



Children's Flexible Attention to Numerical and Spatial Magnitudes in Early Childhood

Mary Wagner Fuhs , Nadia Tavassolie , Yiqiao Wang , Victoria Bartek , Natalie A. Sheeks & Elizabeth A. Gunderson

To cite this article: Mary Wagner Fuhs , Nadia Tavassolie , Yiqiao Wang , Victoria Bartek , Natalie A. Sheeks & Elizabeth A. Gunderson (2021) Children's Flexible Attention to Numerical and Spatial Magnitudes in Early Childhood, *Journal of Cognition and Development*, 22:1, 22-47, DOI: [10.1080/15248372.2020.1844712](https://doi.org/10.1080/15248372.2020.1844712)

To link to this article: <https://doi.org/10.1080/15248372.2020.1844712>



Published online: 12 Dec 2020.



Submit your article to this journal 



Article views: 47



View related articles 



View Crossmark data 



This article has been awarded the Centre for Open Science 'Open Data' badge.



This article has been awarded the Centre for Open Science 'Open Materials' badge.



Children's Flexible Attention to Numerical and Spatial Magnitudes in Early Childhood

Mary Wagner Fuhs^a, Nadia Tavassolie^b, Yiqiao Wang^c, Victoria Bartek^b, Natalie A. Sheeks^d, and Elizabeth A. Gunderson^b

^aUniversity of Dayton; ^bTemple University; ^cHarvard University; ^dUniversity of Indianapolis

ABSTRACT

Young children are sensitive to both numerical and spatial magnitude cues early in development, but many questions remain about how children's attention to magnitudes relates to their early math achievement. In two studies, we tested three hypotheses related to the flexible attention to magnitudes (FAM) account, which suggests that young children's flexible attention to *both* numerical *and* spatial magnitudes is an important predictor of early math success. In Study 1, we recruited 318 preschool-age children (51.5% female; $M_{age} = 54.7$ months, $SD_{age} = 4.24$ months) and assessed them at two time points on a battery of math, executive function, and language measures, including a novel assessment of FAM ability developed for this study. Consistent with our hypotheses, we found that young children had specific difficulty flexibly shifting between numerical and spatial magnitudes, that their FAM ability was related to both their executive function and math skills controlling for covariates, and finally that their FAM ability at the first time point was predictive of their math achievement growth over time controlling for executive function skills and covariates. In Study 2, we recruited 157 preschool-age children (47.4% female; $M_{age} = 53.36$ months, $SD_{age} = 7.17$ months). We replicated the findings of Study 1 and extended them to account for children's non-symbolic numerical magnitude discrimination skills. Implications of the results of these studies for early math activities development are discussed.

Strong math skills are important for later academic and career success, but many children struggle from a young age to learn math concepts (Jordan & Levine, 2009). It is therefore critical from both a theoretical and practical standpoint to better understand the emergence of young children's math skills. The flexible attention to magnitudes (FAM) account proposed here argues that a fundamental challenge young children face in early math – that is not addressed by current interventions – is disentangling numerical magnitudes (discrete, countable units) and spatial magnitudes (continuous quantities including surface area, density, width, etc.) and making decisions about which are relevant in early math problem solving. We define FAM as the ability to flexibly shift attention between different dimensions of magnitude that may be incongruent with one another, including numerical magnitudes and spatial magnitudes. Three specific hypotheses stemming from this account are tested in two distinct studies: a) Attending to and flexibly shifting between *both* numerical *and* spatial magnitudes is a unique challenge in early childhood math, b)

CONTACT Mary Wagner Fuhs  mfuhs1@udayton.edu  Department of Psychology, University of Dayton, SJ 319, 300 College Park Avenue, Dayton, OH 45469.

© 2020 Taylor & Francis

Children's FAM ability is related to both their executive function (EF) skills and math achievement, and c) Children's FAM ability predicts their math achievement after controlling for EF skills and non-symbolic numerical magnitude discrimination skills (i.e., approximate number system acuity).

