatively, some multilingual children may struggle when science instruction is provided only in English, interfering with their comprehension of science content and hindering their comprehension of items on English standardized tests, which also explains some of the differences between multilingual and English monolingual children on math tests (e.g., Abedi, 2002; Abedi & Lord, 2001; Horn, 2003).

4. Limitations of prior research

Prior research suggests that students practice language, math, and reading skills when they engage in science activities (Gelman & Brenneman, 2004; Inan et al., 2010). However, a foundation of basic research has not been established to support intentionally integrating these domains. The interconnected relationships between science and other domains of learning have not been empirically demonstrated during the early elementary school years among multilingual and English monolingual children. The stronger the association between science, math, and reading, the stronger the argument that integrated science curricula should be used over the traditional approach that uses siloed, non-integrated curricula.

In addition to the need for the present study, extant research investigating the associations between science, math, and reading achievement as well as the prediction of science achievement are limited. One, they have not focused on multilingual children in the U.S. (e.g., Barnard-Brak et al., 2017; Koerber & Osterhaus, 2019; Mayer et al., 2014; Piekny & Maehler, 2013; van der Graaf et al., 2016). In the one study that included a focus on multilingual children in the U.S. (Hwang, 2020), the measure used as an indicator of science knowledge was a measure of general knowledge, including items that assessed science as well as social studies knowledge. Two, studies focused on the prediction of science achievement tended to focus on concurrent, not longitudinal relations (Koerber & Osterhaus, 2019; Mayer et al., 2014; van der Graaf et al., 2016). Finally, studies on these topics tend to be siloed into content areas and do not cut across academic domains (e.g., Koerber & Osterhaus, 2019; Mayer et al., 2014; Piekny & Maehler, 2013; van der Graaf et al., 2016). Because the previous characteristics limit the inferences that can be drawn from the extant research related to science achievement, additional research is needed.

5. Study's purpose and research questions

The present study is the first to investigate the extent to which cognitive and linguistic abilities predict science achievement in kindergarten through second grade among multilingual and English monolingual children in the U.S. The relations between science, math, and reading achievement, as well as their association with cognitive and linguistic abilities, may differ between multilingual children, who are developing in English as well as another language, and English monolingual children (Maerten-Rivera et al., 2010). The extent to which these relations deviate for multilingual and English monolinguals could inform the development of STEM curriculum, differentiated for student populations if necessary. STEM instruction derived from the empirical basis of domain interconnectedness could help to ameliorate science achievement inequities (e.g., National Association of Educational Progress [NAEP], 2015). Research from outside the U.S. suggests that language proficiency and EF predict scientific thinking in elementary school–age children (Koerber & Osterhaus, 2019; Mayer et al., 2014; Piekny, Grube et al., 2013; van der Graaf et al., 2016). The development of scientific thinking is a cumulative process of intentional knowledge-seeking (e.g., Bullock et al., 2009; Kuhn, 2011; Zimmerman, 2007), and individual differences in science achievement across elementary school can be traced to kindergarten (Morgan et al., 2016). Kindergarten is therefore the ideal time to identify malleable predictors of children's science achievement across early elementary school. Given that children from low-socioeconomic status (SES) households typically have less access to resources associated with learning traditionally valued by the U.S. education system, i.e., English language books and educational toys (Arnold & Doctoroff, 2003), SES is included as a predictor in our regression analyses. Our analyses were guided by the following research questions and hypotheses. We did not have a priori hypotheses regarding the strength of relations changing from one grade level to the next.

1. To what extent does science represent a unique domain of achievement relative to math and reading in kindergarten, first grade, and second grade as measured in multilingual and English monolingual children?

Science tasks often require children to handle numerical data (Kanari & Millar, 2004) and science instruction requires children to read informational text. Therefore, we hypothesized that science, math, and reading achievement would be highly correlated in both populations. However, because of domain-specific (e.g., core knowledge) theories, we expected domain knowledge within science to be unique from math and reading achievement in both populations.

Table 1
Sample demographic characteristics.

	Multilingual Children ($n = 1022$)	Monolingual English Children ($n = 12,343$)
Mean Age in Months (SD)	65.60 (4.62)	66.35 (4.57)
Gender		
Male	500 (49 %)	6364 (52 %)
Female	522 (51 %)	5979 (48 %)
Race/Ethnicity		
White	44 (4 %)	7308 (59 %)
Black	33 (3 %)	1817 (15 %)
Hispanic	631 (61 %)	1847 (15 %)
Asian	293 (29 %)	469 (4 %)
Hawaiian/Pacific Islander	7 (< 1 %)	69 (< 1 %)
American Indian	3 (< 1 %)	110 (< 1 %)
Multiracial	11 (1 %)	719 (6 %)

2. To what extent does kindergarten English language proficiency and EF predict science achievement in kindergarten, first grade, and second grade when accounting for SES? Do these relationships differ for multilingual children and English monolinguals?

Given the results of similar studies of math achievement with multilingual children (Foster et al., 2019) and English monolinguals (Foster et al., 2015), we expected the relationship between language proficiency, EF, and science (as well as math and reading) achievement to be stronger among multilingual children.

6. Method section

The data analyzed in this study came from the public use file of the ECLS-K class of 2010–2011 (ECLS-K: 2011; National Center for Education Statistics, 2011). The ECLS-K: 2011 is a nationally representative cohort of children who entered kindergarten during the 2010–2011 school year.[1] The cohort was followed from kindergarten through fifth grade, but we only used data from the spring of kindergarten, first grade, and second grade to answer the current research questions. The spring measurement occasions were chosen in an effort to be consistent and because the science test was not administered in the fall of kindergarten. Measures for language proficiency and EF are only from the spring of kindergarten (2011), whereas science, math, and reading achievement were measured in each grade (2011–2013).

6.1. Participants

Only participants who completed the ECLS-K direct cognitive assessment in English (i.e., passed the English screener) were included in this study ($n = 13,365$). multilingual children ($n = 1022$) were identified by the ECLS-K researchers as living in a home in which a language other than English[2] was the primary language spoken and who also had parents who reported speaking a non-English language as the primary language in the home during the fall[3] or spring[4] of the student's kindergarten school year (see Table 1 for demographic characteristics). Of the multilingual children, 61% reported Spanish as the primary language used in the home. Twenty-two percent reported primarily using a language originating in Asia in the home. The remaining 17% of families reported primarily speaking a language not captured by the previous categories. Parents of English monolingual children ($n = 12,343$) reported only speaking English in the home and not using another language.

6.2. Measures

In the following subsections, we briefly describe each measure and its reliability as reported in ECLS-K user's manual and psychometric reports. For science, math, and reading, ECLS-K: 2011, examiners asked students questions related to images that were presented on a small easel. The students responded by pointing or telling the examiner their answers. For additional information on the measures and the psychometrics for each variable, see Najarian, Tourangeau, Nord, and Wallner-Allen (2018), Tourangeau et al., (2015, 2017). All measures were administered in English.

[1] The cohort of students was selected using a three-stage process. Stage one included 90 primary sampling units (PSUs), or geographic areas that are counties or groups of contiguous counties. In stage two, samples of public and private schools with kindergarten programs or that educated children of kindergarten age in ungraded settings were selected within the PSUs. In the final stage, children enrolled in kindergarten and kindergarten-age children in ungraded schools or classrooms were selected within each sampled school (see Tourangeau, 2015).

[2] The variable (i.e., X12LANGST) was a composite created to indicate whether English or a non-English language was a primary language spoken in the home.

[3] Primary language of the household fall of kindergarten (i.e., P1PRMLNG).

[4] Primary language of the household spring of kindergarten (i.e., P2PRIMLN).

6.2.1. *Language proficiency*

Language was measured using the "Simon Says" (receptive) and "Art Show" (expressive) tasks from the Preschool Language Assessment Scale (*pre*LAS; Duncan & De Avila, 2000). The Simon Says task[5] tests receptive language, listening comprehension, following directions, and total physical response by having children follow simple directions provided by the examiner in English (e.g., "Point to the floor"). The Art Show task tests expressive language through a picture vocabulary task designed to elicit labels for concrete nouns and single-word responses. Reliability of the *pre*LAS total raw number-right scores was.89 for the spring of kindergarten.

6.2.2. *Executive functioning*

The Dimensional Change Card Sort (DCCS; Zelazo, 2006) task was administered to assess cognitive ability, namely EF (i.e., cognitive flexibility and inhibitory control). In this task, children were instructed to sort a series of 22 bivalent picture cards according to different rules. First, children play the "Color Game" and are instructed to sort the first set of cards based on their color (i.e., red or blue). Second, children play the "Shape Game" and are instructed to sort the cards based on their shapes (i.e., rabbit or boat). Finally, children play the "Border Game" and are instructed to sort the cards based on their color or shapte depending on whether or not the card has a black border around the edges. If the card has a border, the child is to sort by color; if there is no border on the card, the child is to sort by shape. To the best of our knowledge, the reliability of the DCCS total raw number-right scores was not reported for the spring of kindergarten.

