Cognitive Development 62 (2022) 101154

Contents lists available at ScienceDirect

Cognitive Development

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

Numeracy skills mediate the relation between executive function and mathematics achievement in early childhood

Jenny Yun-Chen Chan [a,*,1], Nicole R. Scalise [b]

[a] Institute of Child Development, University of Minnesota – Twin Cities, 51 East River Parkway, Minneapolis, MN 55455, USA
[b] Department of Cognitive Sciences, University of California, Irvine Social and Behavioral Sciences Gateway, Irvine, CA 92697-5100, USA

A R T I C L E I N F O

Keywords:
Executive function
Numeracy skills
Preschool
Kindergarten
Mathematics achievement
Early childhood

A B S T R A C T

We examine a potential mechanism underlying the relation between executive function (EF) skills and mathematics achievement in early childhood. Across two samples of three- to six-year-olds in preschool and kindergarten, we found that children's EF skills predict their concurrent skills in set counting, numeral identification, number comparison, and number line estimation. The effects of EF on later numeral identification, number comparison, and number line estimation skills remained significant for these two samples of children, but these effects attenuated when controlling for the respective earlier numeracy skill. Further, aspects of numeracy skills mediated the association between EF and mathematics achievement in both samples. Together, these findings provide evidence on the nuanced relations between EF, numeracy, and mathematics achievement, and suggest attention to each numeracy skill in order to support early mathematical development.

1. Introduction

A wealth of research has shown that children's executive function (EF) skills relate to their mathematics achievement both concurrently (Best, Miller, & Naglieri, 2011; Bull & Scerif, 2001) and over time (Clark et al., 2014; Mazzocco & Kover, 2007; Ribner, Willoughby, & Blair, 2017). There are various theoretical models to explain this relation, with some suggesting that EF supports mathematical skills development (e.g., Case, 1978; Goodrich, Peng, Bohaty, Leiva, & Thayer, 2021; Nguyen, Duncan, & Bailey, 2019; Ribner, 2020), and with others suggesting bidirectional relations (e.g., Cameron, Kim, Duncan, Becker, & McClelland, 2019; Clements, Sarama, & Germeroth, 2016; Fuhs, Nesbitt, Farran, & Dong, 2014; Nesbitt, Fuhs, & Farran, 2019). One potential mechanism underlying this relation is that children's EF skills support their developing numeracy skills, such as counting and comparing numbers, and in turn, these numeracy skills predict children's mathematics achievement (Chu, vanMarle, & Geary, 2016; Fuhs, Hornburg, & McNeil, 2016).

Here, we investigate the relations between EF, numeracy skills, and mathematics achievement in early childhood with a focus on four numeracy skills—set counting, numeral identification, number comparison, and number line estimation—as potential pathways between EF and mathematics achievement. We focus on these numeracy skills because they have been identified as important early predictors of later mathematics achievement (Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Mazzocco & Thompson, 2005; Schneider et al., 2018). Specifically, we examine (a) the concurrent association and (b) longitudinal relation between EF and these four numeracy

* Corresponding author.
E-mail address: jchan2@wpi.edu (J.Y.-C. Chan).
[1] Dr. Chan is currently at the Department of Social Science and Policy Studies, Worcester Polytechnic Institute, Worcester, MA 01609, USA.

https://doi.org/10.1016/j.cogdev.2022.101154
Received 13 April 2021; Received in revised form 21 December 2021; Accepted 5 January 2022
Available online 15 January 2022

skills; we also test (c) these skills as potential pathways between EF and mathematics achievement. By doing so in two samples of children, we aim to better understand the roles of EF in mathematical development among children varying in age and demographic background.

1. .1. Executive function and its role in mathematics achievement

Executive function skills are a set of top-down mental processes that support explicit and effortful control of one's attention, thoughts, and actions (Carlson, Zelazo, & Faja, 2013; Diamond, 2013; Miyake et al., 2000). They are typically conceptualized as comprising three distinct but related components: working memory, inhibition, and cognitive flexibility. *Working memory* refers to the ability to retain and manipulate information during a task. *Inhibition* refers to the ability to suppress or delay habitual or pre-potent response. *Cognitive flexibility* refers to the ability to consider multiple options simultaneously or shift attention flexibly between them. Although theoretically distinct, these three components of EF skills appear to emerge as a unitary construct in two- to six-year-olds (Wiebe, Espy, & Charak, 2008; Wiebe et al., 2011), and begin to differentiate into three separable components during middle childhood and into adulthood (Clark, Sheffield, Wiebe, & Espy, 2013; Lee, Bull, & Ho, 2013; Shing, Lindenberger, Diamond, Li, & Davidson, 2010).

EF skills are theorized to be central to the information processing required by mathematics, and may influence children's mathematics achievement in several ways. For instance, working memory allows children to keep different quantities in mind and manipulate them while solving problems; inhibition helps children to ignore irrelevant information or suppress inappropriate strategies; cognitive flexibility allows children to shift attention between solution strategies, response options, or aspects of a problem (Bull & Lee, 2014). The association between EF and mathematical skills is supported by a substantial body of work on their concurrent correlations and predictive relations (Cragg & Gilmore, 2014; Fuhs et al., 2014; Geary, Hoard, & Nugent, 2012). For instance, preschoolers' EF skills significantly predict their concurrent mathematics achievement even after controlling for child age and intelligence (Bull, Espy, Wiebe, Sheffield, & Nelson, 2011). A longitudinal study that followed children over a two-year period found that children's EF skills at age three significantly predicted their numeracy skills nine months and eighteen months later as well as their kindergarten mathematics achievement, controlling for socioeconomic status (SES), language skills, and processing speed (Clark et al., 2013). Together, these studies support the notion that EF skills influence mathematics achievement in early childhood, but offer little empirical evidence on the mechanisms underlying this relation. In the current study, we examine the relative influence of EF on four early numeracy skills, and ask whether the relation between EF and mathematics achievement may be mediated through these numeracy skills.

1.2. Early numeracy skills as mediators of EF and mathematics achievement

Early numeracy skills encompass informal skills such as counting and comparing numerical magnitudes, and children often acquire these skills prior to formal schooling (Purpura & Lonigan, 2013; Raghubar & Barnes, 2017). These early numeracy skills are consistent predictors of children's later mathematics achievement formally acquired through explicit instruction (e.g. Duncan et al., 2007; Geary, 2011; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Watts, Duncan, Siegler, & Davis-Kean, 2014). In the current study, we focus on four distinct numeracy skills—set counting, numeral identification, number comparison, and number line estimation—as potential pathways between EF and mathematics achievement among three- to five-year-old preschoolers (Study 1) and five- to six-year-old kindergartners (Study 2). Each of the selected numeracy skills represents a building block key to supporting children's mathematics achievement. Below, we provide an overview on each numeracy skill and its relation with EF and mathematics achievement, followed by a review of prior work that examined a combination of these numeracy skills in one study.

1.2.1. Set counting. One of the first numeracy skills that children develop is set counting, mapping number words with sets of objects. Although children can recite the counting sequence as early as two years of age (Gelman & Gallistel, 1978), it may take them a year or more to correctly count sets of objects up to five (Baroody & Price, 1983; Schaeffer, Eggleston, & Scott, 1974). Children's ability to count sets of objects in the fall of preschool predicts their mathematics achievement at the end of the school year (Chu, vanMarle, & Geary, 2015), and children's counting skills at the entry of kindergarten significantly relate to their general mathematics achievement across kindergarten and first grade (Jordan, Kaplan, Locuniak, & Ramineni, 2007).

While counting is a foundational skill that contributes to mathematics achievement, its developmental process is complex and protracted as it may rely on other cognitive skills, such as EF. For young children who are still developing representations of number words, counting sets of objects may require children to remember the verbal counting sequence, assign each number to one (and only one) object, and inhibit the urge to count each object multiple times (Gelman & Gallistel, 1978). Indeed, studies have shown a correlation between children's EF and set counting skills (Purpura, Schmitt, & Ganley, 2017; Scalise & Ramani, 2021).

1.2.2. Numeral identification. In addition to learning the meanings of number words, young children must also learn about written number symbols, such as Arabic numerals. Although knowledge of number words via the counting sequence occurs first in development, young children quickly develop knowledge of numerals. By age four, approximately one-quarter of children can correctly label numerals 1 to 9 (Ginsburg & Baroody, 2003). Numeral knowledge represents an important link to children's success in formal mathematics, since much of school-based mathematics requires children to correctly identify and produce written numerals (Purpura, Baroody, & Lonigan, 2013). Indeed, preschoolers' ability to identify written numerals is a significant predictor of their later mathematics ability (Chard et al., 2005; Clarke & Shinn, 2004).

The process of learning to identify numerals may require EF skills as children need to remember the verbal labels and connect them with visual representations of numbers. For some children, this process may involve reciting the counting sequence out loud or mentally to remind themselves of the numeral name. Moreover, children's numeral identification skill is significantly correlated with aspects of their EF skills, suggesting that EF skills may support children's learning of numerals (Purpura et al., 2017).

1.2.3. Number comparison. After children understand how individual number words and numerals map onto quantities, they begin to understand the relations between numbers. In the integrated theory of numerical development, Siegler and colleagues (2011, 2014, 2016) propose that an understanding of the relative magnitude of numbers underlies mathematical learning across the lifespan. An accurate representation of numerical magnitude may support more advanced mathematics problem-solving by helping children learn arithmetic, select appropriate calculation strategies, and assess the plausibility of their answers (Mussolin, Nys, Leybaert, & Content, 2013). Magnitude understanding is correlated with mathematical competence in children and adults (Schneider et al., 2017, 2018), and the effects of magnitude interventions have generalized to other mathematical skills, including arithmetic (Honoré & Noël, 2016; Park & Brannon, 2013). One method of assessing symbolic magnitude knowledge is through comparisons of number words and written symbols (i.e., "which is more, 2 or 7?"). Knowledge of number comparison strongly predicts performance on mathematics achievement measures, such as mental arithmetic (Schneider et al., 2017).

To compare numbers, such as four and five, children may need to recall their memorized counting sequence, and shift attention between the two numbers while inhibiting other proximal numbers, such as three and six. Children's number comparison skill is correlated with their concurrent inhibition (Scalise & Ramani, 2021) and working memory (Purpura et al., 2017) in preschool, and both components of EF skills in kindergarten (Gashaj, Oberer, Mast, & Roebers, 2019a; Gashaj, Uehlinger, & Roebers, 2016). These findings suggest the potential importance of EF skills in the development of magnitude understanding in early childhood.

1.2.4. Number line estimation. Another method of measuring symbolic magnitude knowledge is number line estimation tasks, in which children mark where a target number would fall on a number line. Children with a strong understanding of numerical magnitude are expected to make estimates that are close to where the target number would actually appear on a fully partitioned number line. Although young children with knowledge of the counting sequence may correctly order the numbers, they often compress the space between their estimates of larger numbers (Siegler, 2016). As children gain more experience with numbers through counting and comparison, their estimates become more accurate and linear (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Laski & Siegler, 2007; Siegler & Booth, 2004). A meta-analysis by Schneider and colleagues (2018) estimated the strength of the relation between number line estimation skill and mathematical competence to be $r = 0.443$ among four- to fourteen-year-olds, suggesting their close relation across development.