Magnitude discrimination and math achievement

Young children are sensitive to both numerical magnitudes and spatial magnitudes early in life, prior to formal math instruction (Newcombe, Levine, & Mix, 2015). Children's ability to discriminate numerical and spatial magnitudes is dependent on the ratio of the stimuli to be compared (Weber's law) (Bonny & Lourenco, 2015; Feigenson, Dehaene, & Spelke, 2004), and the acuity of this representation increases with age (Bonny & Lourenco, 2015; Halberda & Feigenson, 2008). These ratio-limited representations are evolutionarily-old and shared with non-human animals (Feigenson et al., 2004). More recent research suggests neural overlap in the processing of non-symbolic numerical magnitudes, symbolic numerical magnitudes, and non-numerical magnitudes (e.g., size) in both the parietal and frontal areas of the brain (Sokolowski, Fias, Ononye, & Ansari, 2017).

A substantial body of research has investigated whether and how humans' numerical magnitude discrimination skills (i.e., the ability to compare quantities without counting) relate to their symbolic math achievement. Results from correlational studies linking children's non-symbolic magnitude discrimination skills with symbolic math achievement have been mixed, with some finding significant relations (Halberda, Mazzocco, & Feigenson, 2008; Libertus, Feigenson, & Halberda, 2011; Starr, DeWind, & Brannon, 2017) and others finding no relation (Sasanguie, De Smedt, Defever, & Reynvoet, 2012; Sasanguie, Defever, Maertens, & Reynvoet, 2014). Although recent meta-analyses reveal a small but positive relation between non-symbolic numerical magnitude discrimination skills and math achievement among young children (Chen & Li, 2014; Fazio, Bailey, Thompson, & Siegler, 2014; Schneider et al., 2017), much debate still surrounds the interpretation of this correlation. In part, this debate stems from findings that the associations between non-symbolic magnitude discrimination skills and math achievement are significantly smaller in effect size ($r = .241$) compared to associations between symbolic magnitude discrimination skills and math achievement ($r = .302$) (Schneider et al., 2017).

Here, we argue that the relation between young children's non-symbolic numerical magnitude discrimination skills and math achievement is largely driven by their ability to attend to numerical magnitudes in the face of conflicting spatial information, a component of a broader FAM ability. Supporting this, the relation between non-symbolic numerical magnitude discrimination skills and math achievement depends on variations in congruency and incongruency between numerical and spatial magnitudes in commonly-used numerical discrimination tasks (Fuhs & McNeil, 2013; Gilmore et al., 2013). When explicitly compared, conditions of salient incongruency between numerical and spatial magnitudes – for example, asking which set has more stars, 4 large stars or 8 small stars – result in the lowest performance for young children (Fuhs & McNeil, 2013; Gilmore et al., 2013; Szucs, Nobes, Devine, Gabriel, & Gebuis, 2013). Importantly, performance on these incongruent trials drives the relation between non-symbolic numerical magnitude discrimination tasks and children's symbolic math achievement (Fuhs & McNeil, 2013). In contrast, the ability to compare numerical magnitudes that are congruent with spatial magnitudes does

not consistently relate to math achievement. In the case of congruent trials, children do not have to choose between attending to numerical or spatial information as focusing on either will yield the correct response.

Some studies have suggested that focusing on numerical magnitudes in the face of incongruent spatial information relies heavily on domain-general EF (Fuhs & McNeil, 2013; Gilmore et al., 2013). EF skills can be broadly defined as skills associated with pre-frontal cortex functioning that contribute to goal-directed behavior and problem solving and include working memory, inhibitory control, and cognitive flexibility or attention shifting (Diamond, 2013; Garon, Bryson, & Smith, 2008; Miyake & Friedman, 2012). This is consistent with overlap in the activation of neural networks in the pre-frontal cortex for both processing EF tasks as well as numerical and non-numerical magnitude tasks (Diamond, 2013; Sokolowski et al., 2017). Although not without debate, empirical research generally supports viewing EF skills as representing a unitary construct in early childhood that becomes more differentiated over time (Lee, Bull, & Ho, 2013). In the current study, we interpret our results in line with the empirical support for a unitary construct of EF skills in early childhood while acknowledging the limitations of using a single measure of EF skills.

Given the robust association between children's EF skills and early math achievement (e.g., Cragg & Gilmore, 2014; Fuhs, Hornburg, & McNeil, 2016; Fuhs, Nesbitt, Farran, & Dong, 2014; Miller, Rittle-Johnson, Loehr, & Fyfe, 2016; Nesbitt, Fuhs, & Farran, 2019; Prager, Sera, & Carlson, 2016; Schmitt, Geldhof, Purpura, Duncan, & McClelland, 2017; Verdine, Irwin, Golinkoff, & Hirsch-Pasek, 2014), one possibility is that general EF skills partially or fully explain why children's non-symbolic numerical magnitude discrimination skills are sometimes linked to their math achievement (Clayton & Gilmore, 2015; Fuhs & McNeil, 2013; Gilmore et al., 2013; Leibovich, Katzin, Harel, & Henik, 2017). We propose that FAM, a skill that is related to EF skills, better explains this relation.