6.2.3. *SES*

To measure SES, the recommended composite variable as detailed in the Tourangeau et al. (2015) was used, which was based on parent reports from the fall 2010 and spring 2011 parent interviews. Five components were used to create the recommended SES composite scores: (1) parent 1/guardian's education; (2) parent 2/guardian's education; (3) parent 1/guardian's occupational prestige score; (4) parent 2/guardian's occupational prestige score; and (5) household income. Reliability statistics were not reported for the composite SES measure, as the reliability of SES scores, though continuous in this study, are not routinely reported.

6.2.4. *Science achievement*

The science assessment was similar to the science items measured as part of the "General Knowledge Test" from the ECLS-K Class of 1998–1999. Items were chosen based on areas identified as important in the 1996–2005 NAEP science frameworks. The assessment included equal proportions of questions on scientific inquiry and conceptual understanding of physical science, life science, and environmental science. In short, students were expected to conduct simple investigations and draw inferences and conclusions. They were also expected to understand scientific topics such as the basic properties of solids, liquids, and gasses; the weather's effect on people's daily activities; and the relationship between the Sun and the Earth. IRT scaled scores were derived from a two-stage administration of the science assessment. Items in the first stage covered a broad range of difficulty, and each child's scores on these items determined which second-stage form (low, middle, or high difficulty) the child received. Reliability of the IRT scale scores obtained in spring of kindergarten, first, and second grade were.75, .83, and .83, respectively.

6.2.5. *Math achievement*

The math assessment was based in part on the math assessment from the ECLS-K Class of 1998–1999, but it was grounded in the NAEP math framework (see National Assessment Governing Board [NAGB], 1994) and extended down into earlier grades for the ECLS-K: 2011. The current math assessment was also updated to be consistent with the Mathematics Framework for the 2005 NAEP (NAGB, 2004); National Council of Teachers of Mathematics Principles and Standards for School Mathematics (2001); National Mathematics Advisory Panel (2008); and the math standards espoused in several states. The assessment measured skills in conceptual knowledge, procedural knowledge, and problem solving. Questions included in the assessment focused on number sense and operations; measurement; geometry and spatial sense; data analysis, statistics, and probability; and patterns, algebra, and functions. The same two-stage procedure that was used for the science assessment was also used for the math assessment. Reliability of the IRT scale scores obtained in spring of kindergarten, first, and second grade were.94, .93, and .94, respectively.

6.2.6. *Reading achievement*

The ECLS-K reading assessment was largely based on the 2009 Reading Frameworks for NAEP (NAGB, 2008). In addition, experts in reading assessment development were consulted to ensure that the assessment reflected current curriculum standards and would be suitable for children in each grade level. The knowledge and skills assessed in the reading assessment included basic skills, vocabulary, and comprehension. Basic skills included early literacy skills such as phonological awareness, letter recognition, rhyming, and word recognition. To measure vocabulary, children were asked to convey their knowledge both verbally and nonverbally. Finally, reading comprehension skills were assessed by having children read literary and informational texts, and then asking them questions about what they read. The assessment for kindergarten and first grade students began with more emphasis on basic reading skills, while greater emphasis was placed on reading comprehension in the third, fifth, and eighth grades—grade levels outside the scope of the

[5] The *pre*LAS "Simon Says" task is not the same as the "Simon Says" task often used in the literature on executive function in young children (e.g., Carlson, 2005). For the task in the *pre*LAS, the child is *not* required to perform the experimenter's direction only if it is preceded by the phrase "Simon says".

Table 2
Sample's descriptive statistics by group.

	Grade	Multilingual Children (*n* = 1022)				Monolingual English Children (*n* = 12,343)			
		Mean	*SD*	*Skew*	*Kurtosis*	*Mean*	*SD*	*Skew*	*Kurtosis*
1. SES	K	-0.54	0.81	0.63	0.02	0.06	0.77	0.24	-0.40
2. Language Proficiency	K	17.85	2.48	-2.00	-8.03	19.58	1.28	-7.90	7.30
3. Executive Functioning	K	14.60	3.06	-1.78	3.08	15.34	2.67	-2.09	5.18
4. Science	K	-1.16	0.72	0.03	-0.67	-0.44	0.67	-0.70	0.38
5. Math	K	-0.50	0.61	-0.37	0.95	-0.32	0.61	-0.59	1.56
6. Reading	K	-0.41	0.69	-0.57	0.88	-0.21	0.61	-0.83	2.50
7. Science	1st	-0.35	0.84	-0.46	0.37	0.26	0.72	-0.83	1.39
8. Math	1st	0.39	0.54	-0.66	2.92	0.56	0.56	-1.15	4.35
9. Reading	1st	0.43	0.51	-1.61	5.52	0.60	0.47	-1.94	8.57
10. Science	2nd	0.47	0.81	-0.62	0.95	0.86	0.68	-0.91	2.42
11. Math	2nd	0.96	0.54	-0.27	-0.16	1.09	0.53	-1.03	2.70
12. Reading	2nd	0.85	0.39	-1.63	6.20	0.98	0.36	-1.53	5.84

Note. K is kindergarten; 1st is first grade; 2nd is second grade. SES, Language Proficiency, and Executive Functioning were raw scores. The remaining measures are IRT-based theta scores.

present study. The same two-stage procedure that was used for the science and math assessments was also used for the reading assessment. Reliability of the IRT scale scores obtained in spring of kindergarten, first, and second grade were .95 ,.93, and .91, respectively.

6.3. Data analytic overview

To address the research questions, we first put the predictors (i.e., SES, *pre*LAS total score, DCCS total score) on the same metric by transforming their raw scores to z-scores. As recommended in Tourangeau et al. (2015), IRT-based theta scores were used for science, math, and reading achievement, which are appropriate for both cross-sectional and longitudinal analyses, and are appropriate for examining differences in overall achievement among subgroups of children. After conducting preliminary analyses, we examined the correlations among the variables, conducted three multigroup regression analyses, and conducted four measurement invariance tests to determine if the predicted relationships differed statistically between multilingual and English monolingual children. We also cross-validated our results with a subgroup of multilingual children who received services for limited English proficiency (LEP) and a subgroup of English monolingual children who did not receive LEP services. The effect of SES was accounted for in all regression analyses, which were estimated using maximum likelihood estimation with robust standard errors (MLR) in Mplus (Version 8.2). Using the "Type = Complex" command in conjunction with the "Cluster = Class" command allowed us to correct standard errors of all parameter estimates for non-independence of students due to nesting within classrooms.

The tenability of each analytic model was judged relative to the collective information provided by commonly used fit indices, including the Bayesian information criterion (BIC), sample size adjusted BIC (SSABIC), root mean square error of approximation (RMSEA), comparative fit index (CFI), and standardized root mean square residual (SRMR). Lower values of the BIC and SSABIC indicate better model fit than higher values. The RMSEA tests the extent to which the model fits reasonably well in the population (Brown, 2006). RMSEA values ≤ 0.08 are preferred over larger values (Marsh et al., 2004). The CFI ranges from 0.0 to 1.0 (Brown, 2006) and values closer to 1.0 imply good fit. The SRMR is viewed as the average discrepancy between the observed and predicted model correlations. The smaller the SRMR value, the better, with 0.0 indicating perfect model fit (for review, see Brown, 2006). R^2 was interpreted according to the recommendation by Cohen et al. (2003) for small (0.02), medium (0.13), and large (0.26) when only a few independent variables are used.

7. Results

7.1. Preliminary analyses

The sample's descriptive statistics are displayed by group in Table 2. The only evidence of data non-normality was the skew (−7.90) for the *pre*LAS total scores for the English monolingual children, which exceeded Weston and Gore (2006) recommended interval of $|3.00|$ for skewness but not for kurtosis $|10.00|$. Skewness and kurtosis for all the variables for the multilingual children ($|2.00|$ and $|8.03|$) and the remaining variables for the English monolingual children ($|2.09|$ and $|7.30|$) were below Weston and Gore (2006) recommendations. Given that MLR estimation in Mplus is well known for its ability to reliably estimate models with non-normal data, we did not transform the *pre*LAS z-scores for either group of children.

Table 3
Correlations among all variables for multilingual and english monolingual children.