To accurately estimate numbers on a number line, children must keep both endpoints and the target number in mind while determining an appropriate placement for the target number; this process may involve thinking about the relations of these numbers in the counting sequence. Moreover, children must inhibit the urge to place the target number at either endpoint or in the middle of the page. Working memory and cognitive flexibility are correlated with concurrent performance on number line estimation in four- to eight-year-olds (Friso-van den Bos, Kolkman, Kroesbergen & Leseman, 2014; Gashaj et al., 2016; Gashaj, Oberer, Mast, & Roebers, 2019a). Further, children's working memory capacity positively predicts their improvement on number line estimation after six sessions of number training (Kolkman, Hoijtink, Kroesbergen, & Leseman, 2013), suggesting its importance in the development of number line estimation skill.

1.2.5. Investigations of multiple numeracy skills. Prior research has also considered the interrelations of these numeracy skills and their unique relations with EF and mathematics achievement. For instance, children's set counting, numeral identification, and number comparison skills in kindergarten each uniquely predict their mathematics achievement in second and third grade (Mazzocco & Thompson, 2005). Similarly, children's number comparison and number line estimation skills each uniquely predict their arithmetic performance in first and second grade (Lyons et al., 2014). All of these numeracy skills develop in early childhood, but each skill taps a distinct aspect of numeracy and may require different levels of EF skills at different developmental stages (Purpura et al., 2017).

When examining numeracy skills as potential mediators between EF and mathematics achievement, Gashaj and colleagues (2019b) found that children's number comparison and number line estimation skills significantly mediated the association between EF skills in preschool and mathematics achievement in second grade. However, Fuhs and colleagues (2016) found that EF skills predicted children's number line estimation skill, and set counting skill predicted children's mathematics achievement in kindergarten, yet neither numeracy skill mediated the association between EF and mathematics achievement. Whereas set counting and numeral identification tap the skills of mapping between quantities, numerals, and number words, number comparison and number line estimation tap the skills of connecting numbers in terms of cardinal values and spatial positions (Aunio et al., 2006; Purpura & Lonigan, 2013). By studying these four early numeracy skills and their relations with EF and mathematics achievement, we aim to examine the roles of these distinct skills in the association between EF and mathematics achievement.

In summary, the four numeracy skills reviewed above have each been linked to children's EF skills and mathematics achievement. However, the ways in which EF supports these four numeracy skills may differ by the particular skill and by children's developmental level. Because set counting and numeral identification tend to precede number comparison and number line estimation in development (Siegel, 1971; Fuson, 1988), EF skills may be called on to support set counting and numeral identification early in development. As children acquire set counting and numeral identification skills, these tasks may not require as much EF skills, whereas number comparison and number line estimation may continue to demand EF skills later in development. Further, given the skills involved in number line estimation (e.g., attending to endpoints and the target number, shifting between numbers while comparing their cardinal values and ordinal positions), number line estimation may be more closely related to and consistently engage children's EF skills compared to other numeracy skills. Prior work has demonstrated significant associations between set counting, numeral identification, and EF skills among three- to five-year-olds (Purpura et al., 2017), and between number comparison, number line estimation, and EF skills among six-year-olds (Gashaj et al., 2016). Further, among six-year-olds, EF skills appear to be a stronger predictor of number line estimation than of number comparison skills (Gashaj et al., 2019a). Here, we extend prior work by examining the relations between EF, mathematics achievement, and the four numeracy skills in two samples of children to test the robustness and consistency of these relations in early childhood.

1.3. The current study

Our overarching research questions (RQ) and corresponding hypotheses are as follows:

1. *Does the relation between EF and numeracy skills vary depending on the specific numeracy skill?*
   Given the complexity of number line estimation and the strong association between EF and number line estimation skills (Gashaj et al., 2019a), we hypothesize that EF may be more strongly related to number line estimation than other numeracy skills.
2. *Does the pattern of the relation between EF and numeracy skills change over time?*
   As children become better at counting and reading numerals, the effects of EF on these skills may decrease over time. Given the prior findings on the association between EF and number line estimation skills in kindergartners (Gashaj et al., 2016, 2019a), the influences of EF on number line estimation skill may remain strong over time.
3. *Do the four numeracy skills independently mediate the association between EF and mathematics achievement?*
   Because each numeracy skill may tap different components of EF skills (e.g., Purpura et al., 2017; Gashaj et al., 2016) and contribute to distinct aspects of mathematics achievement, the four numeracy skills may independently mediate the association between EF and mathematics achievement.

We first addressed our research questions with a sample of three- to five-year-old preschoolers (Study 1), then repeated the analyses with another sample of five- to six-year-old kindergartners (Study 2). These two studies differed in participant age and demographic backgrounds as well as the measures of EF and mathematics achievement, limiting direct comparisons of the findings. For example, we measured children's addition skills as a proxy of their mathematics achievement in Study 1, and used a standardized assessment that tapped a broad set of number principles in Study 2. While both types of measures focused on number rather than other domains of mathematics (e.g., geometry), our approaches in the two studies were consistent with prior research that focused on children's knowledge of number principles as indicators of their broader mathematics achievement (e.g., Best et al., 2011; Fuhs et al., 2016; Gashaj, Oberer, Mast, & Roebers, 2019b; Jordan et al., 2007; Ribner et al., 2017). Together, the two studies provided opportunities to explore similar patterns of the relations between EF, numeracy, and mathematics achievement in early childhood.

2. Study 1: methods

2.1. Participants

Participants were 140 three- to five-year-olds at the first assessment (M = 4 years 5 months; range = 3 years 2 months – 5 years 4 months; 49% female, 51% male). Seventy-five percent of parents reported their child's race; of those children, 60% were identified as African American or Black, 19% as Asian or Pacific Islander, 9% as Caucasian or White, 1% as Native American or Alaskan, and 11% as Biracial or Multiracial. Eighty-two percent of parents reported their child's ethnicity; of those children, 21% were identified as Hispanic or Latino and 79% were identified as non-Hispanic/Latino. An additional four children were recruited but excluded from the study because the child: was extremely distracted and unable to complete the assessments (n = 1), had little to no English comprehension and production capacity (n = 2), or repeatedly declined to participate (n = 1).

Children were recruited from four local Head Start centers in a mid-Atlantic state as a part of a larger study on EF and mathematical development (Scalise & Ramani, 2021). Head Start is a federally funded early childhood education program for children living in households at or below the federal poverty line—an annual household income of $25,100 or less for a family of four in 2018. All parents of three- to five- year-olds attending one of the participating Head Start centers received a letter describing the study, along with a consent form and survey. All materials were available in English and Spanish. Children received a sticker after each assessment session, and the lead teacher in each classroom received a $25 gift card for classroom materials.

2.2. Procedure

This study was approved by the Institutional Review Board of the University of Maryland, College Park. Written informed consent was first collected from the parent or guardian of each child, and child verbal assent was collected prior to every session. Children completed three 15- to 20-minute sessions of tasks individually with an experimenter in a quiet area of their classroom or school.

Time 1 of the study (November 2018 – January 2019) included two sessions assessing children's set counting, numeral identification, number comparison, number line estimation, mathematics achievement, and EF skills. Time 2 (March – May 2019) included one session, assessing the four numeracy skills and mathematics achievement. On average, the gap between two time points was 3.7 months (SD = 0.23; range = 3.5 to 5.4).

2.3. Measures

2.3.1. Executive function skills. Three EF tasks were administered targeting inhibition, working memory, and cognitive flexibility, respectively (Diamond, 2013). EF tasks were administered on a tablet computer using the National Institutes of Health Toolbox Cognition Battery (Weintraub et al., 2013). Children's performance on each task was scored using the assessment protocols for uncorrected standardized scores.

2.3.1.1. Flanker. Children saw displays with five fish and were asked to touch a button indicating the direction that the middle fish was facing, which was either in the same direction as surrounding fish (i.e., congruent trials) or the opposite direction (i.e., incongruent trials; adapted from Eriksen & Eriksen, 1974; Rueda et al., 2004). Children completed four practice trials with accuracy feedback and 20 test trials. For each accurate test trial, children received 0.125 points, which were summed over all test trials for a total accuracy score (range: 0 to 2.5 points). Children with high accuracy on the initial test trials proceeded to 20 additional trials involving the direction of arrows instead of fish; on these trials, children received both an accuracy score (range: 0 to 2.5 points) and a reaction time score based on their median reaction time on the incongruent trials that they answered correctly (range: 0 to 5 points). The performance measure was children's uncorrected standardized scores, calculated as the sum of the accuracy and reaction time scores (range: 0 to 10 points) and converted into a normative score. Test-retest reliability for this task is 0.91 (Bauer & Zelazo, 2013).

2.3.1.2. List sorting. Children saw a series of stimuli (e.g., pictures of animals) while listening to verbal labels (e.g., dog, horse), then repeated the verbal labels back to the experimenter from smallest to largest (Tulsky et al., 2013). Children had to remember the animals *and* update the order of the animals based on size, thus the task tapped both storage and manipulation of working memory. Trials began with a list of two stimuli; if a child correctly ordered the stimuli by size, the next trial increased the list length by one. If a child incorrectly ordered the stimuli, the next trial repeated the same list length. The task ended if a child incorrectly responded to two trials of the same list length. The performance measure was the uncorrected standardized score, calculated as the normalized sum of scores across all lists presented. Test-retest reliability for this task is 0.86 (Bauer & Zelazo, 2013).

2.3.1.3. Dimensional Change Card Sort. Children were asked to sort pictures of objects by either color or shape (Zelazo, 2006). Children received four practice trials to sort by color and four practice trials to sort by shape. If a child responded correctly to three or more practice trials, they completed additional test trials where they sorted pictures by color or by shape. For each accurate trial, children received 0.125 points, which was summed over all test trials for a total accuracy score (range: 0 to 5 points). Children with high accuracy proceeded to more challenging test trials and also received a reaction time score based on their median reaction time on trials of the less frequently cued dimension that they answered correctly (range: 0 to 5 points). The performance measure was children's uncorrected standardized scores, calculated as the sum of the accuracy and reaction time scores that are then converted into a normative score. Test-retest reliability for this task is 0.92 (Bauer & Zelazo, 2013).

2.3.2. Numeracy skills

2.3.2.1. Set counting. Children were shown pictures of 2, 3, 4, 5, and 6 stars in random order and asked to count the stars (adapted from Wynn, 1992). The performance measure was the percentage of trials in which the child correctly counted the number of stars. The reliability of the items was KR-20 = 0.81 at Time 1 and KR-20 = 0.71 at Time 2.

2.3.2.2. Numeral identification. Children were shown cards with the numerals 1 to 10 presented in random order and asked what number was on the card (Ramani & Siegler, 2008). The performance measure was the percentage of trials in which the child correctly identified the numeral. The reliability of the items was KR-20 = 0.93 at both Time 1 and Time 2.