Flexible attention to magnitudes (FAM)

As noted previously, we define FAM as the ability to attend to and flexibly switch focus between dimensions of magnitude, both numerical and spatial, in the face of potentially incongruent information from other magnitude dimensions. One key distinction between FAM and numerical magnitude discrimination skills is that FAM involves the ability to focus not only on numerical magnitudes, but also on spatial magnitudes. Though much less well-studied than for numerical magnitudes, the ability to discriminate spatial magnitudes (e.g., surface area) is associated with symbolic math achievement in both adults (Lourenco, Bonny, Fernandez, & Rao, 2012) and children (Bonny & Lourenco, 2015).

Of course, spatial and numerical magnitudes are deeply related, both in terms of their mental representation and in the physical world. Spatial and numerical magnitudes are represented in overlapping neural areas (Pinel, Piazza, Le Bihan, & Dehaene, 2004; Sokolowski et al., 2017), and single-cell recordings in non-human primates indicate that some neurons in these areas are responsive to both types of magnitudes (Tudusciuc & Nieder, 2007). This representational overlap may be driven in part by the fact that, in the real world, numerical and spatial magnitudes are frequently positively correlated (e.g., 10 apples is more than 5 apples in number, total surface area, and total volume) (Clearfield & Mix, 1999). Researchers have theorized that humans begin with a “generalized magnitude system” or “sense of magnitude”, with overlapping mental representations of numerical and



spatial magnitudes, and a key challenge in development is for children to understand the distinctions between these dimensions of magnitude (e.g., number, surface area, height, width, density) (Leibovich et al., 2017; Newcombe et al., 2015). Building on these theories, a key goal of this study is to directly test children's ability to overcome the challenge of disentangling numerical and spatial magnitudes, in other words, their FAM ability.

Relation of FAM to math achievement and EF skills

We expect FAM ability to relate to math skills not only in early childhood, but also throughout development. This is because many math skills, such as number line estimation, measurement, and proportional reasoning, require children to pay attention in a flexible way to both numerical and spatial magnitudes. Number line estimation requires coordination between numerical magnitudes (specific numbers to be placed on the number line) and spatial magnitudes (line lengths associated with distances between points on the number line), and strongly predicts math achievement from preschool onward (e.g., Ramani & Siegler, 2008). Similarly, linear measurement (e.g., using a ruler) involves imposing discrete, numerical units onto a continuous space and understanding the concept of spatial units (Solomon, Vasilyeva, Huttenlocher, & Levine, 2015). The ability to focus on spatial magnitudes is also beneficial for proportional reasoning: children are more successful in reasoning about spatial proportions than about numerical proportions, reflecting over-reliance on numerical quantities even when they are inappropriate for the task at hand (Boyer & Levine, 2015).

We expect FAM ability to correlate with general EF skills, but also to predict later math achievement over and above general EF. Flexibly shifting between numerical and spatial magnitudes when needed, and ignoring irrelevant magnitude information, is expected to draw on children's general EF skills. Specifically, we expect attention switching to be helpful in shifting focus from one magnitude dimension to another. For these reasons, we expect general EF skills to positively correlate with FAM. However, we expect the relation between FAM and EF to be moderate. A driving rationale for this prediction is that the experiences that ought to strengthen children's ability to attend to numerical and spatial magnitudes, such as reading counting books that emphasize incongruity between size and number, differ from those that we would expect to strengthen general EF (e.g., music and movement games that do not involve numerical or spatial magnitudes, Schmitt, McClelland, Tominey, & Acock, 2015). Importantly, we expect FAM ability to predict later math achievement even after accounting for general EF skills.

Another way to conceptualize FAM is as a type of domain-specific EF skill. Therefore, this proposal is consistent with recent research linking number-specific EF to children's math growth over and above general (non-numerical) EF skills (Wilkey, Pollack, & Price, 2020; Wilkey & Price, 2019). The results of these studies suggest that domain-specific EF skills related to attending to numerical, and, as we argue, spatial magnitudes – in other words, FAM – will be a stronger predictor of children's math skills growth compared to EF skills alone.