Sample	Variables	Time Point	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
Multilingual	1. SES	K	---											
	2. Language	K	.21	---										
	3. EF	K	.12	.25	---									
	4. Science	K	.33	.56	.31	---								
	5. Math	K	.28	.38	.37	.57	---							
	6. Reading	K	.31	.48	.32	.56	.75	---						
	7. Science	1^{st}	.32	.51	.37	**.72**	.64	.62	---					
	8. Math	1^{st}	.31	.31	.39	.52	**.82**	.65	.66	---				
	9. Reading	1^{st}	.31	.41	.35	.51	.71	**.76**	.66	.75	---			
	10. Science	2^{nd}	.30	.45	.37	**.67**	.66	.61	**.81**	.69	.65	---		
	11. Math	2^{nd}	.34	.31	.35	.51	**.79**	.62	.65	**.85**	.71	.72	---	
	12. Reading	2^{nd}	.32	.35	.32	.51	.69	**.69**	.64	.69	**.85**	.67	.72	---
	Means(SD)		-.62(1.00)	-.78(1.41)	-.20(1.10)	-1.16(0.72)	-.50(0.61)	-.41(0.69)	-.35(0.84)	.39(0.54)	.43(0.51)	.47(0.81)	.96(0.54)	0.85(0.39)
		Time Point	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
English Monolingual	1. SES	K	---											
	2. Language	K	.11	---										
	3. EF	K	.17	.27	---									
	4. Science	K	.35	.42	.35	---								
	5. Math	K	.37	.37	.39	.63	---							
	6. Reading	K	.33	.38	.32	.55	.77	---						
	7. Science	1^{st}	.37	.40	.37	**.76**	.66	.59	---					
	8. Math	1^{st}	.36	.37	.37	.61	**.84**	.70	.69	---				
	9. Reading	1^{st}	.34	.36	.32	.53	.73	**.81**	.62	.77	---			
	10. Science	2^{nd}	.37	.49	.37	**.73**	.67	.61	**.83**	.70	.64	---		
	11. Math	2^{nd}	.38	.35	.36	.61	**.80**	.67	.68	**.87**	.72	.73	---	
	12. Reading	2^{nd}	.37	.33	.32	.56	.70	**.75**	.63	.73	**.87**	.69	.74	---
	Means(SD)		.12(0.95)	.21(0.73)	.07(0.96)	-.44(0.67)	.26(0.72)	.86(0.68)	-.32(0.61)	.56(0.56)	1.09(0.53)	-.21(0.61)	.60(0.48)	.98(0.36)

Note. EF is executive functioning; K is kindergarten; 1st is first grade; 2nd is second grade. Bold is autoregressive correlations. Boxes are concurrent correlations among achievement measures. SES, Lang, and EF were transformed to z-scores. The remaining measures are IRT-based theta scores.

Table 4
Multigroup Multivariate Regression Analyses for Multilingual and English Monolingual Children.

Variable	Science				Math				Reading			
	β	*SE*	*p*	R^2	*β*	*SE*	*p*	R^2	*β*	*SE*	*p*	R^2
Model 1 – Kindergarten												
Multilingual				.39				.26				.31
SES	.21	.03	<.001		.20	.03	<.001		.21	.03	<.001	
Language Proficiency	.48	.03	<.001		.27	.03	<.001		.39	.03	<.001	
Executive Functioning	.16	.03	<.001		.28	.03	<.001		.19	.03	<.001	
English Monolingual				.34				.33				.28
SES	.27	.01	<.001		.30	.01	<.001		.26	.01	<.001	
Language Proficiency	.37	.02	<.001		.30	.02	<.001		.32	.02	<.001	
Executive Functioning	.21	.01	<.001		.26	.01	<.001		.20	.01	<.001	
Model 2 – 1st Grade												
Multilingual				.36				.25				.27
SES	.20	.03	<.001		.24	.03	<.001		.22	.03	<.001	
Language Proficiency	.40	.03	<.001		.16	.04	<.001		.28	.04	<.001	
Executive Functioning	.25	.03	<.001		.32	.04	<.001		.26	.04	<.001	
English Monolingual				.32				.30				.25
SES	.29	.01	<.001		.29	.01	<.001		.28	.01	<.001	
Language Proficiency	.30	.02	<.001		.26	.02	<.001		.27	.03	<.001	
Executive Functioning	.24	.01	<.001		.25	.01	<.001		.20	.01	<.001	
Model 3 – 2rd Grade												
Multilingual				.31				.25				.24
SES	.20	.03	<.001		.26	.03	<.001		.23	.03	<.001	
Language Proficiency	.34	.04	<.001		.18	.05	<.001		.24	.05	<.001	
Executive Functioning	.27	.03	<.001		.28	.04	<.001		.24	.04	<.001	
English Monolingual				.32				.30				.27
SES	.29	.01	<.001		.31	.01	<.001		.31	.01	<.001	
Language Proficiency	.30	.02	<.001		.25	.02	<.001		.25	.02	<.001	
Executive Functioning	.25	.01	<.001		.25	.01	<.001		.21	.01	<.001	

Note. Completely standardized results reported. SES, language proficiency, and executive functioning were measured during the spring of participant's kindergarten school year.

Table 5
Tests of measurement invariance for multilingual and english monolingual children.

Model	BIC	SSABIC	RMSEA	CFI	SRMR
1. Unstructured covariance matrices	176,450	176,087	.023	.997	.025
2. Predicting kindergarten achievement	155,756	155,613	.041	.994	.037
3. Predicting first-grade achievement	141,715	141,572	.035	.994	.030
4. Predicting second-grade achievement	129,509	129,366	.026	.997	.030

7.2. Correlations

As expected, Table 3 shows high autoregressive and within-domain correlations. The correlations among science, math, and reading for multilingual and English monolingual children are large. These correlations correspond to R^2 of .30 to.56 for multilingual children and R^2 of .30 to.59 for English monolingual children. With respect to the first research question, for multilingual children, kindergarten science and math were correlated at .57, which translates to R^2 of .32. Thus, 68% of the variance in kindergarten science was distinct from kindergarten math.[6] Similarly, 69% of the variance in kindergarten science was distinct from kindergarten reading for the group of multilingual children. However, in first grade, the uniqueness of science from math and reading was smaller, equating to 56% for each correlation. In second grade, the uniqueness of science from the other two domains was even less, equating to 49 % and 55 % unique variance from math and reading, respectively. For English monolingual children, the same pattern emerged, with science becoming less distinct from math and reading in each school year from kindergarten to second grade. Thus, multingual and English monolingual children show the same developmental pattern when it comes to performance on English tests of science, math, and reading achievement. One final noteworthy finding was the correlation between math and reading. In both groups of children, this correlation was rather high, suggesting that these constructs, as measured in the ECLS-K, share a lot of variance in multilinguals (52–56%) and English monolinguals (55–59%).

[6] Please note that all reports of unique (or unexplained) variance includes error variance.

7.3. Regression analyses

Three multigroup multivariate regression analyses were conducted to address research question 2 and model the prediction of the three outcomes at each of the three grade levels from SES, kindergarten language proficiency, and EF. Completely standardized results are reported in Table 4. When accounting for SES, language proficiency and EF are statistically significant predictors of science, math, and reading achievement in multilingual and English monolingual children. The amount of variance that the group of predictors accounted for in science was similar for multilingual (31–39%, depending on the grade level) and English monolingual (32–34%, depending on the grade level) children. In general, the group of predictors accounted for a small to large proportion of the variance in each outcome at each grade level for both groups of children.

To test if the predictions of science, math, and reading achievement were equivalent across multilingual and English monolingual children, four tests of measurement invariance were conducted. The first test constrained the unstructured covariance matrices to equality across the two groups of children across all grade levels. Then, the latter three tests constrained the regression pathways from kindergarten SES, language proficiency, and EF to science, math, and reading achievement in kindergarten, first grade, and second grade to equality across the two groups of children. The model fit indices from these tests are reported in Table 5. The initial model that constrained the covariance matrices to equality fit the data well (see Model 1). The subsequent models that held the pathways from SES, language proficiency, and EF to kindergarten (see Model 2), first grade (see Model 3), and second grade (see Model 4) science, math, and reading achievement to equality also fit the data well. Thus, in contrast to our hypothesis, there were not significant differences between multilingual and English monolingual children with regard to the relations between predictors and each domain of achievement.

7.4. Cross validation regression analyses

Given potential concerns that the multilingual children may have had moderate English proficiency, which influenced the null findings above, we cross-validated our results with a subgroup of multilingual children who received services for limited English proficiency (LEP) and were therefore English language learners (ELLs) ($n = 524$) and a subgroup of English monolingual children who did not receive LEP services ($n = 294$). Most of the ELLs were Hispanic (66%), followed by Asian (27%), White (3%), and Black (3%). The English monolingual non-ELLs included Hispanic (36%), Asian (26%), White (22%), Black (10%), Multiracial (4%), and Hawaiian/Pacific Islanders. The results (i.e., correlations, regression analyses, and tests of measurement invariance) are displayed in the Appendix. The full report of the results is available by request from the first author. In brief, the same pattern of results emerged in the subgroup analyses as found for the multilingual and English monolingual children, thereby strengthening the inferences that can be drawn from this study.