2.3.2.3. Number comparison. Children were asked to compare pairs of symbolic numerals ranging from 1 to 9 presented on a paper flipbook (Ramani & Siegler, 2008). After two practice trials with accuracy feedback, children were shown 22 test pairs of numbers and asked to indicate which number is larger. Each number was counterbalanced for side of presentation (i.e., 3|8, 8|3), with the ratio between pairs ranging from 1.1 (e.g., 8|9) to 9.0 (e.g., 9|1). The performance measure was the percentage of correct comparisons. The reliability of the items was KR-20 = 0.81 and 0.85 at Times 1 and 2, respectively.

2.3.2.4. Number line estimation. Children were shown 20 cm lines on a tablet computer, with a 0 labeled at the left end and 10 labeled at the right end, and asked to make a mark on the line where a target number would go (Ramani & Siegler, 2008). After practice making marks on an example trial without a target number, children received 18 test trials with numbers ranging from 1 to 9 presented in random order. On each trial, the experimenter identified the number at the top then asked, "If this is where 0 goes (pointing) and this is where 10 goes (pointing), where does N go?" The performance measure was the percentage of absolute error (PAE = (|estimate - target| / scale of estimates) × 100%). Lower PAE indicated higher accuracy; therefore, PAE were reversed prior to analyses to aid in interpretability (e.g., 100% − PAE). Test-retest reliability in the current sample was $r = 0.29$, $p < .001$. We note that the reliability was suboptimal, possibly due to the challenging nature of this task for preschoolers (Berteletti et al., 2010). We discuss the implications of this finding in the Discussion section.

2.3.3. Mathematics achievement. Children's mathematics achievement was assessed with two addition tasks. Basic arithmetic skills are a hallmark of early elementary school mathematics instruction (Ginsburg, Lee, & Boyd, 2008), and addition tasks supported with nonverbal representations are consistently related to other measures of early mathematics achievement (Chu et al., 2016; Jordan et al., 2009). Thus, although these tasks focus on arithmetic, we used them as proxy measures of mathematics achievement, as in prior research (e.g., Gashaj et al., 2019b).

2.3.3.1. Forced-choice addition. Children were asked which of two imaginary children answered arithmetic problems correctly (adapted from Daubert, 2018; Prather & Alibali, 2011). Children were shown two equations (e.g. 2 + 1 = 3, 2 + 1 = 1), one of which was answered correctly. All equations were shown in the form of symbolic numerals and their corresponding non-symbolic quantities (i.e., sets of cookies). The experimenter read both equations aloud as they pointed to the equation they were reading. The participating children were then asked to indicate which imaginary child was right. Children completed seven items, and the performance measure was the percentage of items children answered correctly. The reliability of the items was KR-20 = 0.37 and 0.46 at Times 1 and 2, respectively. The relatively low reliability might be due to the limited items in this task.

2.3.3.2. Story problem addition. Children were asked to solve addition problems embedded in a story context. The experimenter described a child who has a starting number of tokens (e.g., 3) and gets some number of additional tokens (e.g., 2), then asked the participant how many tokens the child had altogether. On each trial, the experimenter set out the number of tokens referenced for the participant to use if they chose. Each child completed 10 trials, and the performance measure was the percentage of trials that the child answered correctly. The reliability of the items was KR-20 = 0.86 and 0.85 at Times 1 and 2, respectively.

2.4. Analytic approach

The analytic plan was pre-registered on the Open Science Framework (https://osf.io/6r84y). First, we conducted descriptive and correlation analyses to examine the distributions of the data and the associations between the measures. A subset of the 140 participants had missing data on one or more measures of EF (4 children), numeracy (9 children), and mathematics achievement (9 children); however, less than 3% of scores on individual tasks were missing and only 11 children were missing scores on at least one task across the two time points, representing less than 8% of the overall sample. Because the missingness was relatively small, we used mean imputation for missing data instead of excluding these participants.

Next, because EF and mathematics achievement were measured with multiple tasks, we created composite scores for the respective domains of skills to reduce the data for our analyses. We created the EF composite by scaling the three individual task scores to a common *Z* score, then averaging the *Z* scores to form the EF composite. Similarly, we created the mathematics achievement composite by averaging the *Z* scores of the forced-choice and story problem tasks.

To examine the concurrent relation between EF and numeracy skills (RQ1), we conducted OLS regressions on the data obtained at Time 1. In these models, we predicted children's performance on the four numeracy tasks with their concurrent EF composite score while controlling for age and gender. To examine effects of EF on numeracy skills over time (RQ2), we repeated the OLS regressions with the Time 2 numeracy task scores as the dependent variables. The original pre-registration included additional analyses for RQ2 in which we explored the effects of EF on numeracy skills among children with high versus low initial numeracy skills. However, the sample-based median split decreased the sample size for detecting significant effects and the scores within each subgroup were restricted to the higher versus lower range, raising concerns on the reliability of the results. Due to these concerns, we omitted these analyses in the results section. Instead, for RQ2, we added Time 1 numeracy task scores in the regression models to test the unique effect of early EF skills on later numeracy skills.

Finally, we conducted a series of mediation analyses to test the potential mediating effects of the four numeracy skills (i.e., set counting, numeral identification, number comparison, number line estimation) on the longitudinal association between EF and mathematics achievement (RQ3). Because we were interested in the effect of the individual numeracy skill on this association, we first conducted four mediation models to separately test the potential mediating effects of the four numeracy skills. In each model, the Time 1 EF composite was the exogenous variable, Time 2 mathematics achievement composite was the endogenous variable, age and gender were the covariates of mathematics achievement composite, and the Time 1 numeracy task score was the mediator. These four separate models allowed us to examine the individual effect of each numeracy skill on the association between EF and mathematics achievement. Next, we tested significant mediators simultaneously in one model to examine their unique effects on the association between EF and mathematics achievement. Because children's numeracy skills might be correlated with each other, we also estimated the covariance between the mediators. We reported the unstandardized coefficients in the mediation models as it is the recommended practice (Hayes, 2017; Kim & Mueller, 1976).

All primary analyses were conducted in (R Core Team, 2020). We used the *lme4* package (Bates et al., 2015) for the regression analysis, and the *lavaan* package (Rosseel, 2012) for the mediation analysis. Lavaan uses the Delta method to compute parameters. Because the regression and mediation analyses were repeated for four numeracy skills, we used Bonferroni-corrected *p*-values when

Table 1
Mean and standard deviation of scores for each task at Time 1 and Time 2 in Study 1.

Tasks	Time 1 *Mean* (*SD*)	Time 2 *Mean* (*SD*)
Executive function skills		
Flanker task	37.12 (17.39)	–
List sorting task	22.79 (25.95)	–
Dimensional Change Card Sort	40.74 (18.88)	–
Early numeracy skills		
Set counting (%)	81.46 (28.65)	86.32 (22.42)
Numeral identification (%)	66.23 (36.11)	75.49 (32.75)
Number comparison (%)	67.03 (20.74)	73.43 (20.78)
Number line estimation (%)	70.25 (11.24)	71.24 (10.89)
Mathematics achievement		
Forced-choice addition (%)	67.60 (21.23)	72.28 (21.18)
Story problem addition (%)	51.62 (30.87)	57.20 (29.44)

Note: Number line estimation is reported as reversed percent absolute error (see Measures above).

reporting significant effects and included 95% confidence intervals when interpreting the strength of the effects.

3. Study 1: results and discussion

3.1. Preliminary analysis

Descriptive statistics for each measure are shown in Table 1. The means and standard deviations suggest that the scores were widely distributed and none of the tasks were subject to ceiling or floor effects.

Table 2 presents Pearson's bivariate correlations between each measure within each time point. Most correlations were significant, $0.17 < rs < 0.63$, $ps < 0.05$, except for the correlations between the scores on the list sort working memory and number line estimation tasks at Time 1, and the correlations between the forced-choice addition task with the list sort working memory, Dimensional Change Card Sort, set counting, numeral identification, and number line estimation tasks. We noted that the correlation between the two measures of mathematics achievement was weak, likely due to the differences in the task demands. Whereas the forced-choice task involved symbolic equations, the story problem task was more informal (e.g., asking how many tokens there were altogether instead of asking what X + Y equals). Given that the two tasks measure distinct yet related aspects of children's arithmetic skills and together may provide a more complete measure of their mathematics achievement (Scalise & Ramani, 2021), we created composites using the two task scores for the mediation analyses. However, due to concerns with the low reliability of the forced-choice addition task and the low correlation between the two addition tasks, we also repeated the mediation analyses using only the standardized score on the story problem task to test the consistency of the findings. In summary, the descriptive statistics and correlations suggested that children's performance varied across the sample, as did the relations between aspects of children's EF and numeracy skills.

3.2. EF skills predict concurrent numeracy skills

OLS regression analyses revealed that EF skills significantly predicted all four concurrent numeracy skills above and beyond children's age and gender. Specifically, a one standard deviation increase in the EF score was associated with a 0.46 standard deviation ($SE = 0.11$, 95%CI = [0.23,0.68]) increase in accuracy on set counting, 0.54 standard deviation ($SE = 0.11$, 95%CI = [0.32,0.76]) increase in accuracy on numeral identification, 0.60 standard deviation ($SE = 0.10$, 95%CI = [0.41,0.80]) increase in accuracy on number comparison, and 0.38 standard deviation ($SE = 0.10$, 95%CI = [0.16,0.63]) improvement on number line estimation (Table 3). The effect of EF on all four numeracy skills was significant after Bonferroni correction ($p < 0.0125$). Overall, the analyses suggested that children with higher EF skills performed better on these numeracy tasks.

3.3. EF skills predict later numeracy skills

OLS regression analyses revealed that EF skills significantly predicted later performance on all four numeracy tasks above and beyond children's age and gender. Specifically, a one standard deviation increase in the EF score was associated with a 0.40 standard deviation ($SE = 0.11$, 95%CI = [.17,.63]) increase in set counting, 0.39 standard deviation ($SE = 0.11$, 95%CI = [.17,.61]) increase in numeral identification, 0.53 standard deviation ($SE = 0.10$, 95%CI = [.33,.73]) increase in number comparison, and 0.57 standard deviation ($SE = 0.11$, 95%CI = [.14,.63]) improvement on number line estimation (Table 3). The effect of EF on the four numeracy skills remained significant after Bonferroni correction ($p < .0125$). In summary, EF skills remained as a significant predictor of children's later numeracy skills.

Next, we added the respective Time 1 numeracy skill in each regression model to examine the unique influence of EF on later numeracy skills. The models revealed that children's Time 1 numeracy skills significantly predicted their Time 2 numeracy skills for set counting, numeral identification, and number comparison, $ps < 0.001$. When controlling for children's Time 1 numeracy skills and applying the Bonferroni correction, children's age, gender, and EF skills at Time 1 no longer predicted their later numeracy skills in set counting, numeral identification, and number comparison. However, EF skills remained a significant predictor of children's later number line estimation skill ($\beta = 0.52$, $SE = 0.11$, 95%CI = [.29,.74]). In sum, children's early EF skills positively predicted their later

Table 2
Pearson correlations between all task scores at Time 1 (upper triangle) and Time 2 (lower triangle) in Study 1.