Study 1

In order to test children's FAM ability, we developed a measure that requires children to choose between two sets based on number for some trials, and based on size for other trials. The sets are always incongruent with respect to size and number, e.g., 6 small stars versus 3

large stars, which requires children to attend to the dimension of interest (e.g., size) while inhibiting the other dimension (e.g., number). Importantly, we designed this measure so that individual differences in children's *acuity* of magnitude discrimination for either number or size should not impact their performance, by employing ratio differences in both number and surface area (1:2, 1:3, 2:3) that are very easy to discriminate for preschoolers (Halberda & Feigenson, 2008).

Our novel FAM measure is untimed and involves small numerical quantities (less than 10), making it more naturalistic than typical non-symbolic numerical magnitude discrimination tasks that involve approximating large discrete quantities without counting. We chose this approach because preschoolers encounter numbers less than 10 much more frequently than larger numbers, both in informal settings such as naturalistic parent-child interactions (Levine, Suriyakham, Rowe, Huttenlocher, & Gunderson, 2010) as well as in more formal interactions such as reading counting books (Ward, Mazzocco, Bock, & Prokes, 2017). Further, early math activities, whether formal or informal, often involve sets of objects or tangible math manipulatives, and rarely require children to make estimations while being prevented from counting. Therefore, this approach provides more straightforward theoretical and empirical connections to the activities children spend their time working on in early math.

We assessed children in the fall (Time 1) and spring (Time 2) of their preschool year on our novel measure of FAM as well as their math achievement. We also assessed children on their EF skills and vocabulary skills in the fall, and we conducted cross-lagged panel path analyses to test our three hypotheses of interest.

Study 1 method

Participants

All participants were recruited from preschools in Philadelphia, PA, and Dayton, OH metropolitan areas. Parents or guardians of all participants gave written consent. Demographic information was obtained from parents or guardians via an optional questionnaire. There were 318 children whose parents consented to the study. We assessed 309 children aged 42–65 months (51.4% female; $M_{age} = 54.7$ months, $SD_{age} = 4.24$ months) on at least one assessment in the fall of their preschool year. There were 293 children who were assessed on at least one assessment in the spring of the same preschool year.

Families reported on their income bracket and highest level of parent education. The largest percentage of families (37.1%) were living with yearly incomes below 25,000, USD followed by 17.6% of families within the 25,001 USD – 42,000 USD bracket, 6.9% in the 42,001 USD – 60,000 USD bracket, 5.7% in the 60,001 USD – 79,000 USD bracket, and 18.6% in the 79,001 USD and up bracket (14.2% did not report their family income). The majority of families had a primary parent/guardian with either a high school diploma or GED or some college (up to an Associate's Degree) as their highest level of education (68.3%), followed by 15.4% having a Bachelors, 1.6% having some graduate training, and 12.6% having a graduate degree (2.2% did not report their highest level of education). Parents also reported their child's race/ethnicity: 36.8% were Black or African-American,



33.3% White, 7.5% Hispanic or Latinx, 2.2% Asian or Asian American, and 12.9% multi-racial or other (7.3% did not report their child's race/ethnicity).

Stimuli and procedure

Children completed assessments in a single session that lasted between 20 and 30 minutes in the fall (Time 1) and spring (Time 2) of their school year. The fall and spring sessions were on average 5.56 months apart ($SD = 0.66$ months). At both Time 1 and Time 2, children completed measures of their FAM ability and math achievement as a primary interest of the study. We also assessed children's EF skills and general vocabulary at Time 1.

It should be noted that as part of a larger longitudinal study of preschool practices at the OH research site, children's literacy skills were assessed at both time points, and their EF skills and general vocabulary were also assessed at the spring time point. The PA research site also collected these measures to maintain consistency in data collection practices across sites. However, as those assessments were collected to address separate research questions about the impact of preschool practices, they are not reported here. Each child received measures of EF skills first, then vocabulary, literacy skills, math achievement, and FAM at the end. Tasks were administered in a single order for all participants.