8. Discussion

This study extends the literature, as it is the first study to examine science achievement, its association with math and reading, and its prediction from language proficiency and EF when accounting for SES in multilingual and English monolingual children in the U.S. Consistent with prior research (Barnard-Brak et al., 2017; Hwang, 2020), when using a broad test of science, our results show that science is a distinct domain of achievement, though it is shares large associations with math and reading in multilingual and English monolingual children. The results were also consistent with those of studies that examined the prediction of scientific thinking (e.g., Mayer et al., 2014; Koerber & Osterhaus, 2019; van der Graaf et al., 2016), showing that language proficiency and EF in kindergarten predicted science as well as math and reading achievement across early elementary school. Finally, in contrast to our hypothesis, the results indicated that the prediction from kindergarten language proficiency and EF to science achievement did not differ for multilingual and English monolingual children despite the additional challenge of taking science, math, and reading assessments in English for multilingual (see Maerten-Rivera et al., 2010).

8.1. Associations between science, math, and reading achievement

In contrast to prior studies (Barnard-Brak et al., 2017; Hwang, 2020), the present study was able to unpack the correlations between science, math, and reading achievement during kindergarten, first and second grade, into variance that is shared among the variables and variance that is distinct, we were able to show that the three domains of achievement were highly associated with one another in multilingual and English monolingual children. The large correlations suggest that cognitive development is supported by domain-general mechanisms, such as language proficiency and EF, as posited by Piaget (1970), Vygotsky (1962), and Pascual-Leone and Johnson (2021), and to some extent facilitate learning and concept acquisition in science, math, and reading. The association between science and math may be in part due to a third variable—relational reasoning, which is the ability to jointly consider multiple sets of relations between mental representations of objects, attributes, and events (e.g., Hummel & Holyoak, 1997; Miller Singley & Bunge, 2014). Relational reasoning is important in solving science and math problems (Blums et al., 2017). For example, highlighting the relational features of analogies, which is central to solving science problems such as hypothesis testing (Chen & Klahr, 1999), has been shown to improve problem-solving performance in five- and six-year-olds (Chen et al., 2016). Relational reasoning is also essential in solving problems with algebraic functions. For example, McNeil and Alibali (2005) results suggested that older and arguably more advanced students viewed the equal sign (i.e., =) as a relation symbol of equivalence, whereas younger, less advanced students tended to view the equal sign as operational or signaling the need to carry out a calculation. Relational reasoning may therefore support math (Miller Singley & Bunge, 2014) and science achievement.

The high correlations between science and reading achievement are a reminder that reading skills are necessary for comprehending science problems (e.g., Miller Singley & Bunge, 2014; Voss & Silfies, 1996). Basic literacy skills (e.g., phonological awareness) are necessary to fluently read science problems. Increased oral language and vocabulary skills help students understand science texts, make science-related inferences (Tate et al., 2012), and likely help students comprehend oral discourse during science instruction. In fact, skilled reading may serve in a compensatory role for children with less knowledge of science (O'Reilly & McNamara, 2007). Reciprocally, students with strong knowledge of science may be more capable of making inferences when reading text that is difficult to understand, omits important background information, and fails to clearly connect related concepts (e.g., Beck et al., 1989; VanLehn, 1996, 1998). Thus, success in science and reading are likely reciprocal. Optimal learning may occur when children have a solid background in science and strong reading skills that enable them to easily integrate information in text to their existing knowledge and experiences (Cervetti & Wright, 2020; McNamara & Kintsch, 1996; Snow, 2002).

The correlations between science, math, and reading achievement also provided evidence of domain specificity—that the three domains were somewhat distinct from one another (i.e., $r \leq 0.85$; for review, see Brown, 2006). This is consistent with other research that shows that one's level of knowledge in a domain can dramatically impact the comprehension and learning of new information within that domain (e.g., Shapiro, 2004; Thompson & Zamboanga, 2004). The high within-domain correlations across time also suggest that students' standing relative to their peers is rather stable during early elementary school. High-performing students in kindergarten continue to be high performing in first and second grade, which underscores the importance of early intervention for students who experience difficulties. Finding the same pattern of correlations for multilingual and English monolingual children (as well as ELLs and English monolingual non-ELLs) is also noteworthy, with science becoming less distinct from math and reading in each subsequent school year from kindergarten to second grade. Domain-general skills such as oral language and EF may exert more influence on academics in later grades than when children are younger. Unfortunately, the ECLS-K data did not allow us to investigate such an interaction directly, but the increasing associations between science, math, and reading suggest that increases in oral language and EF may benefit students' science achievement. Oral language is the most likely candidate because other researchers have documented a rise in the contribution of oral language to reading comprehension over time (Adlof et al., 2006; Foorman et al., 2018). As students master word reading skills and the linguistic and background knowledge demands increase, they need strong language skills to read and to comprehend science and math content.

8.2. Prediction of science achievement

In contrast to our hypothesis, the results suggest that domain-general abilities such as language proficiency and EF are similarly involved in performance on science assessments for multilingual and English monolingual children. One possible reason we did not find differences between the two groups in our study is because the multilingual in the ECLS data may not be English language learners, students who meet Title III criteria of the Every Student Succeeds Act of 2015 to receive supplemental services in the public school system to improve their English language proficiency and academic achievement. We therefore cross-validated our results with a subgroup of ELLs and a subgroup of English monolingual non-ELLs, thereby strengthening the inferences that can be drawn from our results. As that may be, because of the manner in which the data were collected, all children who passed the English screener were included, which means even the multilingual children and ELLs in our study had some level of English language proficiency. It is important that we limit the generalizability of our findings to students who live in homes where a non-English language is the primary language spoken but who have some English proficiency. Nonetheless, it is well established that oral language proficiency benefits all children as they acquire academic knowledge. For example, *observe*, *predict*, and *check* are science process skills and vocabulary words that students must understand to be able to engage in the processes. To learn the NGSS cross-cutting concept of cause and effect, students must be able to identify independent and dependent variables, and express them using oral language (e.g., "A steeper ramp makes the marble go faster"). In carrying out hypothesis testing, students make *predictions* about, for example, the change in motion anticipated when two objects touch or collide. The ability to make predictions is a critical oral language skill (LervAag et al., 2018). Students can then carry out their investigation and *observe* the results to *check* their predictions. Thus, science terms, concepts, language, and cross-cutting processes are interconnected and mutually supporting (Gelman & Brenneman, 2004).

With respect to EF, it allows children to sustain attention, attend to specific important task features, avoid distractions, and hold information in memory during classroom instruction (McClelland et al., 2015). EF also supports holding and integrating new information with prior knowledge (Diamond, 2013; Gathercole et al., 2004; Gathercole, Pickering, Ambridge et al., 2004; Pascual-Leone & Johnson, 2021) and enables one to think analytically (Gropen et al., 2011). In addition, EF may subserve science just as it does math and reading. For example, EF permits the child to hold "step one" answers in memory when completing multistep math problems (Cragg & Gilmore, 2014). For science, EF likely helps children revise (or update) initial hypotheses when a response generated through analytical processing (i.e., based on data or experiments) conflicts with one based on experiential processing (i.e., memory). Higher inhibitory control is also associated with greater task perseverance among children (McClelland et al., 2007), which may help children engage in complex reasoning, involving the inhibition of experiential processing (Amsel et al., 2008) when children test and evaluate hypotheses (Gropen et al., 2011; Koerber & Osterhaus, 2019). Regardless of the mechanism, EF is clearly an important constellation of cognitive capacities that subserves most if not all areas of academics.

8.3. Implications

The results of the present study are consistent with longitudinal correlational studies that point to the importance of language in the development of academic skills (e.g., Duncan et al., 2007; Foster et al., 2019; Pace et al., 2019). They are also consistent with reports made by the NASEM (2018) indicating that language proficiency is important to children's science achievement. Awareness of the need to provide effective STEM instruction in elementary school is increasing. However, far fewer teachers feel prepared to teach science (31%) than math (73%) or reading/language arts (77%; Banilower et al., 2018). This translates to less instructional time devoted to science (18 min) than to math (57 min) and reading/language arts (89 min) in each school day (Banilower et al., 2018), which results in roughly 213 fewer hours devoted to science than reading/language arts instruction in a school year. Therefore, there is a critical need to develop innovative science curricula and professional development materials that would increase attention to science in elementary schools and make it easier for teachers to deliver science instruction well. Pre-service and in-service teachers may feel more prepared and confident to teach science if they understand that science knowledge is malleable and critical to developing reading comprehension (Cabell & Hwang, 2020; Cervetti & Wright, 2020).