	Flanker	List Sort	DCCS	SC	NID	NC	NLE	FC	SP
Flanker	–	0.36***	0.50***	0.39***	0.47***	0.44***	0.28***	0.19*	0.52***
List Sort	0.36***	–	0.38***	0.35***	0.42***	0.52***	0.16	0.19*	0.52***
DCCS	0.50***	0.38***	–	0.32***	0.34***	0.47***	0.31***	0.02	0.44***
SC	0.30***	0.30***	0.35***	–	0.61***	0.46***	0.24**	0.08	0.53***
NID	0.36***	0.39***	0.31***	0.47***	–	0.50***	0.28***	0.09	0.58***
NC	0.47***	0.50***	0.38***	0.46***	0.59***	–	0.41***	0.23**	0.58***
NLE	0.43***	0.49***	0.27**	0.23**	0.43***	0.54***	–	0.11	0.28***
FC	0.20*	0.14	0.13	−0.02	0.05	0.20*	0.17*	–	0.21*
SP	0.49***	0.41***	0.37***	0.46***	0.61***	0.63***	0.48***	0.21*	–

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Note: Flanker, List Sort, and DCCS were measured at Time 1 only, thus the values in the lower triangle represented the correlations between these EF skills at Time 1 and the other measures at Time 2.

Abbreviations: DCCS = Dimensional Change Card Sort; SC = set counting; NID = numeral identification; NC = number comparison; NLE = number line estimation; FC = forced-choice addition; SP = story problem addition

Table 3
Regression models predicting children's numeracy skills at Time 1 and Time 2 in Study 1 ($N = 140$).

	Time 1 Numeracy Skills				Time 2 Numeracy Skills			
Outcome	SC	NID	NC	NLE	SC	NID	NC	NLE
Time 1 Predictors	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]
Age	.17 (.09)	.19 (.08)*	.27 (.08)***	.05 (.10)	.14 (.09)	.27 (.09)**	.29 (.08)***	.13 (.09)
	[−.00,.35]	[.02,.36]	[.12,.42]	[−.14,.24]	[−.04,.31]	[.10,.44]	[.13,.44]	[−.14,.24]
Gender	−.20 (.15)	−.05 (.14)	−.08 (.13)	.06 (.16)	−.43 (.15) **	−.15 (.15)	−.16 (.13)	.03 (.15)
	[−.50,.09]	[−.33,.24]	[−.34,.18]	[−.27,.38]	[−.73,.13]	[−.45,.14]	[−.43,.11]	[−.27,.38]
EF	.46 (.11) ***	.54 (.11)***	.60 (.10)***	.38 (.12)**	.40 (.12)***	.39 (.11)***	.53 (.10)***	.57 (.11)***
	[.23,.68]	[.32,.76]	[.41,.80]	[.16,.63]	[.17,.63]	[.17,.61]	[.33,.73]	[.14,.63]
R^2	.24	.30	.43	.10	.22	.26	.39	.27
Age					.06 (.08)	.11 (.05)*	.16 (.07)*	.12 (.09)
					[−.10,.22]	[.01,.21]	[.02,.31]	[−.05,.29]
Gender					−.33 (.14)*	−.11 (.08)	−.12 (.12)	.02 (.15)
					[−.60, −.06]	[−.28,.05]	[−.36,.11]	[−.26,.31]
EF					.19 (.11)	−.07 (.07)	.26 (.10)*	.52 (.11)***
					[−.03,.40]	[−.21,.06]	[.05,.46]	[.29,.74]
Numeracy					.47 (.08)***	.85 (.05)***	.46 (.08)***	.14 (.07)
					[.32,.62]	[.76,.95]	[.31,.62]	[−.01,.29]
R^2					.39	.77	.51	.29

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. Gender: Male = 1, Female = 0. Abbreviations: SC = set counting; NID = numeral identification; NC = number comparison; NLE = number line estimation; Numeracy = Time 1 numeracy task score corresponding to the Time 2 outcome.

set counting, number identification, number comparison, and number line estimation skills. However, only the effect of EF on number line estimation skill remained significant when controlling for the respective early numeracy skills.

3.4. Numeracy skills mediate the association between EF and mathematics achievement

The first step of the mediation analysis revealed that EF skills predicted later mathematics achievement above and beyond age and gender (Fig. 1a). Next, four separate mediation analyses revealed that children's skills in set counting (Fig. 1b), numeral identification (Fig. 1c), and number comparison (Fig. 1d) partially mediated the association between EF and mathematics achievement; children's number line estimation skill did not mediate the association between EF and mathematics achievement (Fig. 1e). Finally, we tested the

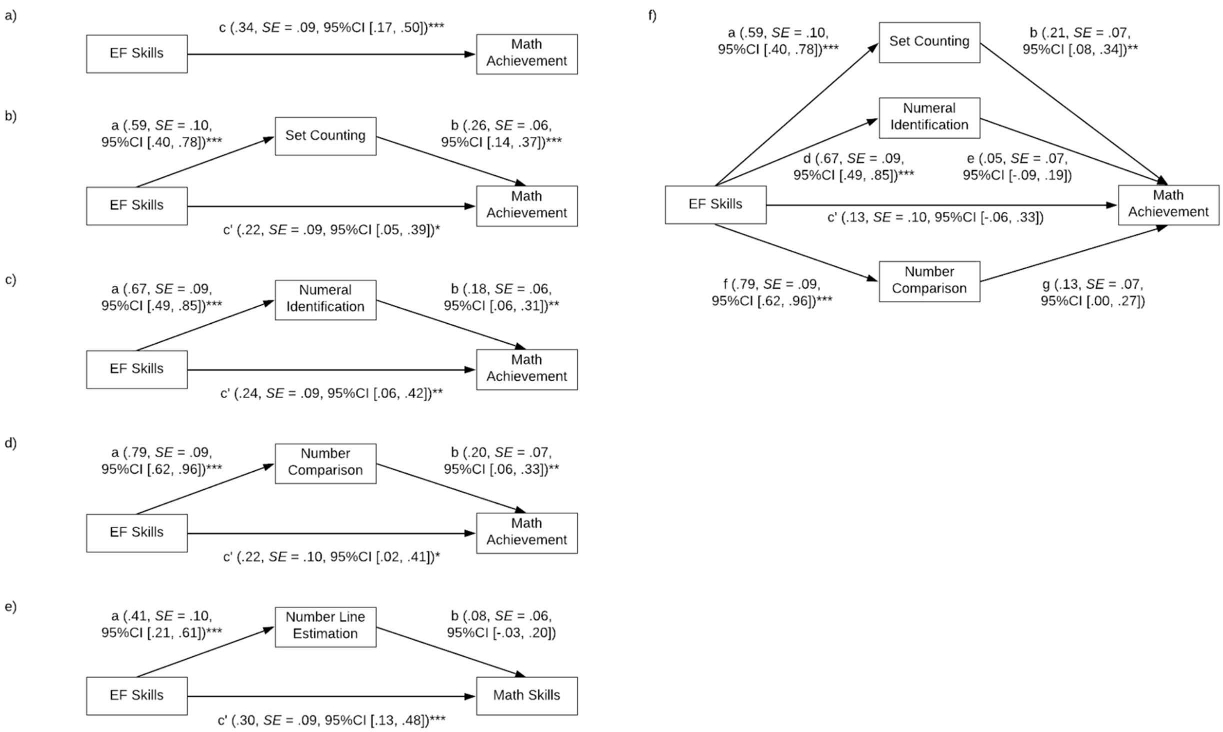

Fig. 1. Mediation analyses between EF and later mathematics achievement through early numeracy skills in Study 1. All models include age and gender as covariates of mathematics achievement. The path values represent non-standardized estimates. *Note.* *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

three significant mediators simultaneously (Fig. 1f), and found that only set counting skill mediated the association between EF and mathematics achievement. Informed by poor reliability of the forced-choice addition task and the low correlation between the two measures of children's mathematics achievement, we repeated the mediation analyses using only the scores on the story problem addition task. Because the results were replicated whether we used the composite score or the story problem addition score, we reported the findings using the composite scores above.

In summary, we found that children's set counting, numeral identification, and number comparison skills each partially mediated the association between EF and mathematics achievement. However, when the three skills were tested simultaneously, only the set counting skill mediated the association between EF and mathematics achievement, whereas the pathways from numeral identification and number comparison skills to children's later mathematics achievement were not significant. The findings suggest that children's set counting skill may be a potential pathway through which EF influences mathematics achievement among three- to five-year-olds.

4. Study 2: methods

4.1. Participants

A total of 109 kindergartners were recruited from four public schools in a midwestern state in the United States. The percentage of students identified as English Language Learners ranged between 9% and 57% at each school; the percentage of students eligible for free and reduced lunch ranged between 22% and 89%. Five children did not complete the study due to family relocation ($n = 1$) or their request to end the sessions ($n = 4$), thus the analyses excluded these five children and the final sample included 104 children (56% female, 44% male).

The participating children were five or six years of age ($M = 5$ years 11 months, $SD = 4$ months); 36% were identified as White, 25% as Black, 15% as Asian or Asian American, 2% as American Indian or Alaskan Native, and 22% as other races. In terms of ethnicity, 21% were identified as Spanish, Hispanic, or Latino, and 79% were not. We did not provide monetary compensation for children who participated in the study. We gave the lead teacher in each classroom a $50 gift card for classroom materials.

4.2. Procedure

This study was approved by the Institutional Review Board of the University of Minnesota, Twin Cities. Written informed consent was collected from the parent or guardian of each child. All measures were collected in the context of a larger training study in which children received four 15-minute training sessions in one of three conditions that focused on (1) number only, (2) number combined with EF and relational language (e.g., 5 is *more* than 4 and 5 comes *after* 4), or (3) alphabetic (non-numerical) ordering (Chan, 2019; Chan, Sera, & Mazzocco, in press). The goal of the training conditions was to test whether the combination training was more effective at improving children's numeracy skills compared to the number only training. To examine the training effects and inform the primary analyses, we conducted a 2 (Time: pretest vs. posttest) × 3 (Condition) repeated measures MANOVA on the four numeracy skills. The MANOVA revealed a main effect of time, $F(4,98) = 18.64$, $p < 0.001$, $\eta_p^2 = 0.432$; the effects of condition and Time × Condition interaction were not significant, $ps > 0.10$. The follow up ANOVAs revealed that children, regardless of the condition, improved on numeral identification, $F(1,101) = 19.83$, $p < 0.001$, $\eta_p^2 = 0.164$, and number line estimation skills, $F(1,101) = 58.90$, $p < 0.001$, $\eta_p^2 = 0.368$. Neither of these skills were practiced in the alphabetic condition, suggesting that the comparable gains across the three conditions might be independent of the training activities. Thus, in the current study, we collapsed across the three training conditions and focused on children's performance at pretest and posttest only (hereafter Time 1 and Time 2, respectively).