FAM task

We developed a novel measure to capture children's FAM skills. Modeled after reliable and valid EF tasks for young children (Carlson & Zelazo, 2014), this paper-based task involved three blocks of assessment that increased in complexity. In each block, children were shown two pictures of boxes that contained stars (see [Figure 1](#)). One box had more stars, while the other box had stars with a greater spatial magnitude (i.e., greater total surface area), making all trials incongruent with respect to numerical and spatial magnitudes. To calculate total surface area, we calculated the area for each object around which one could draw a square that fully captured the edges of each star, and then summed the areas of each star, consistent with calculation measures used in prior studies (e.g., Halberda et al., 2008). Individual item size was homogeneous within boxes. Convex hull was not precisely controlled, although attempts were made to visually equate convex hull across object sets within each trial. The number of stars within each box ranged from 1–10 and the ratio differences for numerical and spatial magnitudes of stars between the two boxes were 1:3, 1:2, or 2:3. The background color of the boxes was either black or white depending on whether children were asked to make judgments based on numerical or spatial magnitudes. We created two versions of the task which differed in the pairing of background color (black or white) with magnitude judgment (number or size). Children were randomly assigned to complete one of the two versions.

In the first block (size trials), children were asked to play the size game. The experimenter explained that the size game entailed pointing to the box that had bigger stars, and illustrated this with a demonstration trial. Children then received two familiarization trials. If they responded correctly, the experimenter gave positive feedback and repeated the rule of the size game. If children responded incorrectly, the experimenter corrected the child while emphasizing the spatial and numerical features of the stars in the two boxes (e.g., "Well, there are more stars in that box, but the stars in the other box are bigger. There are fewer of them, but they're bigger," adapted from Negen & Sarnecka, 2015). The two

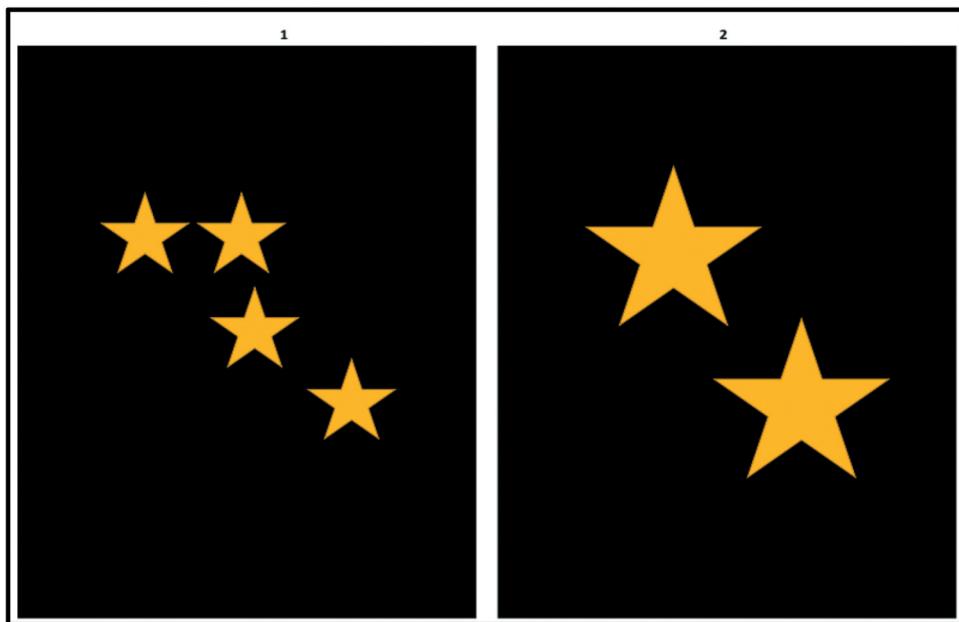


Figure 1. FAM task example item. Experimenter: “This is the **size (number)** game. Which box has bigger (more) stars?”

familiarization trials were followed by six test trials. The experimenter asked which box has bigger stars before each test trial. No feedback was provided on test trials.

In the second block (number trials), children were asked to switch rules and play the number game by pointing to the box with more stars. As in the size game, the experimenter first demonstrated a trial, then children completed two familiarization trials and six test trials.

On the third block of trials (mixed trials), children were asked to switch between the size game and the number game based on the background color of the boxes. Specifically, they were asked to play the size game and point to the box with bigger stars if they saw one color of boxes, and to play the number game and point to the box with more stars if they saw the other color of boxes. The colors associated with each rule (black and white) were the same as those in the first and second blocks of trials. The mixed game was comprised of 12 test trials, which were preceded by two experimenter demonstration trials and two familiarization trials. The experimenter repeated the rule twice during the test trials, once before the first trial and once before the sixth trial.