Given that language proficiency is an underlying competency needed for arguably most, if not all, academic skills and that science, math, and reading achievement are highly correlated, STEM curricula that integrate language, math, and reading targets have the potential to be highly effective (e.g., Cabell & Hwang, 2020; Cabell, White, Kim et al., 2019; Doabler et al., 2021; Neuman & Kaefer, 2018). For example, content-rich English language arts (ELAs) reading instruction can be designed to foster students' linguistic comprehension while also building their STEM knowledge. The construction-integration model of reading comprehension (Kintsch, 1988) and the contemporary science of reading research (see Cervetti, Pearson et al., 2020) bring awareness to a constellation of core repertoires that when intentionally targeted during instruction, benefits to multiple domains can be achieved simultaneously (Wackerle-Hollman et al., 2021; Wright & Gotwals, 2017).

Building students' STEM knowledge can help them connect ideas across sentences in ways that better help them understand science texts. Likewise, instructional practices that intentionally promote academic vocabulary, inferencing, and knowledge through oral retelling and read alouds (Foorman et al., 2016) can facilitate comprehension of advanced STEM content and promote scientific thinking (see Kintsch, 1988). For example, Wright and Gotwals found that an integrated language, literacy, and science curriculum enhanced kindergarten students' receptive vocabulary, application of science concepts, and ability to make claims and provide evidence supports. One promising feature of integrated instruction is that redundancies can be reduced, which will likely result in accelerated learning (Pace et al., 2019), and in turn curtail achievement inequities. Of great importance is the finding that the influence of language proficiency was not different for multilingual compared to monolingual English speakers. While unexpected, this suggests that a single integrated curriculum that centralizes the core constellations of oral language and bridge processes would be sufficient for both groups of children. It would not be necessary to spend valuable resources teaching multilingual one way and monolingual English children another way.

The high within-domain correlations reported in this study suggest the measurement of science across early elementary school is stable as is students' academic standing relative to that of their peers regardless of the academic domain. That is, high-performing students in kindergarten continue to be high performing in first and second grade. Coupled with the stability of these tests, there was evidence of high across-domain associations that increase across elementary school, suggesting that there is redundancy in the content of the assessments administered in the present study. High-stakes standardized tests have posed particular challenges for students from racial minority, language minority, and low-income backgrounds (Amrein & Berliner, 2002; Valenzuela, 2005), leading to calls for test reform (see Au, 2016; Bach, 2020). At the very least, the present results suggest that annual high-stakes testing in every content domain is not necessary. Some researchers suggest that testing in oral language may be the most useful because it is linked to all other domains (Adlof & Hogan, 2019; Pace et al., 2019; Silverman et al., 2019). In addition, given federal initiatives focused on STEM education (e.g., the America COMPETES Reauthorization Act of 2010), assessment design that focuses on vocabulary, comprehension, and academic knowledge aligned to grade-level science standards should be encouraged to monitor the impact of federal investments.

8.4. Limitations and future directions

Despite the breadth of data available in the ECLS-K: 2011 public use data set, it does not include every science or cognitive and linguistic ability one may want access to. In particular, the ECLS-K: 2011 public use data set does not include measures for specific areas of science achievement. It may be that the domain-general abilities such as language proficiency and EF are more strongly related to some areas of science achievement than others. For example, tasks that assess children's understanding of the nature of science may be more reliant on language proficiency whereas tasks related to experimentation, such as understanding the difference between conclusive and inconclusive tests (Sodian et al., 1991) and data interpretation (i.e., interpreting complex patterns of covariation in data, e.g., Koerber et al., 2005), may be more reliant on EF. Future research may benefit from looking more closely at the relationship between language proficiency, EF, and tasks that measure one's understanding of the nature of science, experimentation, and data interpretation. Similarly, including additional cognitive and linguistic abilities shown to be predictive of math and reading (e.g., phonological processing; Cirino et al., 2018) may also support science achievement. Measures of language proficiency that more closely index social and academic language as well as the four domains (reading, writing, speaking, and listening) emphasized in current English language proficiency (ELP) content standards could help disentangle the relationship between language and science achievement. Such studies may be able to show which dimensions of ELP are relevant to science achievement. Finally, while we posit that the relations between science and math may be due to the association between each of the variables and relational reasoning, it was not measured in the ECLS-K: 2011 study. Future studies may benefit from measuring more cognitive and linguistic abilities than those that are available in the ECLS-K: 2011 public use data set.

8.5. Conclusion

While it has been theorized that integrated curricula are more effective than the traditional siloed approach to instruction (Wackerle-Hollman et al., 2021; Wright & Gotwals, 2017), we provide the first empirical evidence in support of integrated science instruction. Moreover, we interpret our results as suggesting that siloed (non-integrated) science curricula are misaligned with the nature of science achievement in early elementary school. Based on a clear body of research that indicates oral language is critical for learning, we posit that intentionally organizing instruction such that science, math, and reading instruction takes advantage of children's oral language may help reduce redundancies and increase instruction time dedicated to science content, thereby contributing to increased learning. At the very least, integrated instruction and cross-domain effects is a promising new area of research (Cabell & Hwang, 2020; Cervetti & Wright, 2020; Tyner & Kabourek, 2020).

Funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Matthew Foster: Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft, Visualization. **Sara Smith:** Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing. **Trina Spencer:** Writing – original draft, Writing – review & editing.

Author note

We have no known conflict of interest to disclose. This research was enhanced by training during the American Educational Research Association (AERA) Institute on Statistical Analysis: Development of Mathematics Competencies in Early Childhood. AERA receives funds for its AERA Grants Program from the National Science Foundation (NSF) under award NSF-DRL #1749275. Opinions reflect those of the author and do not necessarily reflect those of AERA or NSF.

Declaration of interest

None.

Data availability

The data are publicly available.

Appendix

Subgroup Analyses Results.
See Tables A1, A2, A3.

Table A1

Correlations Among all Variables for English Language Learners (ELLs) and English Monolingual non-ELLs (non-ELLs).

Sample	Variables	Time Point	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
ELLs	1. SES	K	**---**											
	2. Language	K	.15	---										
	3. EF	K	.10	.19	---									
	4. Science	K	.23	.54	.24	---								
	5. Math	K	.20	.32	.33	.48	---							
	6. Reading	K	.18	.44	.28	.46	.70	---						
	7. Science	1^{st}	.26	.52	.34	**.65**	.59	.57	---					
	8. Math	1^{st}	.25	.26	.39	.42	**.80**	.60	.63	---				
	9. Reading	1^{st}	.26	.41	.33	.41	.69	**.75**	.63	.71	---			
	10. Science	2^{nd}	.25	.42	.33	**.62**	.60	.52	**.82**	.65	.61	---		
	11. Math	2^{nd}	.27	.27	.35	.46	**.76**	.54	.63	**.84**	.68	.70	---	
	12. Reading	2^{nd}	.26	.33	.29	.41	.67	**.64**	.60	.67	**.87**	.65	.71	---
	Means(SD)		-.84(0.87)	-1.02(1.40)	-0.25(1.07)	-1.33(0.66)	-.58(0.59)	-.50(0.67)	-.46(0.82)	.33(0.52)	.37(0.49)	.35(0.79)	.91(0.52)	.80(0.41)
		Time Point	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
Non-ELLs	1. SES	K	**---**											
	2. Language	K	.17	**---**										
	3. EF	K	.18	.19	**---**									
	4. Science	K	.34	.40	.26	**---**								
	5. Math	K	.40	.34	.33	.55	**---**							
	6. Reading	K	.33	.36	.32	.51	.76	**---**						
	7. Science	1^{st}	.36	.36	.27	**.70**	.58	.53	**---**					
	8. Math	1^{st}	.38	.31	.33	.55	**.82**	.70	.62	---				
	9. Reading	1^{st}	.29	.33	.25	.46	.72	**.84**	.50	.77	**---**			
	10. Science	2^{nd}	.34	.40	.25	**.66**	.61	.57	**.77**	.64	.55	---		
	11. Math	2^{nd}	.40	.36	.36	.53	**.77**	.68	.62	**.85**	.73	.67	---	
	12. Reading	2^{nd}	.35	.30	.25	.46	.67	**.75**	.52	.71	**.86**	.59	.73	**---**
	Means(SD)		.08(0.96)	.08(0.63)	-.05(1.00)	-.62(.67)	-.37(.60)	-.20(.67)	.10(.72)	.51(.49)	.59(.43)	.76(.68)	1.67(.51)	.96(.36)

Note. EF is executive functioning; K is kindergarten; 1st is first grade; 2nd is second grade. Bold is autoregressive correlations. Boxes are concurrent correlations among achievement measures. SES, Lang, and EF were transformed to z-scores. The remaining measures are IRT-based theta scores.