We measured children's EF, numeracy, and mathematics achievement at Time 1, and EF and numeracy skills at Time 2, approximately six weeks later. Children were tested individually in the child's school, in a relatively quiet area away from others. Most children completed the Time 1 measures in two sessions, and the Time 2 measures in one session. Some children participated in an additional session due to breaks they requested or time constraints set by teachers.

4.3. Measures

4.3.1. Executive function skills.

Head Toes Knees Shoulders and Minnesota Executive Function Scale were used as two composite measures of EF skills.

4.3.1.1. Head Toes Knees Shoulders (HTKS). In this standardized task, children were to do something different from what the experimenter asked (Cameron Ponitz et al., 2008). For instance, children were to touch their shoulders when the experimenter said, "touch your knees." A total of 47 trials were administered (17 practice and 30 test trials), and children received 0 (*incorrect*), 1 (*self-correct*), or 2 (*correct*) points per trial, with a maximum of 94 points. Following the task developer's recommendation, the practice trials were included in the score to increase the test's range and variability. The internal consistency reliability of the task ranges from 0.87 to 0.92 (Cameron Ponitz et al., 2008).

4.3.1.2. Minnesota Executive Function Scale (MEFS). This is an adaptive iPad version of Dimensional Change Card Sort (Carlson & Zelazo, 2014; Zelazo, 2006). Similar to the traditional Dimensional Change Card Sort, children sorted virtual cards following the stated rules. Extending from the traditional task, there were seven levels of difficulty and all children in the current study started at Level 4, the entry level for five-year-olds. Children progressed forward to the next level or backward to an easier level depending on their performance. The total scores were computed based on highest level passed, highest level attempted, errors, and reaction time, and an age-referenced standard score was used as the indicator of children's performance. The test-retest reliability for MEFS is $ICC = 0.93$ (Beck, Schaefer, Pang, & Carlson, 2011).

4.3.2. Numeracy skills. The four numeracy tasks were similar to those in Study 1, except that the range of the numbers tested was larger.

4.3.2.1. Set counting. The experimenter showed children a page with a set of black dots and asked them to count the dots carefully with their fingers and to tell the experimenter how many there were. The set size ranged from 8 to 17. A total of four trials were presented, and the performance measure was the percentage of trials in which the child correctly counted the number of dots. The reliability of the items was KR-20 = 0.42 and 0.46 at Times 1 and 2, respectively. The relatively low reliability might be due to the limited number of items.

4.3.2.2. Numeral identification. The experimenter presented printed numerals in a random order, and asked children to name the number. The numbers ranged from 1 to 99. A total of 20 numerals were tested (1 to 9 and 11 numbers between 10 and 99). The performance measure was the percentage of trials in which the child correctly identified the numeral. The reliability of the items was KR-20 = 0.91 and 0.90 at Times 1 and 2, respectively.

4.3.2.3. Number comparison. The experimenter asked children, "which number means more", then stated two numbers between 1 and 99. All 16 pairs of numbers differed by two (e.g., 3 vs. 5), and the performance measure was the percentage of correct response. The reliability of the items was KR-20 = 0.74 and 0.77 at Times 1 and 2, respectively.

4.3.2.4. Number line estimation. The experimenter presented a number line with a target number on a letter-size paper in landscape orientation, then asked children where the target number goes on the number line. The length of the line was 25.40 cm, and two points were marked at 6.35 cm from each end of the line, resulting in a 12.70 cm span between the two points. Children received twelve 0–20 number line trials with targets between 3 and 19, and fourteen 0–100 number line trials with targets between 7 and 91, for a total of 26 trials. On half of each set of trials, the correct midpoint (i.e., either 10 or 50) was marked and labeled to support estimation and to maximize variation in children's performance. As in Study 1, we calculated PAE for each trial. Children's PAE across these four types of trials (i.e., 0–20 number line and 0–100 number line with and without the midpoint) were significantly correlated with each other (Time 1: $0.41 < |r_S s| < 0.74$; Time 2: $0.54 < |r_S s| < 0.76$, $ps < 0.001$), thus we used the average PAE as an indicator of children's performance on this task. As in Study 1, PAE scores were reversed (e.g., 100% − PAE) prior to analyses. Test-retest reliability in the current sample was $r = 0.71$, $p < 0.001$.

4.3.3. Mathematics achievement. We administered the Test of Early Mathematics Ability–3rd Edition (TEMA-3; Ginsburg & Baroody, 2003) as a measure of children's general mathematics achievement. The TEMA-3 includes items such as verbal counting, story problems, and arithmetic fact retrieval, and is a comprehensive assessment of children's mathematical skills. Similar types of items were grouped for the ease of testing but the beginning and the end of the test followed the standard procedure. Form A of the TEMA-3 was administered. The standard score based on an age-referenced mean of 100 and a standard deviation of 15 was used in the analyses. The internal-consistency reliability of the TEMA-3 is 0.94 (Ginsburg & Baroody, 2003).

We noted that five TEMA-3 items overlapped with tasks that assessed set counting (two items) and numeral identification skills (three items). However, our assessments of numeracy skills provided more granular information on each numeracy skill (e.g., number of correct responses rather than pass or fail on each item). Further, TEMA-3 included additional items that tapped other skills, forming a single measure of general mathematical skills.

4.4. Analytic approach

We conducted parallel analyses outlined in Study 1 to address RQ1 and RQ2. For RQ3, because mathematics achievement was only measured at Time 1, we conducted mediation analyses using concurrent data at Time 1. Because we could not infer the directionality of the effects based on concurrent data, we also conducted reverse mediations in which EF skills served as a mediator between numeracy and mathematics achievement to examine the plausible directions of the influences. One child did not complete the number line estimation task at Time 2, thus we used mean imputation for the missing data. We created composites for EF skills by computing the average Z scores of the two EF tasks.

Table 4
Mean and standard deviation of scores for each task at Time 1 and Time 2 in Study 2.

Tasks	Time 1 *Mean* (*SD*)	Time 2 *Mean* (*SD*)
Executive function skills		
HTKS total score	60.59 (26.13)	68.00 (20.93)
MEFS standard score	100.40 (9.88)	104.77 (12.50)
Early numeracy skills		
Set counting (%)	71.39 (24.74)	69.47 (23.35)
Numeral identification (%)	85.19 (20.32)	89.38 (17.09)
Number comparison (%)	76.68 (18.06)	78.79 (18.57)
Number line estimation (%)	84.56 (7.07)	88.46 (6.16)
Mathematics achievement		
TEMA standard score	99.59 (13.69)	–

Note: Number line estimation is reported as reversed percent absolute error (see Measures above).
Abbreviations: HTKS = Head Toes Knees Shoulders task; MEFS = Minnesota Executive Function Scale; TEMA = Test of Early Mathematics Ability.

5. Study 2: results and discussion

5.1. Preliminary analysis

The descriptive analyses revealed that none of the tasks was subject to ceiling or floor effects, as reflected by the mean and standard deviation of scores (Table 4). Spearman correlation analysis revealed that most bivariate correlations were significant, $0.22 < r_S s < 0.75$, $ps < 0.05$, except for the correlations between the scores on set counting and numeral identification at Time 2, and set counting at Time 2 and TEMA at Time 1 (Table 5). These analyses provided evidence for a wide range of performance levels in the sample, and relations between EF, numeracy, and mathematics achievement.

5.2. EF skills predict concurrent numeracy skills

The four OLS regression models revealed that EF skills significantly predicted all four numeracy skills above and beyond children's age and gender. Specifically, a one standard deviation increase in EF skills was associated with 0.34 standard deviation ($SE = 0.12$; 95%CI = [0.10,0.58]) increase in set counting, 0.53 standard deviation ($SE = 0.11$; 95%CI = [0.31,0.75]) increase in numeral identification, 0.61 standard deviation ($SE = 0.10$; 95%CI = [0.41,0.81]) increase in number comparison, and 0.55 standard deviation ($SE = 0.11$; 95%CI = [0.34,0.77]) improvement on number line estimation (Table 6). The effect of EF on the four numeracy skills was significant after Bonferroni correction ($p < 0.0125$). Overall, the analyses suggested that children with higher EF skills performed better on these four numeracy tasks.

5.3. EF skills predict later numeracy skills

OLS regression analyses revealed that when controlling for children's age and gender, children's EF skills significantly predicted their later numeracy skills except for set counting. A one standard deviation increase in EF skills was associated with 0.50 standard deviation ($SE = 0.11$, 95%CI = [0.27,0.72]) increase in numeral identification, 0.49 standard deviation ($SE = 0.11$, 95%CI = [0.28,0.70]) increase in number comparison, and 0.46 standard deviation ($SE = 0.11$, 95%CI = [0.23,0.68]) improvement on number line estimation skills (Table 6). After Bonferroni correction ($p < 0.0125$), the effect of EF on the three numeracy skills remained significant.

Next, we added the respective Time 1 numeracy skill in each regression model to examine the unique influence of EF on later numeracy skills. The models revealed that children's Time 1 numeracy skills significantly predicted their Time 2 numeracy skills, $ps < 0.010$. When accounting for children's Time 1 numeracy skills and applying the Bonferroni correction, children's age, gender, and EF skills at Time 1 no longer predicted their later numeracy skills. In sum, children's early EF skills positively predicted their later numeral identification, number comparison, and number line estimation skills; however, the effect of EF on these numeracy skills was no longer significant when controlling for the respective early numeracy skills.

5.4. Numeracy skills mediate the association between EF and mathematics achievement

The first step of the mediation analysis revealed that EF skills predicted concurrent mathematics achievement above and beyond age and gender (Fig. 2a). Next, we found that skills in set counting (Fig. 2b), numeral identification (Fig. 2c), number comparison (Fig. 2d), and number line estimation (Fig. 2e) each partially mediated the concurrent association between EF and mathematics achievement. Finally, we tested the four mediators simultaneously (Fig. 2f), and found that all four numeracy skills significantly mediated the association between EF and mathematics achievement, and the path between EF and mathematics achievement was no longer significant.

Because EF, numeracy, and mathematics achievement were all measured at Time 1, the concurrent measures precluded causal inferences. We conducted follow-up analyses to explore whether the reverse mediation—EF as a pathway between numeracy and mathematics achievement—would also be statistically significant. The first step of the analysis revealed that, after controlling for age and gender, set counting ($B = 0.10$, $SE = 0.03$, 95%CI = [0.04,0.15]), numeral identification ($B = 0.31$, $SE = 0.05$, 95%CI = [0.22,0.40]), number comparison ($B = 0.20$, $SE = 0.05$, 95%CI = [0.10,0.31]), and number line estimation skills ($B = 0.48$, $SE = 0.12$, 95%CI = [0.23,0.73]) significantly predicted children's mathematics achievement, $ps \leq 0.001$. When we added EF skills as a mediator, EF did not mediate the association between the four numeracy skills and mathematics achievement. Specifically, only number

Table 5
Spearman correlations between all task scores at Time 1 (upper triangle) and Time2 (lower triangle) in Study 2.