During all trials, children were allowed to count the items (i.e., stars) if they chose, but were not prompted to do so. The experimenter recorded whether children counted on each item (e.g., “one, two, three”) or stated cardinal values of sets (e.g., “There are three stars” or “one, two, three, three stars”). The experimenter recorded which box children had pointed to in each trial and the percentage of accurate responses was calculated for each block. Children’s scores were calculated as the percent correct from the combined number and mixed trials. This decision was both empirical and theoretical. Empirically, a continuous post-switch trial accuracy rate has been used for executive function tasks

(e.g., DCCS) in prior studies (Fuhs, Hornburg, et al., 2016). Theoretically, we are interested in children's ability to both shift from attending to one dimension to another as well as to flexibly shift between them within the same task. In the size trials, although children are asked to attend to spatial information while ignoring numerical information, they are not asked to shift from previously focusing on numerical information. Therefore, only the Number and Mixed trials involved flexible shifting from one dimension to another. However, we also explored other methods of scoring in the exploratory follow-up analyses.

Math achievement

We assessed children's math achievement using the Woodcock-Johnson IV Tests of Early Cognitive and Academic Development (ECAD) Number Sense subtest. Number Sense is a standardized, norm-referenced measure of children's number development skills and has demonstrated reliability and validity (Schrank, McGrew, & Mather, 2015). The Number Sense test includes visually- and orally-presented problems that involve mathematical skills including recognizing numbers, counting, ordering numbers, and estimating quantities. W scores (a Rasch-based metric; Wendling, Mather, LaForte, McGrew, & Schrank, 2015) were calculated from the Woodcock-Johnson Online Scoring and Reporting Program (Schrank & Dailey, 2014, 2015) and were used in the analyses.

EF skills

Children completed the Minnesota Executive Function Scale (MEFS) as a measure of EF skills via the MEFSTM App (Carlson & Zelazo, 2014). The MEFS was developed based on the Dimensional Change Card Sort Task (DCCS; Zelazo, 2006). We chose this measure specifically because it focuses on assessing children's cognitive flexibility and attention shifting skills, skills we viewed as similar to the ones required for the FAM task. It is a computerized measure administered on an iPad and takes 2–7 minutes to complete. It is nationally-normed for ages 2 years and up and is adaptive based on children's performance. There are 7 levels in this task, each with increasing difficulty. The starting level is pre-determined by age: older children begin at a higher level. Based on performance, children move up or down within the levels. On each level children are asked to sort cards based on a specific characteristic (i.e., color, shape, or size). Conflict is introduced such that children must switch from one dimension to another. For example, on one level, children are asked to play the "color game" and sort cards by color, and then to switch and play the "shape game" by sorting them by shape. On a more difficult, mixed-rule level, children are asked to play the color game if they see cards with a black border around them and play the shape game if they see cards with no black border. We used nationally-normed Standard Scores based on child age as our measure of children's performance. Scoring is determined by an algorithm that combines both response time and accuracy, yielding a standard score with a mean of 100 and a standard deviation of 15. These standard scores were used in all analyses.

Vocabulary

Children completed the Picture Vocabulary subtest of the Woodcock-Johnson IV ECAD as a measure of their language skills. The Picture Vocabulary subtest is normed for ages 2 through 10 (Schrank et al., 2015). In this task, children are asked to name pictures of

increasing difficulty until basal and ceiling criteria are reached. W scores were calculated and used in the analyses.

Missingness

There were 318 children whose parents provided consent, and of those, 293 had complete assessment data at Time 1. Reasons for complete missingness (not assessed on any measures) at Time 1 were that children were outside of our desired age range ($n = 4$), absent ($n = 3$), or did not give assent ($n = 2$). Partial missingness at Time 1 (children stopped assessments before session was complete) was due to refusal to complete one or more assessments ($n = 13$) or experimenter error ($n = 3$; 2 instances of the WJ-ECAD Number Sense test basal or ceiling not established and 1 instance of missing trial in size trials of FAM). At Time 2, 278 children had complete assessment data. Spring complete missingness (not assessed on any measures) was due to withdrawal from pre-k program ($n = 12$), absences ($n = 6$), children not being in age range ($n = 4$), and not giving assent ($n = 3$). Partial missingness at Time 2 (children stopped assessments before session was complete) was due to refusal to complete one or more assessment ($n = 6$) or experimenter error ($n = 9$, 7 instances of the WJ-ECAD Number Sense test basal or ceiling not established, 1 instance of missing trial in size trials of FAM, and 1 instance of missing FAM and WJ-ECAD data). Summarizing across the fall and spring assessments, 263 children had complete data on all direct assessments used in analyses for the current study. Compared to children who were missing one or more assessment scores, children with complete assessment data scored significantly higher on a number of assessments (fall MEFS, fall WJ-IV ECAD Number Sense, spring WJ-IV ECAD Number Sense, all $p < .05$). Finally, 52 parents/guardians (16.4% of sample) chose not to answer one or more demographic questions relating to gender, income, and education. Missing data were treated as missing at random (Little's MCAR test – $\chi^2(42) = 87.09, p < .001$), and therefore Full Information Maximum Likelihood estimation procedures were used in Mplus v. 8 for correlations and path analyses to maximize usage of available data and limit bias.