Table A2
Multigroup Multivariate Regression Analyses for English Language Learners (ELLs) and English Monolingual non-ELLs (non-ELLs).

Variable	Science				Math				Reading			
	β	*SE*	*p*	R^2	*β*	*SE*	*p*	R^2	*β*	*SE*	*p*	R^2
Model 1 – Kindergarten												
ELLs				.32				.18				.23
SES	.15	.04	<.001		.15	.05	.001		.12	.04	.004	
Language Proficiency	.48	.04	<.001		.22	.05	<.001		.36	.04	<.001	
Executive Functioning	.13	.04	.001		.27	.05	<.001		.20	.04	<.001	
Non-ELLs				.27				.29				.26
SES	.26	.05	<.001		.31	.05	<.001		.25	.07	<.001	
Language Proficiency	.34	.06	<.001		.25	.06	<.001		.28	.06	<.001	
Executive Functioning	.15	.05	.002		.23	.07	.001		.22	.06	<.001	
Model 2 – 1st Grade												
ELLs				.34				.21				.24
SES	.16	.04	.001		.19	.05	<.001		.19	.04	<.001	
Language Proficiency	.43	.04	<.001		.14	.05	.002		.30	.04	<.001	
Executive Functioning	.25	.04	<.001		.34	.05	<.001		.25	.05	<.001	
Non-ELLs				.23				.24				.17
SES	.28	.06	<.001		.29	.06	<.001		.22	.08	.008	
Language Proficiency	.28	.06	<.001		.21	.06	<.001		.24	.07	<.001	
Executive Functioning	.15	.05	.002		.15	.05	<.001		.15	.07	.031	
Model 3 – 2rd Grade												
ELLs				.27				.22				.20
SES	.17	.05	.001		.21	.05	<.001		.18	.05	<.001	
Language Proficiency	.35	.05	<.001		.20	.06	<.001		.26	.07	<.001	
Executive Functioning	.24	.04	<.001		.29	.05	<.001		.22	.05	<.001	
Non-ELLs				.25				.31				.20
SES	.26	.07	<.001		.31	.06	<.001		.27	.09	.002	
Language Proficiency	.33	.07	<.001		.26	.06	<.001		.22	.06	<.001	
Executive Functioning	.12	.06	.025		.25	.07	<.001		.16	.09	<.061	

Note. Completely standardized results reported. SES, language proficiency, and executive functioning were measured during the spring of participant's kindergarten school year.

Table A3
Tests of Measurement Invariance for English Language Learners (ELLs) and English Monolingual non-ELLs.

Model	BIC	SSABIC	RMSEA	CFI	SRMR
1. Unstructured covariance matrices	13,207	12,845	.061	.981	.070
2. Predicting kindergarten achievement	10,662	10,519	.036	.995	.042
3. Predicting first-grade achievement	9685	9542	.023	.998	.024
4. Predicting second-grade achievement	9195	9052	.019	.998	.030

References

Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational Assessment, 8*(3), 231–257. ⟨https://rb.gy/1impfr⟩.

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

Adlof, S. M., Catts, H. W., & Little, T. D. (2006). Should the simple view of reading include a fluency component? *Reading and Writing, 19*, 933–958. ⟨https://rb.gy/8otnab⟩.

Adlof, S. M., & Hogan, T. P. (2019). If we don't look, we won't see: Measuring language development to inform literacy instruction. *Policy Insights from the Behavioral and Brain Sciences, 6*(2), 210–217. https://doi.org/10.1177/2372732219839075

Amrein, A.L., & Berliner, D.C. (2002). *An analysis of some unintended and negative consequences of high-stakes testing.* Education Policy Research Unit, Arizona State University.

Amsel, E., Klaczynski, P. A., Johnston, A., Bench, S., Close, J., Sadler, E., & Walker, R. (2008). A dual-process account of the development of scientific reasoning: The nature and development of metacognitive intercession skills. *Cognitive Development, 23*, 452–471. https://doi.org/10.1016/j.cogdev.2008.09.002

Ardasheva, Y., Tretter, T. R., & Kinny, M. (2012). English language learners and academic achievement: Revisiting the Threshold Hypothesis. *Language Learning, 3*, 769–812. https://doi.org/10.1111/j.1467-9922.2011.00652.x

Arnold, D. H., & Doctoroff, G. L. (2003). The early education of socioeconomically disadvantaged children. *Annual Review of Psychology, 54*, 517–545. https://doi.org/10.1146/annurev.psych.54.111301.145442

Au, W. (2016). Meritocracy 2.0: High-stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educational Policy, 30*(1), 39–62.

Bach, A. J. (2020). High-stakes, standardized testing and emergent bilingual students in Texas: A call for action. *Texas Journal for Literacy Education, 8*(1), 18–37. ⟨https://rb.gy/cye5tr⟩.

Banilower, E.R., Smith, P.S., Malzahn, K.A., Plumley, C.L., Gordon, E.M., & Hayes, M.L. (2018). Report of the 2018 NSSME. *Horizon Research.*

Barnard-Brak, L., Stevens, T., & Ritter, W. (2017). Reading and mathematics equally important to science achievement: Results from nationally-representative data. *Learning and Individual Differences, 58*, 1–9. https://doi.org/10.1016/j.lindif.2017.07.001

Beck, I., McKeown, M., & Gromoll, E. (1989). Learning from social studies texts. *Cognition and Instruction, 6*, 99–158. https://doi.org/10.1207/s1532690xci0602_1

Blums, A., Belsky, J., Grimm, K., & Chen, Z. (2017). Building links between early socioeconomic status, cognitive ability, and math and science achievement. *Journal of Cognition and Development, 1*, 16–40. https://doi.org/10.1080/15248372.2016.1228652

Brown, T.A. (2006). *Confirmatory factor analysis for applied research.* Guilford Press.

Bull, R., Espy, K. A., & Wiebe, S. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 3*, 205–228 (https://doi.org/dqgnkn).

Bullock, M., Sodian, B., & Koerber, S. (2009). Doing experiments and understanding science: Development of scientific reasoning from childhood to adulthood. In W. Schneider, & M. Bullock (Eds.), *Human development from early childhood to early adulthood. Findings from the Munich longitudinal study* (pp. 173–197). Lawrence Erlbaum.

Cabell, S. Q., & Hwang, H. (2020). Building content knowledge to boost comprehension in the primary grades. *Reading Research Quarterly, 72*(4), 499–507. https://doi.org/gk53q6.

Cabell, S.Q., White, T.G., Kim, J., Hwang, H., & Gale, C. (2019, December). Impact of the Core Knowledge Language Arts read-aloud program on kindergarteners' vocabulary, listening comprehension, and general knowledge. Paper presented at the annual meeting of the Literacy Research Association, Tampa, FL.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*, 595–616. https://doi.org/10.1207/s15326942dn2802_3

Cervetti, G. N., Pearson, P. D., Palincsar, A. S., Afflerbach, P., Kendeou, P., Biancarosa, G., … Berman, A. I. (2020). How the Reading for Understanding initiative's research complicates the simple view of reading invoked in the science of reading. *Reading Research Quarterly, 55*, S161–S172 (https://doi.org/gkzkmn).

Cervetti, G. N., & Wright, T. S. (2020). The role of knowledge in understanding and learning from text. In E. B. Moje, P. Afflerbach, P. Enciso, & N. K. Leseaux (Eds.), *Handbook of reading research* (Vol. 5, pp. 237–260). Routledge.

Chen, Z., Hononichl, R., Kennedy, D., & Tan, E. (2016). Aiming to complete the matrix: Eye- movement analysis of processing strategies in children's relational thinking. *Developmental Psychology, 52*(6), 867–878. https://doi.org/10.1037/dev0000113

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development, 70*, 1098–1120. https://doi.org/dqk698.

Choi, J. Y., Jeon, S., & Lippard, C. (2018). Dual language learning, inhibitory control, and math achievement in Head Start and kindergarten. *Early Childhood Research Quarterly, 42*, 66–78. https://doi.org/10.1016/j.ecresq.2017.09.001

Cirino, P. T., Child, A. E., & Macdonald, K. T. (2018). Longitudinal predictors of the overlap between reading and math skills. *Contemporary Educational Psychology, 54*, 99–111. https://doi.org/10.1016/j.cedpsych.2018.06.002

Clark, C. A. C., Pritchard, V. E., & Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology, 5*, 1176–1191. https://doi.org/10.1037/a0019672

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Multiple regression/correlation with two or more independent variables. *In applied multiple regression/correlation analysis for the behavioral science* (3rd ed., pp. 64–99). Lawerence Erlbaum Associates,.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education, 3*(2), 63–68. https://doi.org/10.1016/j.tine.2013.12.001

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 6*, 934–945. https://doi.org/10.1037/0012-1649.33.6.934

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Doabler, C. T., Therrien, W. J., Longhi, M. A., Roberts, G., Hess, K., Maddox, S. A., … Toprac, P. (2021). Efficacy of a second-grade science program: Increasing science outcomes for all learners. *Remedial and Special Education, 42*(3), 140–154. https://doi.org/10.1177/0741932521989091

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., … Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duncan, S.E., & De Avila, E.A. (2000). *preLAS Cue Picture Book English Form C*. CTB/McGraw-Hill.