	HTKS	MEFS	SC	NID	NC	NLE	TEMA
HTKS	–	0.44***	0.28**	0.60***	0.57***	0.55***	0.53***
MEFS	0.52***	–	0.22*	0.28**	0.45***	0.49***	0.38***
SC	0.27**	0.24*	–	0.29**	0.25*	0.31**	0.36***
NID	0.57***	0.41***	0.14	–	0.68***	0.56***	0.75***
NC	0.58***	0.47***	0.22*	0.71***	–	0.66***	0.64***
NLE	0.53***	0.49***	0.38***	0.61***	0.63***	–	0.65***
TEMA	0.56***	0.48***	0.16	0.73***	0.69***	0.61***	–

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. *Note:* TEMA was measured at Time 1 only, thus the values in the lower triangle represented the correlations between TEMA at Time 1 and the other measures at Time 2. Abbreviations: HTKS = Head Toes Knees Shoulders task; MEFS = Minnesota Executive Function Scale; SC = set counting; NID = numeral identification; NC = number comparison; NLE = number line estimation; TEMA = Test of Early Mathematics Ability.

Table 6
Regression models predicting children's numeracy skills at Time 1 and Time 2 in Study 2 ($N = 104$).

	Time 1 Numeracy Skills				Time 2 Numeracy Skills			
Outcome	SC	NID	NC	NLE	SC	NID	NC	NLE
Time 1 Predictors	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]	β(SE) [95%CI]
Age	.05 (.10) [−.15,.25]	.10 (.09) [−.09,.29]	.18 (.09)* [.01,.35]	.06 (.09) [−.12,.24]	.21 (.10)* [.04,.41]	.11 (.09) [−.08,.30]	.27 (.09)** [.10,.45]	.07 (.10) [−.12,.26]
Gender	−.04 (.19) [−.42,.34]	.08 (.18) [−.27,.42]	−.06 (.16) [−.38,.26]	−.38 (.17)* [−.72,−.04]	−.20 (.19) [−.58,.18]	.04 (.18) [−.31,.40]	−.11 (.17) [−.43,.22]	−.40 (.18)* [−.75, −0.05]
EF	.34 (.12)** [.10,.58]	.53 (.11)*** [.31,.75]	.61 (.10)*** [.41,.81]	.55 (.11)*** [.34,.77]	.19 (.12) [−.05,.43]	.50 (.11)*** [.27,.72]	.49 (.11)*** [.28,.70]	.46 (.11)*** [.23,.68]
R^2	.10	.24	.36	.28	.10	.22	.32	.22
Age					.19 (.10) [−.00,.38]	.02 (.05) [−.08,.13]	.16 (.07)* [.02,.30]	.04 (.08) [−.11,.19]
Gender					−.18 (.18) [−.54,.18]	−.02 (.09) [−.21,.17]	−.07 (.13) [−.32,.19]	−.16 (.14) [−.44,.13]
EF					.08 (.12) [−.15,.32]	.05 (.07) [−.09,.18]	.09 (.10) [−.10,.28]	.10 (.10) [−.10,.30]
Numeracy					.32 (.09)** [.13,.51]	.86 (.05)*** [.75,.96]	.65 (.08)*** [.49,.81]	.64 (.08)*** [.48,.81]
R^2					.20	.78	.60	.52

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. Gender: Male = 1, Female = 0
Abbreviations: SC = set counting; NID = numeral identification; NC = number comparison; NLE = number line estimation; Numeracy = Time 1 numeracy task score corresponding to the Time 2 outcome.

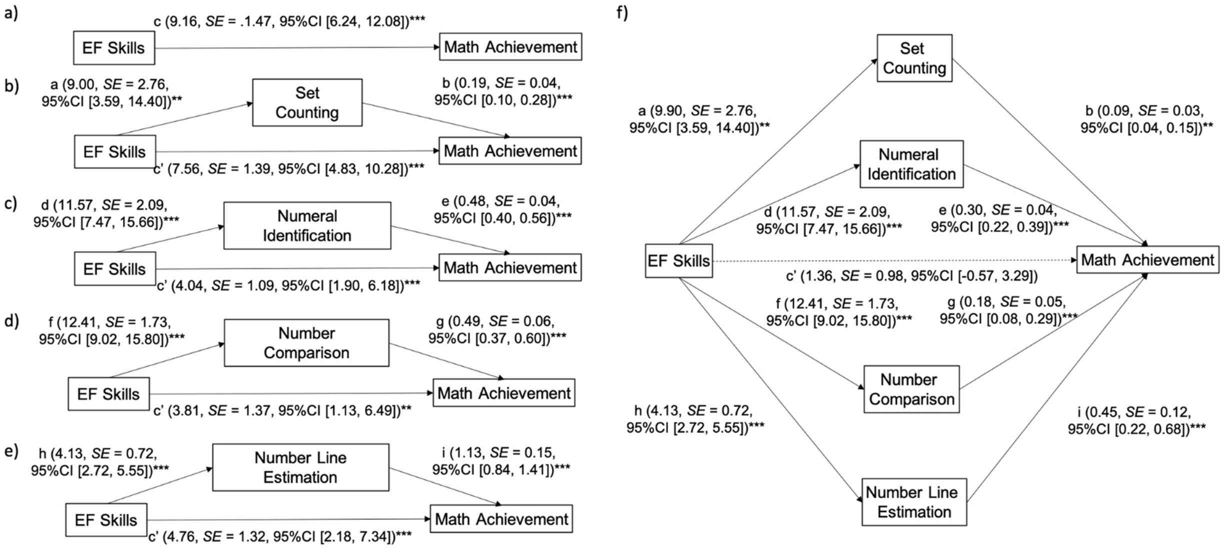

Fig. 2. Mediation analyses between EF and mathematics achievement through early numeracy skills at Time 1 in Study 2. All models include age and gender as covariates of mathematics achievement. The path values represent non-standardized estimates. *Note.* **$p < 0.01$, ***$p < 0.001$.

comparison skill significantly predicted children's EF skills ($B = 0.02$, $SE = 0.01$, 95%CI = [0.01,0.03]), $p = 0.001$; EF skills did not significantly predict children's mathematics achievement, $p = 0.147$.

In summary, when the numeracy skills were examined separately, these skills partially mediated the association between EF and mathematics achievement. When these numeracy skills were tested simultaneously, they all significantly mediated the association between EF and mathematics achievement, and together fully mediated this association. Further, the reverse mediation revealed that EF was not a mediator between numeracy and mathematics achievement. The findings suggest that children's numeracy skills may be potential pathways through which EF influences mathematics achievement among five- to six-year-olds.

6. General discussion

The goal of the current study was to unpack the relations between EF, numeracy, and mathematics achievement in early childhood. Across the two samples of children, the results revealed significant relations between EF and numeracy skills, although the patterns varied between samples and specific types of skills. We discuss these findings in more detail below.

6.1. Concurrent association between EF and numeracy skills

In both studies, we found that the concurrent association between EF and numeracy was significant and positive for the four numeracy skills examined. The findings align with prior work on the association between EF and numeracy skills in preschool and kindergarten children (Friso-van den Bos et al., 2014; Gashaj et al., 2016; Purpura et al., 2017), and provide additional evidence on the robustness of this association across specific numeracy skills in early childhood. Further, the standardized beta coefficients for EF skills range from 0.34 to 0.61 across the two studies; these values appear comparable to the correlations reported in previous studies on the association between EF and set counting ($0.28 < r < 0.60$; Purpura et al., 2017), numeral identification ($0.14 < r < 0.38$; Purpura et al., 2017), number comparison ($r = 0.30$; Gashaj et al., 2019a), and number line estimation skills ($0.27 < r < 0.35$; Gashaj et al., 2016).

Contrary to our original hypothesis, we did not find that EF skills were more strongly related to number line estimation compared to other numeracy skills. In fact, the confidence intervals for the effects of EF on each numeracy skill overlapped with one another in both samples, suggesting comparable strength of the relations between EF and each numeracy skill in preschoolers and kindergartners (Tables 3 and 6, respectively). Although the results differed from our hypothesis, the study still contributed to the existing literature (Gashaj et al., 2016, 2019a; Purpura et al., 2017) by providing findings on the association between EF and number line estimation skills in preschoolers as well as the association between EF and set counting and numeral identification skills in kindergartners.

One interpretation of our current findings is that many children are still developing these four numeracy skills in preschool and kindergarten, and EF skills may play an important role in supporting their overall learning process. Alternatively, the influence of EF skills on early numeracy development may be more nuanced in that aspects of numeracy skills may draw on certain components of EF skills. For instance, while EF skills may support set counting by helping children remember the counting sequence and keep track of counted objects, the ways in which EF skills are involved in numeral identification, number comparison, and number line estimation skills may be different. Purpura and colleagues (2017) have found that inhibition may be more important for counting whereas working memory may be more important for number comparison and arithmetic problem-solving. We used composite scores as an overall indicator of EF skills because prior work suggested that EF skills might emerge (Lee et al., 2013) and contribute to mathematics achievement as a unitary construct in early childhood (Nguyen et al., 2019). Future research should further explore how components of EF skills relate to aspects of numeracy skills.

6.2. Longitudinal association between EF and numeracy skills

In both studies, we found that EF skills predicted aspects of later numeracy skills. Specifically, EF skills remained as a significant and strong predictor of all four numeracy skills three to five months later among preschoolers. Similarly, EF skills remained as a significant and strong predictor of numeral identification, number comparison, and number line estimation skills, but not set counting skill among kindergartners. One potential explanation for this finding may be that set counting is one of the earliest numeracy skills children develop, and its demand on EF skills may decrease as children become better at counting sets of objects. Indeed, other research has shown that the effects of EF skills are weaker in preschoolers with higher, compared to lower, EF and mathematical skills (Dong et al., 2021), suggesting a decreasing trend of the association between EF and mathematical skills as children develop these skills.

When controlling for early numeracy skills, the effect of EF was only significant on later number line estimation skill among preschoolers, and it was not significant on the numeracy skills among kindergartners. The significant effect of EF on number line estimation aligned with a previous report on kindergartners (Gashaj et al., 2016, 2019a), but it was worth noting that the test-retest reliability for number line estimation was relatively low among preschoolers in Study 1. The low reliability of this task among preschoolers echoed the previous finding that 52% of four-year-olds' estimation on 0-10 number lines could not be categorized as logarithmic (i.e., correct order of numbers but smaller spacing between larger numbers) or linear (i.e., correct order and equal spacing between numbers; Berteletti et al., 2010). Instead, some of these children in the previous study alternated between the left and right endpoints when placing the target numbers on number lines, suggesting that they might not yet fully understand the task or appreciate the number line representation. Indeed, we also observed similar patterns of responses on the number line estimation task in our preschool sample, and to some extent in our kindergarten sample. Therefore, the significant longitudinal relation between EF and number line estimation skill combined with the low reliability of the number line estimation task among preschoolers might extend prior research by suggesting the particular importance of EF in the early stages of learning about number line representations. With little knowledge of numbers or number lines, children with higher versus lower EF skills may be better able to improve their numerical understanding. As children acquire some foundational knowledge of numbers, their emerging number knowledge may become a stronger predictor of their later performance, whereas the influence of EF may weaken overtime. This hypothesis warrants further investigation, but the current research provides a starting point to examine the longitudinal association between EF and aspects of numeracy skills across early childhood. With assessments at multiple time points, future studies can further unpack the association between EF, number line estimation, and other early numeracy skills across development.