Study 1 results

Descriptive analyses

Descriptive analyses were conducted in SPSS v. 24. Fall and spring descriptives are presented in [Table 1](#). Both the MEFS assessment and the WJ-IV ECAD subtests are normed based on age, on a scale with a mean of 100 and a standard deviation of 15. Although W-scores were used in all analyses for the WJ-IV ECAD subtests, we report standardized scores based on age here for comparison with national norms. For the ECAD Number Sense subtest assessing math skills, children's average Standard Score in the fall was 94.79 ($SD = 16.61$), slightly below average. Children's average Standard Score in the spring was also slightly below average ($M = 96.73, SD = 15.40$). Children's language skills, as assessed by the Picture Vocabulary subtest, were around the average at $M = 101.42$ ($SD = 13.46$). Children significantly improved their W scores on the Number Sense subtest from fall to spring ($t(278) = 11.43, p < .001, d = .68$).

**Table 1.** Study 1 descriptive statistics.

	N	Min	Max	M	SD
Fall MEFS Standard Score	309	61.0	124.0	96.98	8.39
Fall Picture Vocabulary W Score	306	399.00	493.00	460.56	13.12
Fall Number Sense W Score	305	362.00	476.00	421.63	17.71
Fall FAM Size Trials Accuracy	307	0.33	1.00	0.96	0.10
Fall FAM Number Trials Accuracy	305	0.00	1.00	0.78	0.31
Fall FAM Mixed Trials Accuracy	296	0.17	1.00	0.56	0.18
Fall FAM Number/Mixed Trials Accuracy	296	0.17	1.00	0.63	0.18
Spring Number Sense W Score	285	362.00	486.00	431.22	16.64
Spring FAM Size Trials Accuracy	289	0.17	1.00	0.97	0.11
Spring FAM Number Trials Accuracy	288	0.00	1.00	0.89	0.22
Spring FAM Mixed Trials Accuracy	288	0.08	1.00	0.66	0.22
Spring FAM Number/Mixed Trials Accuracy	288	0.28	1.00	0.74	0.18

Examining FAM task performance

We first examined potential differences in scores across FAM Forms A and B. We observed no significant differences across different forms for children's accuracy on all trial levels for both fall and spring (all $p > .15$). Therefore, we collapsed across forms in all analyses.

Performance on each FAM trial type in the fall and spring is reported in **Table 1** and **Figure 2**. As predicted, children performed significantly better on the FAM Size trials compared to the FAM Number trials and Mixed trials, and performed better on the FAM Number trials compared to the Mixed trials in both the fall and the spring (see **Table 2** for paired-sample t -tests). We also conducted one-sample t -tests comparing performance to

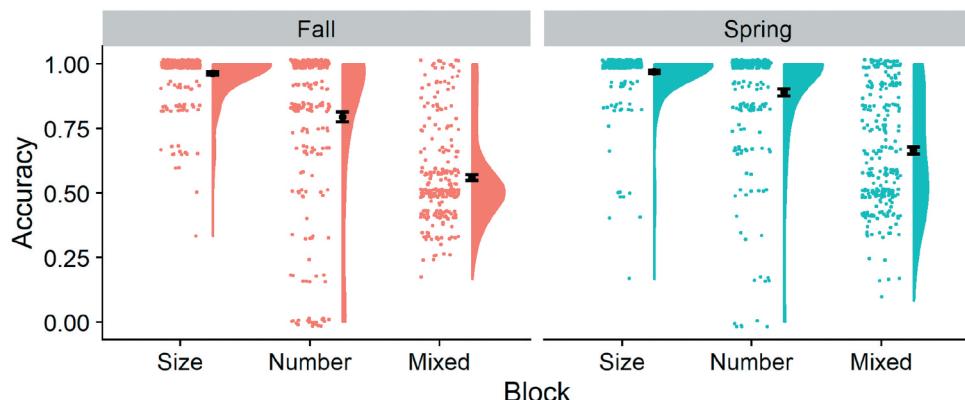


Figure 2. Study 1: children's accuracy on the FAM task by trial type. Note. Black bars show error bars ± 1 SE.