Espinosa, L. M. (2015). Challenges and benefits of early bilingualism in the United States' context. *Global Education Review, 2*, 40–53.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *High School Journal, 91*(1), 29–42. ⟨https://rb.gy/nhudyp⟩.

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C.A., Dimino, J., … Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade (NCEE 2016-4008)*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Foorman, B. R., Petscher, Y., & Herrera, S. (2018). Unique and common effects of decoding and language factors in predicting reading comprehension in grades 1–10. *Learning and Individual Differences, 63*, 12–23. https://doi.org/10.1016/j.lindif.2018.02.011

Foster, M. E., Anthony, J. L., Clements, D. H., & Sarama, J. (2015). Processes in thedevelopment of mathematics in kindergarten children from Title 1 schools. *Journal of Experimental Child Psychology, 140*, 56–73. https://doi.org/10.1016/j.jecp.2015.07.004

Foster, M. E., Anthony, J. L., Zucker, T. A., & Branum-Martin, L. (2019). The prediction of English and Spanish kindergarten math achievement from English and Spanish cognitive and linguistic abilities in Hispanic dual language learners. *Early Childhood Research Quarterly, 46*(1), 213–227. https://doi.org/10.1016/j.ecresq.2018.02.007

Galindo, C. (2010). English language learners' math and reading achievement trajectories in the elementary grades. In E. Garcia, & E. Frede (Eds.), *Young English language learners- Current research and emerging directions for practice and policy* (pp. 42–58). Teachers College Press.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004a). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004b). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*(1), 1–16. https://doi.org/10.1002/acp.934

Geary, D. C. (2007). An evolutionary perspective on learning disabilities in mathematics. *Developmental Neuropsychology, 32*(1), 471–519. ⟨https://rb.gy/3hg99g⟩.

Gelman, R., & Brenneman, K. (2004). Science learning pathways for young children. *Early Childhood Research Quarterly, 19*, 150–158. ⟨https://rb.gy/v65gwm⟩.

Greenfield, D. B., Jiroux, J., Dominguez, M. X., Greenberg, A. C., Maier, M. E., & Fuccillo, J. M. (2009). Science in the preschool classroom: A programmatic research agenda to improve science readiness. *Early Development and Education, 20*(2), 238–264. https://doi.org/10.1080/10409280802595441

Gropen, J., Clark-Chiarelli, N., Hoisington, C., & Ehrlich, S. B. (2011). The importance of executive function in early science education. *Child Development Perspectives, 5*(4), 298–304. https://doi.org/10.1111/j.1750-8606.2011.00201.x

Hartanto, A., Yang, H., & Yang, S. (2018). Bilingualism positively predicts mathematical competence: Evidence from two large-scale studies. *Learning and Individual Differences, 61*, 216–227. https://doi.org/10.1016/j.lindif.2017.12.007

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16*(3), 174–180. ⟨https://rb.gy/forjmh⟩.

Horn, C. (2003). High-stakes testing and students: Stopping or perpetuating a cycle of failure. *Theory into Practice, 42*(1), 30–41. ⟨https://www.jstor.org/stable/1477316⟩.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427–466. https://doi.org/10.1037/0033-295X.104.3.427

Hwang, H. (2020). Early general knowledge predicts English reading growth in bilingual and monolingual students throughout the elementary school years. *The Elementary School Journal, 121*(1), 154–178. https://doi.org/10.1086/709857

Inan, H. Z., Trundle, K. C., & Kantor, R. (2010). Understanding natural sciences education in a Reggio Emilia-inspired preschool. *Journal of Research in Science Teaching, 47*(10), 1186–1208. https://doi.org/10.1002/tea.20375

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. Basic Books.

Kanari, Z., & Millar, R. (2004). Reasoning from data: How students collect and interpret data in science investigations. *Journal of Research in Science Teaching, 7*, 748–769. https://doi.org/10.1002/tea.20020

Kempert, S., Saalbach, H., & Hardy, I. (2011). Cognitive benefits and costs of bilingualism in elementary school students: The case of mathematical word problems. *Journal of Educational Psychology, 103*(3), 547–561. https://doi.org/10.1037/a0023619