6.3. Numeracy skills as mediators of EF and mathematics achievement

The final goal of the current study was to examine the mechanisms underlying the relation between EF and mathematics achievement in early childhood. Specifically, we hypothesized that children's numeracy skills would mediate the relation between EF and mathematics achievement. Our results provide evidence for a more nuanced understanding of the relations between EF, numeracy, and mathematics development. In Study 1, we found that preschoolers' set counting, numeral identification, and number comparison skills each partially mediated the association between EF and later mathematics achievement when tested in separate models. Similarly, in Study 2, we found that kindergartners' set counting, numeral identification, number comparison, and number line estimation skills each partially mediated the concurrent association between EF and mathematics achievement. These parallel findings suggest that each of these numeracy skills may be supported by EF skills and in turn contribute to mathematics achievement.

When we tested the numeracy skills simultaneously in a multiple mediation model, we found somewhat different patterns of results between the two studies. In Study 1, only set counting skill, but not numeral identification or number comparison, predicted later mathematics achievement among preschoolers. In Study 2, all four numeracy skills significantly predicted mathematics achievement among kindergartners. There are several interpretations for the divergent findings. One possibility is that set counting may be fundamental for preschoolers' mathematics achievement (Chu et al., 2015), and it independently influences mathematics achievement beyond other numeracy skills. As children develop other aspects of numeracy skills throughout preschool and kindergarten, other numeracy skills may become more important for mathematics achievement.

Alternatively, the relation between numeracy skills and mathematics achievement may reflect the specifics of the measures used in each study. For example, the addition tasks in Study 1 may rely on counting sets to find the summed total whereas TEMA-3 in Study 2 may tap a broad range of mathematical skills. Further, the reliability of the forced-choice addition task and its correlation with the story problem addition task in Study 1 were relatively low, raising concerns for combining the two scores as a composite measure of preschoolers' mathematics achievement. Given the low reliability of the forced-choice addition task, we repeated the mediation analyses with the standardized score on the story problem task as the outcome, and found that the patterns of results replicated those in Study 1 with the composite measure of addition. In both studies, regardless of whether the mathematics achievement measures are operationalized as a more advanced number principle (i.e., addition in Study 1) or a more general set of number skills (i.e., TEMA-3 in Study 2), the results suggest that EF forms a foundation for these aspects of mathematical understanding. Together, the findings from the two studies suggest that EF supports a collection of early numeracy skills, and in turn support children's deeper and broader understanding of more advanced number principles—a core component of children's mathematics achievement. Our current study provides evidence for the importance of supporting early numeracy skills as a collection of independent but related skills, rather than a single skill driving the relation between EF and mathematics achievement.

6.4. Limitations and future directions

The current study has several limitations. First, although the findings provide some support for the independent relations between EF and aspects of numeracy skills as well as numeracy skills being potential mediators between EF and mathematics achievement, the results were based on two short-term studies with only two time points. Future research should consider longer-term studies with multiple time points to better understand the developmental relations between EF, numeracy, and mathematics achievement throughout early childhood. Second, the two studies used different measures of EF and mathematics achievement and some of the measures had suboptimal reliability. In particular, the challenging nature of the number line estimation task might have contributed to its low reliability in Study 1; the relatively small number of items in the forced-choice addition task in Study 1 and set counting task in Study 2 might have also contributed to their low reliability. Further, while Study 1 used two laboratory-based addition tasks as measures of children's mathematics achievement and Study 2 used a standardized mathematics ability test, none of these measures assessed children's mathematics achievement beyond advanced numerical skills. Our measures did not tap other important mathematical skills such as patterning, geometry, and problem-solving (LeFevre et al., 2010; National Research Council, 2009; Sarama & Clements, 2004). These differences and limitations of our measures did not allow direct comparisons of the results from the two studies, but did allow us to identify similar patterns across samples and measures. Future research with extended, reliable, and consistent measures of EF and mathematics achievement that scale according to children's age would allow direct comparisons across subdomains and development. Third, there were SES differences between the two samples—the preschool sample enrolled children from low-SES households, whereas the kindergarten sample included children from a range of SES groups. Given the vast literature highlighting SES differences in children's mathematics achievement (see Jordan & Levine, 2009), it is possible that some differences in our findings reflect differences in children's SES. Finally, future research should consider the possibility of a more complex model of children's developing skills with bidirectional paths between children's EF, numeracy, and mathematics achievement. Indeed, some evidence suggests that young children's EF and mathematical skills have a bidirectional relation over time (Cameron et al., 2019; Clements et al., 2016; Fuhs et al., 2014; Mulcahy et al., 2021; Nesbitt et al., 2019), however, other researchers have found support for a unidirectional relation with EF supporting children's early mathematical learning (Goodrich, et al., 2021; Ribner, 2020).

6.5. Conclusion

The relation between EF and mathematics achievement is well substantiated in theoretical and empirical literature. In the current study, we examined a potential mechanism underlying this relation in early childhood. In two samples of children in preschool and kindergarten, we found evidence that EF skills predicted children's concurrent and later numeracy skills. Further, individual numeracy skills partially mediated the association between EF and mathematics achievement, and together, these skills fully mediated this association. Such findings may inform practices to better support mathematical learning among children with varying levels of EF skills. For instance, researchers and educators may develop practices that strengthen children's numeracy skills to mitigate the effects of having lower EF skills and in turn lower mathematics achievement. In sum, the study provides evidence on the nuanced relations between EF, numeracy, and mathematics achievement, and suggests attention to each numeracy skill in order to support early mathematical development.

Funding

This work was supported by a Doctoral Dissertation Fellowship from the University of Minnesota, a Dissertation Funding Award from the Society for Research in Child Development Student and Early Career Council, and a Herb Pick Applied Developmental

Psychology Seed Grant from the Institute of Child Development to J. Chan; and a National Science Foundation Graduate Research Fellowship [Grant No. DGE 1322106], a National Science Foundation Social, Behavioral, and Economic Sciences Postdoctoral Research Fellowship [Grant No. SBE 1911869], and a University of Maryland College of Education Doctoral Dissertation Student Award to N. Scalise. The above-listed funding sources had no involvement in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

CRediT authorship contribution statement

Jenny Yun-Chen Chan: Conceptualization, Investigation, Analysis, Writing of Study 2. **Nicole R. Scalise:** Conceptualization, Investigation, Analysis, Writing of Study 1. Both authors contributed equally to the original draft. Chan led the revision of the manuscript.

Declaration of Interests

None.

Data Availability Statement

The research data from this study is confidential. Please contact the authors for any additional information or requests for summary statistics.

Acknowledgments

The research was completed as a partial fulfillment of both authors' doctoral dissertation requirements. We would like to thank Michèle Mazzocco, Geetha Ramani, and Maria Sera for their guidance on the dissertation projects, the members of the labs for their support, and the teachers, staff, students, and families for their participation.

References

Aunio, P., Niemivirta, M., Hautamäki, J., Van Luit, J. E. H., Shi, J., & Zhang, M. (2006). Young children's number sense in China and Finland. *Scandinavian Journal of Educational Research, 50*(5), 483–502. https://doi.org/10.1080/00313830600953576

Baroody, A. J., & Price, J. (1983). The development of the number-word sequence. *Journal for Research in Mathematics Education, 14*(5), 361–368.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Fox, J. (2015). Package "lme4.". *Convergence, 12*(1), 2.

Bauer, P. J., & Zelazo, P. D. (2013). NIH toolbox cognition battery (CB): Summary, conclusions, and implications for cognitive development. *Monographs of the Society for Research in Child Development, 78*(4), 133–146. https://doi.org/10.1111/mono.12039

Beck, D. M., Schaefer, C., Pang, K., & Carlson, S. M. (2011). Executive function in preschool children: Test–retest reliability. *Journal of Cognition and Development, 12* (2), 169–193. https://doi.org/10.1080/15248372.2011.563485

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 46*(2), 545–551. https://doi.org/10.1037/a0017887

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from age 5 to 17 in large, representative national sample. *Learning and Individual Differences, 21*(4), 327–336. https://doi.org/10.1016/j.lindif.2011.01.007

Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., & Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: Sources of variation in emergent mathematics achievement. *Developmental Science, 14*(4), 679–692. https://doi.org/10.1111/j.1467-7687.2010.01012.x

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives, 8*(1), 36–41. https://doi.org/10.1111/cdep.12059

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*(3), 273–293. https://doi.org/10.1207/S15326942DN1903

Cameron, C. E., Kim, H., Duncan, R. J., Becker, D. R., & McClelland, M. M. (2019). Bidirectional and co-developing associations of cognitive, mathematics, and literacy skills during kindergarten. *Journal of Applied Developmental Psychology, 62*, 135–144. https://doi.org/10.1016/j.appdev.2019.02.004

Cameron Ponitz, C. E., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly, 23*(2), 141–158. https://doi.org/10.1016/j.ecresq.2007.01.004

Carlson, S.M., & Zelazo, P.D., 2014, Minnesota executive function scale: Test Manual. St Paul, Minnesota: Reflection Sciences.

Carlson, S. M., Zelazo, P. D., & Faja, S. (2013). Executive function. In P. D. Zelazo (Ed.), *The Oxford Handbook of Developmental Psychology, Vol. 1: Body and Mind* (pp. 706–743). New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199958450.013.0025.

Case, R. (1978). Intellectual development from birth to adulthood: A neo-Piagetian interpretation. *Children's Thinking: What Develops, 14*(2), 37–71.

Chan, J. Y. C. (2019). *The influences of executive function and relational language on number relation skills* [Doctoral dissertation, University of Minnesota-Twin Cities]. *University of Minnesota Digital Conservancy*. https://hdl.handle.net/11299/219330.

J. Y.C. Chan M.D. Sera M.M.M. Mazzocco The influences of relational language on early numeracy skills Child Development.in press.

Chard, D. J., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Intervention, 30*(2), 3–14. https://doi.org/10.1177/073724770503000202

Chu, F. W., vanMarle, K., & Geary, D. C. (2015). Early numerical foundations of young children's mathematical development. *Journal of Experimental Child Psychology, 132*, 205–212. https://doi.org/10.1016/j.jecp.2015.01.006

Chu, F. W., vanMarle, K., & Geary, D. C. (2016). Predicting children's reading and mathematics achievement from early quantitative knowledge and domain-general cognitive abilities. *Frontiers in Psychology, 7*(775), 1–14. https://doi.org/10.3389/fpsyg.2016.00775

Clark, C. A. C., Nelson, J. M., Garza, J., Sheffield, T. D., Wiebe, S. A., & Espy, K. A. (2014). Gaining control: Changing relations between executive control and processing speed and their relevance for mathematics achievement over course of the preschool period. *Frontiers in Psychology, 5*(107), 1–15. https://doi.org/10.3389/fpsyg.2014.00107

Clark, C. A. C., Sheffield, T. D., Wiebe, S. A., & Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Development, 84*(2), 662–677. https://doi.org/10.1111/j.1467-8624.2012.01854.x

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

Clements, D. H., Sarama, J., & Germeroth, C. (2016). Learning executive function and early mathematics: Directions of causal relations. *Early Childhood Research Quarterly, 36*, 79–90. https://doi.org/10.1016/j.ecresq.2015.12.009

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*, 63–68. https://doi.org/10.1016/j.tine.2013.12.001

Daubert, E. N. (2018). *Guided discovery activities supporting mathematical understanding in children [Doctoral dissertation]*. Maryland, U.S.A.: University of Maryland-College Park.

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

Dong, Y., Clements, D.H., Sarama, J., Dumas, D., Banse, H., & Day-Hess, C. (2021). Homogeneity of relation between mathematics and executive function competencies in the context of an intervention. In *The 2021 Annual Meeting of the American Educational Research Association (AERA)*. Orlando, FL. Retrieved from ⟨https://aera21-aera.ipostersessions.com/Default.aspx?s=66-C3–7F-D0–56-A1–73-3A-7F-A3–61-94–45-5A-51-DB⟩.

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

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149. https://doi.org/10.3989/arbor.2000.i650.965

Friso-van den Bos, I., Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2014). Explaining variability: Numerical representations in 4- to 8-year-old children. *Journal of Cognition and Development, 15*(2), 325–344. https://doi.org/10.1080/15248372.2012.742900

Fuhs, M. W., Hornburg, C. B., & McNeil, N. M. (2016). Specific early number skills mediate the association between executive functioning skills and mathematics achievement. *Developmental Psychology, 52*(8), 1217–1235. https://doi.org/10.1037/dev0000145

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology, 50*(6), 1698–1709.

Fuson, K. C. (1988). *Children's Counting and Concepts of Number*. New York: Springer.

Gashaj, V., Oberer, N., Mast, F. W., & Roebers, C. M. (2019a). Individual differences in basic numerical skills: The role of executive functions and motor skills. *Journal of Experimental Child Psychology, 182*, 187–195. https://doi.org/10.1016/j.jecp.2019.01.021

Gashaj, V., Oberer, N., Mast, F. W., & Roebers, C. M. (2019b). The relation between executive functions, fine motor skills, and basic numerical skills and their relevance for later mathematics achievement. *Early Education and Development, 30*(7), 913–926. https://doi.org/10.1080/10409289.2018.1539556

Gashaj, V., Uehlinger, Y., & Roebers, C. M. (2016). Numerical magnitude skills in 6-years-old children: Exploring specific associations with components of executive function. *Journal of Educational and Developmental Psychology, 6*(1), 157–172. https://doi.org/10.5539/jedp.v6n1p157

Geary, D. C. (2011). Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. *Journal of Developmental & Behavioral Pediatrics, 32*(3), 250–263. https://doi.org/10.1097/DBP.0b013e318209edef

Geary, D. C., Hoard, M. K., & Nugent, L. (2012). Independent contributions of the central executive, intelligence, and in-class attentive behavior to developmental change in the strategies used to solve addition problems. *Journal of Experimental Child Psychology, 113*, 49–65. https://doi.org/10.1016/j.jecp.2012.03.003

Gelman, R., & Gallistel, C. R. (1978). *The Child's Understanding of Number*. Cambridge, MA: Harvard University Press.

Ginsburg, H.P., & Baroody, A.J., 2003, Test of Early Mathematics Ability-Third Edition. Austin, TX: Pro-Ed.

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report of the Society for Research in Child Development, 22*(1), 3–22.

Goodrich, J.M., Peng, P., Bohaty, J., Leiva, S., & Thayer, L. (2021). Embedding executive function training into early literacy instruction for at-risk dual language learners: A pilot study. PsyArXiv. ⟨https://doi.org/10.31234/osf.io/xkymz⟩.

Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach*. Guilford Publications,.

Honoré, N., & Noël, M.-P. (2016). Improving preschoolers' arithmetic through number magnitude training: The impact of non-symbolic and symbolic training. *PLoS ONE, 11*(11), Article e0166685. https://doi.org/10.1371/journal.pone.0166685

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice, 22*(1), 36–46. https://doi.org/10.1111/j.1540-5826.2007.00229.x

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850–867. https://doi.org/10.1037/a0014939

Jordan, N. C., & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews, 15*(1), 60–68. https://doi.org/10.1002/ddrr.46

Kim, J. O., & Mueller, C. W. (1976). Standardized and unstandardized coefficients in causal analysis: An expository note. *Sociological Methods & Research, 4*(4), 423–438.

Kolkman, M. E., Hoijtink, H. J. A., Kroesbergen, E. H., & Leseman, P. P. M. (2013). The role of executive functions in numerical magnitude skills. *Learning and Individual Differences, 24*, 145–151. https://doi.org/10.1016/j.lindif.2013.01.004

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 78*(6), 1723–1743. https://doi.org/10.1111/j.1467-8624.2007.01087.x

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development, 84*(6), 1933–1953. https://doi.org/10.1111/cdev.12096

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

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science, 17*(5), 714–726. https://doi.org/10.1111/desc.12152

Mazzocco, M. M. M., & Kover, S. T. (2007). A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychology, 13*, 18–45. https://doi.org/10.1080/09297040600611346

Mazzocco, M. M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice, 20*(3), 142–155. https://doi.org/10.1111/j.1540-5826.2005.00129.x

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Mulcahy, C., Day Hess, C. A., Clements, D. H., Ernst, J. R., Pan, S. E., Mazzocco, M. M. M., & Sarama, J. (2021). Supporting young children's development of executive function through early mathematics. *Policy Insights from the Behavioral and Brain Sciences, 8*(2), 192–199. https://doi.org/10.1177/23727322211033005

Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2013). How approximate and exact number skills are related to each other across development: A review. *Developmental Review, 39*, 1–15. https://doi.org/10.1016/j.dr.2014.11.001

National Research Council. (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. In C. T. Cross, T. A. Woods, & H. Schweingruber (Eds.). Washington, D.C: The National Academies Press.

Nesbitt, K. T., Fuhs, M. W., & Farran, D. C. (2019). Stability and instability in the co-development of mathematics, executive function skills, and visual-motor integration from prekindergarten to first grade. *Early Childhood Research Quarterly, 46*, 262–274. https://doi.org/10.1016/j.ecresq.2018.02.003

Nguyen, T., Duncan, R. J., & Bailey, D. H. (2019). Theoretical and methodological implications of associations between executive function and mathematics in early childhood. *Contemporary Educational Psychology, 58*, 276–287. https://doi.org/10.1016/j.cedpsych.2019.04.002

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science, 24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944

Prather, R., & Alibali, M. W. (2011). Children's acquisition of arithmetic principles: The role of experience. *Journal of Cognition and Development, 12*(3), 332–354. https://doi.org/10.1080/15248372.2010.542214

Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology, 105*(2), 453–464. https://doi.org/10.1037/a0031753

Purpura, D. J., & Lonigan, C. J. (2013). Informal numeracy skills: The structure and relations among numbering, relations, and arithmetic operations in preschool. *American Educational Research Journal, 50*(1), 178–209. https://doi.org/10.3102/0002831212465332

Purpura, D. J., Schmitt, S. A., & Ganley, C. M. (2017). Foundations of mathematics and literacy: The role of executive functioning components. *Journal of Experimental Child Psychology, 153*, 15–34. https://doi.org/10.1016/j.jecp.2016.08.010

R Core Team *R: A language and environment for statistical computing* 2020R R Foundation for Statistical Computing Vienna, Austria.https://www.R-project.org.

Raghubar, K. P., & Barnes, M. A. (2017). Early numeracy skills in preschool-aged children: a review of neurocognitive findings and implications for assessment and intervention. *The Clinical Neuropsychologist, 31*(2), 329–351. https://doi.org/10.1080/13854046.2016.1259387

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable Improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*(2), 375–394. https://doi.org/10.1111/j.1467-8624.2007.01131.x

Ribner, A. D. (2020). Executive function facilitates learning from math instruction in kindergarten: Evidence from the ECLS-K. *Learning and Instruction, 65*, Article 101251. https://doi.org/10.1016/j.learninstruc.2019.101251

Ribner, A. D., Willoughby, M. T., & Blair, C. B. (2017). Executive function buffers the association between early math and later academic skills. *Frontiers in Psychology, 8*(MAY), 1–12. https://doi.org/10.3389/fpsyg.2017.00869

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42*(8), 1029–1040.

Sarama, J., & Clements, D. H. (2004). Building Blocks for early childhood mathematics. *Early Childhood Research Quarterly, 19*, 181–189. https://doi.org/10.1016/j.ecresq.2004.01.014

Scalise, N. R., & Ramani, G. B. (2021). Symbolic magnitude understanding predicts preschoolers' later addition skills. *Journal of Cognition and Development, 22*(2), 185–202. https://doi.org/10.1080/15248372.2021.1888732

Schaeffer, B., Eggleston, V. H., & Scott, J. L. (1974). Number development in young children. *Cognitive Psychology, 6*(3), 357–379. https://doi.org/10.1016/0010-0285(74)90017-6

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science, 20*(3), Article e12372. https://doi.org/10.1111/desc.12372

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development, 89*(5), 1467–1484. https://doi.org/10.1111/cdev.13068

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S.-C., & Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Developmental Neuropsychology, 35*(6), 679–697. https://doi.org/10.1080/87565641.2010.508546

Siegel, L. S. (1971). The sequence of development of certain number concepts in preschool children. *Developmental Psychology, 5*(2), 357–361. https://doi.org/10.1037/h0031426

Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science, 19*(3), 341–361. https://doi.org/10.1017/CBO9781107415324.004

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

Siegler, R. S., & Lortie-Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives, 8*(3), 144–150. https://doi.org/10.1111/cdep.12077

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*(4), 273–296. https://doi.org/10.1016/j.cogpsych.2011.03.001

Tulsky, D. S., Carlozzi, N. E., Chevalier, N., Espy, K. A., Beaumont, J. L., & Mungas, D. (2013). NIH toolbox cognition battery (CB): Measuring working memory. *Monographs of the Society for Research in Child Development, 78*(4), 70–87. https://doi.org/10.1111/mono.12035

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher, 43*(7), 352–360. https://doi.org/10.3102/0013189×14553660

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., & Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology, 80*(11, Supplement 3), S54–S64. https://doi.org/10.1212/WNL.0b013e3182872ded

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. latent structure. *Developmental Psychology, 44*(2), 575–587. https://doi.org/10.1037/0012-1649.44.2.575

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology, 108*, 436–452. https://doi.org/10.1016/j.jecp.2010.08.008

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*(1), 297–301. https://doi.org/10.1038/nprot.2006.46