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction- integration model. *Psychological Review, 95*, 163–182. ⟨https://rb.gy/p15iqt⟩.
Koerber, S., & Osterhaus, C. (2019). Individual differences in early scientific thinking: Assessment, cognitive influences, and their relevance for science learning. *Journal of Cognition and Development, 20*(4), 510–533. ⟨https://rb.gy/emuzgv⟩.
Koerber, S., Sodian, B., Thoermer, C., & Nett, U. (2005). Scientific reasoning in young children: Preschoolers' ability to evaluate covariation evidence. *Swiss Journal of Psychology, 64*, 141–152. https://doi.org/10.1024/1421-0185.64.3.141
Kuhn, D. (2011). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Handbook of childhood cognitive development* (2nd ed., pp. 497–523). Wiley.
LervAag, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the developmental relationship between oral language skills and reading comprehension: It's simple, but complex. *Child Development, 89*(5), 1821–1838. https://doi.org/10.1111/cdev.12861
Maerten-Rivera, J., Myers, N., Lee, O., & Penfield, R. (2010). Student and school predictors of high-stakes assessment in science. *Science Education, 94*(6), 937–962. https://doi.org/10.1002/sce.20408
Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2
Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction, 29*, 43–55. https://doi.org/10.1016/j.learninstruc.2013.07.005
McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43*(4), 947–959. https://doi.org/fc6dnn.
McClelland, M. M., John Geldhof, G., Cameron, C. E., & Wanless, S. B. (2015). Development and self-regulation. In R. M. Lerner (Ed.) (7th ed.,*Handbook of child psychology and developmental science* (7th ed.,, *4* pp. 1–43). John Wiley and Sons. https://doi.org/10.1002/9781118963418.childpsy114.
McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes, 22*, 247–288. ⟨https://rb.gy/w79hrx⟩.
McNeil, N. M., & Alibali, M. W. (2005). Knowledge change as a function of mathematics experience: All contexts are not created equal. In *Journal of Cognitive Development, 6* pp. 285–306). https://doi.org/10.1207/s15327647jcd0602_6
Miller Singley, A. T., & Bunge, S. A. (2014). Neurodevelopment of relational reasoning: Implications for mathematical pedagogy. *Trends in Neuroscience and Education, 3*, 33–37. https://doi.org/10.1016/j.tine.2014.03.001
Morra, S., Panesi, S., Traverso, L., & Usai, M. C. (2018). Which tasks measure what? Reflections on executive function development and a commentary on Podjarny, Kamawar, and Andrews. *Journal of Experimental Child Psychology, 167*, 246–258. https://doi.org/10.1016/j.jecp.2017.11.004
Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher, 45*(1), 18–35. ⟨https://www.jstor.org/stable/43996893⟩.
Morris, B. J., Croker, S., Masnick, A. M., & Zimmerman, C. (2012). The emergence of scientific reasoning. In H. Kloos, B. J. Morris, & J. L. Amaral (Eds.), *Current topics in children's learning and cognition* (pp. 61–82). InTech.
Najarian, M., Tourangeau, K., Nord, C., & Wallner-Allen, K. (2018). *Early Childhood Longitudinal Study, Kindergarten class of 2010–11 (ECLS-K: 2011): First- and second-grade psychometric report (NCES 2018–183)*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
National Academies of Sciences, Engineering, & Medicine. (2018). *English learners in STEM subjects: Transforming classrooms, schools, and lives*. National Academies Press. https://doi.org/10.17226/25182
National Assessment Governing Board. (1994). *Mathematics framework for the 1996 and 2000 National Assessment of Educational Progress.* Retrieved from ⟨https://rb.gy/p1kx7y⟩.
National Assessment Governing Board. (2004). *Mathematics framework for the 2005 National Assessment of Educational Progress*. Retrieved from ⟨https://rb.gy/rnh4qq⟩.
National Assessment Governing Board. (2008). *Reading framework for the 2009 National Assessment of Educational Progress.* Retrieved from ⟨https://rb.gy/pttrhi⟩.
National Center for Education Statistics. (1999). *Kindergarten through eighth grade public-use file* [Data set]. Retrieved from ⟨https://nces.ed.gov/ecls/dataproducts.asp#K-8⟩.
National Center for Education Statistics. (2011). *Kindergarten through fifth grade public-use file* [Data set]. Retrieved from ⟨https://nces.ed.gov/ecls/dataproducts.asp#K-5⟩.
National Council of Teachers of Mathematics. (2001). *Principles and standards for school mathematics*. Retrieved from ⟨https://www.nctm.org/Standards-and-Positions/Principles-and-Standards/⟩.
National Mathematics Advisory Panel. (2008). *The Final Report of the National Mathematics Advisory Panel.* U.S. Department of Education.Retrieved from ⟨https://rb.gy/xeubhc⟩.
National Research Council. (2013). *Next generation science standards: for states, by states*. National Academies Press,. https://doi.org/10.17226/18290
Neuman, S. B., & Kaefer, T. (2018). Developing low-income children's vocabulary and content knowledge through a shared book reading program. *Contemporary Educational Psychology, 52*, 15–24. https://doi.org/10.1016/j.cedpsych.2017.12.001
O'Neal, D., & Ringler, M. (2010). Broadening our view of linguistic diversity. *Phi Delta Kaplan, 91*(7), 48–52. https://doi.org/10.1177/003172171009100710
O'Reilly, T., & McNamara, D. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal, 44*(1), 161–196. https://doi.org/10.3102/0002831206298171
Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., & Hirsh-Pasek, K. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly, 46*(1), 112–125. https://doi.org/10.1016/j.ecresq.2018.04.001
Panesi, S., & Morra, S. (2020). Executive functions and mental attentional capacity in preschoolers. *Journal of Cognition and Development, 21*, 72–91. https://doi.org/h3c5.
Park, M., Zong, J., & Batalova, J. (2018). *Growing superdiversity among young U.S. dual language learners and its implications.* Migration Policy Institute.
Pascual-Leone, J., & Johnson, J. M. (2021). *The working mind: Meaning and mental attention in human development*. MIT Press. ⟨https://rb.gy/4zqr4t⟩.
Pew Research Center. (2018). *Women and men inSTEM often at odds over workplace Equity*. Retrieved from ⟨https://rb.gy/4zqr4t⟩.
Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology, 1*. Wiley.
Piekny, J., Grube, D., & Maehler, C. (2013). The relation between preschool children's false-belief understanding and domain-general experimentation skills. *Metacognition and Learning, 8*, 103–119. https://doi.org/10.1007/s11409-013-9097-4
Piekny, J., & Maehler, C. (2013). Scientific reasoning in early and middle childhood: The development of domain general evidence evaluation, experimentation, and hypothesis generation skills. *British Journal of Developmental Psychology, 31*, 153–179. https://doi.org/10.1111/j.2044-835X.2012.02082.x
Ruffman, T., Perner, J., Olson, D. R., & Doherty, M. (1993). Reflecting on scientific thinking: Children's understanding of the hypothesis-evidence relation. *Child Development, 64*, 1617–1636. https://doi.org/10.1111/j.1467-8624.1993.tb04203.x
Schleicher, A. (2018). *PISA 2018: Insights and interpretations*. Organisation for Economic Co- operation and Development. Retrieved from ⟨https://www.oecd.org/pisa/PISA%202018%20Insights%20and%20Interpretations%20FINAL%20PDF.pdf⟩.
Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal, 41*, 159–189. ⟨https://www.jstor.org/stable/3699387⟩.
Silverman, R. D., McNeish, D., Speece, D. L., & Ritchey, K. D. (2019). Early screening for decoding-and language-related reading difficulties in first and third grades. *Assessment for Effective Intervention, 46*(2), 99–109. https://doi.org/10.1177/1534508419857234
Snow, C. (2002). Reading for understanding: Toward an R&D program in reading comprehension. *RAND*. ⟨https://www.rand.org/pubs/monograph_reports/MR1465.html⟩.
Sobel, D. M., Erb, C. D., Tassin, T., & Weisberg, D. S. (2017). The development of diagnostic inference about uncertain causes. *Journal of Cognition and Development, 18*, 556–576. https://doi.org/10.1080/15248372.2017.1387117
Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development, 62*, 753–766. https://doi.org/10.2307/1131175

Swanson, H. L., Kong, J., & Petcu, S. D. (2020). Math problem-solving and cognition among emerging bilingual children at risk and not at risk for math difficulties. *Child Neuropsychology, 26*(4), 489–517. https://doi.org/10.1080/09297049.2019.1674268

Tate, W. F., Jones, B. D., Thorne-Wallington, E., & Hogrebe, M. C. (2012). Science and the city: Thinking geospatially about opportunity to learn. *Urban Education, 47*, 399–433. https://doi.org/10.1177/0042085911429974

Thompson, R., & Zamboanga, B. (2004). Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology, 96*, 778–784. https://doi.org/10.1037/0022-0663.96.4.778

Tourangeau, K., Nord, C., Lê, T., Sorongon, A.G., Hagedorn, M.C., Daly, P., & Najarian, M. (2015). *Early Childhood Longitudinal Study, Kindergarten class of 2010–11 (ECLS– K: 2011): User's manual for the ECLS–K: 2011 kindergarten data file and electronic codebook, public version* (NCES 2015-074). U.S. Department of Education. National Center for Education Statistics.

Tourangeau, K., Nord, C., Lê, T., Wallner-Allen, K., Vaden-Kiernan, N., Blaker, L., & Najarian, M. (2017). *Early Childhood Longitudinal Study, Kindergarten class of 2010–11 (ECLS-K: 2011): User's manual for the ECLS–K:2011 kindergarten–second grade data file and electronic codebook, public version* (NCES 2017-285). U.S. Department of Education. *National Center for Education Statistics*.

Tyner, A., & Kabourek, S. (2020). Social studies instruction and reading comprehension: Evidence from the Early Childhood Longitudinal Study. *Thomas B Fordham Institute*. ⟨https://rb.gy/361ekf⟩.

U.S. Department of Health & Human Services & U.S. Department of Education. (2017). *Policy statement on supporting the development of children who are dual language learners in early childhood programs*. Retrieved from ⟨https://rb.gy/mot97u⟩.

Valenzuela, A. (2005). *Leaving children behind: How "Texas-style" accountability fails Latino youth*. State University of New York.

van der Graaf, J., Segers, E., & Verhoeven, L. (2015). Scientific reasoning abilities in kindergarten: Dynamic assessment of the control of variables strategy. In *Instructional Science, 43* pp. 381–400). https://doi.org/10.1007/s11251-015-9344-y

van der Graaf, J., Segers, E., & Verhoeven, L. (2016). Scientific reasoning in kindergarten: Cognitive factors in experimentation and evidence evaluation. *Learning and Individual Differences, 49*, 190–200. https://doi.org/10.1016/j.lindif.2016.06.006

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47*, 513–539. https://doi.org/10.1146/annurev.psych.47.1.513

VanLehn, K. (1998). Analogy events: How examples are used during problem solving. *Cognitive Science, 22*, 347–388. https://doi.org/10.1207/s15516709cog2203_4

Voss, J., & Silfies, L. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction, 14*, 45–68. https://doi.org/10.1207/s1532690xci1401_2

Vygotsky, L. (1962). *Thought and language*. MIT Press.

Wackerle-Hollman, A., Spencer, T. D., Artman-Meeker, K., Kelley, E., Durán, L., & Foster, M. E. (2021). Multi-tiered system of support in early childhood: Identifying gaps, considerations for application and solutions. *Early Childhood Research Quarterly, 56*, 201–212. https://doi.org/10.1016/j.ecresq.2021.03.010

Weston, R., & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist, 34*, 719–751. https://doi.org/10.1177/0011000006286345

Willoughby, M. T., Wylie, A. C., & Little, M. H. (2019). Testing longitudinal associations between executive function and academic achievement. *Developmental Psychology, 55*(4), 767–779. https://doi.org/10.1037/dev0000664

Wright, T. S., & Gotwals, A. W. (2017). Supporting kindergartners' science talk in the context of an integrated science and disciplinary literacy curriculum. *The Elementary School Journal, 117*(3), 513–537. https://doi.org/10.1086/690273

Zabrucky, K., & Ratner, H. H. (1992). Effects of passage type on comprehension monitoring and recall in good and poor readers. *Journal of Reading Behavior, 24*, 373–391.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method for assessing executive functioning in children. *Nature Protocols, 1*(1), 297–301. ⟨https://rb.gy/atarxj⟩.

Zimmerman, C. (2007). The development of scientific-thinking skills in elementary and middle school. *Developmental Review, 27*, 172–223. https://doi.org/10.1016/j.dr.2006.12.001