Table 2. Study 1 within-subjects t-tests comparing children's performance across FAM trial types.

Time	Comparison	t	df	p	d
Fall	Size Trials Accuracy – Number Trials Accuracy	9.52	303	<.001	.56
Fall	Size Trials Accuracy – Mixed Trials Accuracy	36.13	294	<.001	2.05
Fall	Number Trials Accuracy – Mixed Trials Accuracy	12.17	295	<.001	.69
Spring	Size Trials Accuracy – Number Trials Accuracy	5.99	286	<.001	.35
Spring	Size Trials Accuracy – Mixed Trials Accuracy	21.41	286	<.001	1.29
Spring	Number Trials Accuracy – Mixed Trials Accuracy	13.67	287	<.001	.83

chance (50%) within each block and time-point; performance was above chance in all cases, $p < .001$. Children made significant improvements from fall to spring in their FAM skills on the Number trials ($t(281) = 5.25, p < .001, d = .33$) and Mixed trials ($t(274) = 7.13, p < .001, d = .46$), but not the Size trials ($t(284) = 0.68, p = .495, d = .04$), as their scores were already near the ceiling.

Counting was not very common (5.4% of trials), nor was labeling the cardinal value (1.6% of trials). Children were more likely to count on number trials than they were on size or mixed trials in the fall ($t(304) = 2.28, p = .024, d = .15$; $t(296) = 3.27, p = .001, d = .20$) and spring ($t(290) = 1.84, p = .067, d = .10$; $t(290) = 3.81, p < .001, d = .26$), and they were more likely to label the cardinal value on number versus mixed trials in the spring ($t(290) = 2.47, p = .014, d = .17$). All other comparisons of fall and spring counting and cardinality across trial types were not significant ($p > .10$).

Correlational and path analyses

Correlations

Correlations among all measures are reported in [Table 3](#). As expected, all of the child assessment measures were significantly correlated with one another. Children's MEFS EF skills standard scores were significantly correlated with math achievement in the WJ-IV ECAD Number Sense assessment both within and across time points. Also as expected, children's FAM task performance was significantly correlated with their math achievement. Age and family income were significantly positively associated with all of the assessment scores, with the exception of one negative correlation between age and the MEFS standard score. Gender was not significantly related to any child assessment scores, and parent education was related to children's performance on the WJ-IV ECAD subtests but not the MEFS standard score or the FAM task.

Path analyses

In our primary analysis, we examined the longitudinal associations between FAM task performance and children's math skills using a cross-lagged panel model where spring Number Sense W scores were regressed on fall FAM Number + Mixed trials accuracy as well as fall Number Sense W scores, fall EF, and fall Picture Vocabulary W scores (see [Figure 3](#)). We also regressed spring FAM Number + Mixed trials accuracy on fall Number Sense W scores, fall FAM Number + Mixed trials accuracy, fall EF, and fall

Table 3. Study 1 correlations among variables.

Variable	1	2	3	4	5	6	7	8	9
1. Fall MEFS Standard Score	1								
2. Fall Picture Vocabulary W Score	.31***	1							
3. Fall Number Sense W Score	.33***	.63***	1						
4. Fall FAM Number/Mixed Trials Accuracy	.32***	.34***	.39***	1					
5. Spring Number Sense W Score	.35***	.58***	.69***	.40***	1				
6. Spring FAM Number/Mixed Trials Accuracy	.29***	.33***	.38***	.42***	.42***	1			
7. Age	-.18**	.29***	.22***	.14**	.23***	.16**	1		
8. Gender	-.04	-.03	.01	-.02	.05	.01	.05	1	
9. Family Income	.27***	.27***	.27***	.16**	.27***	.13*	.06	.03	1
10. Parent Education	.08	.17**	.17**	.05	.21***	.05	.02	.03	.57***

* $p < .05$, ** $p < .01$, *** $p < .001$.

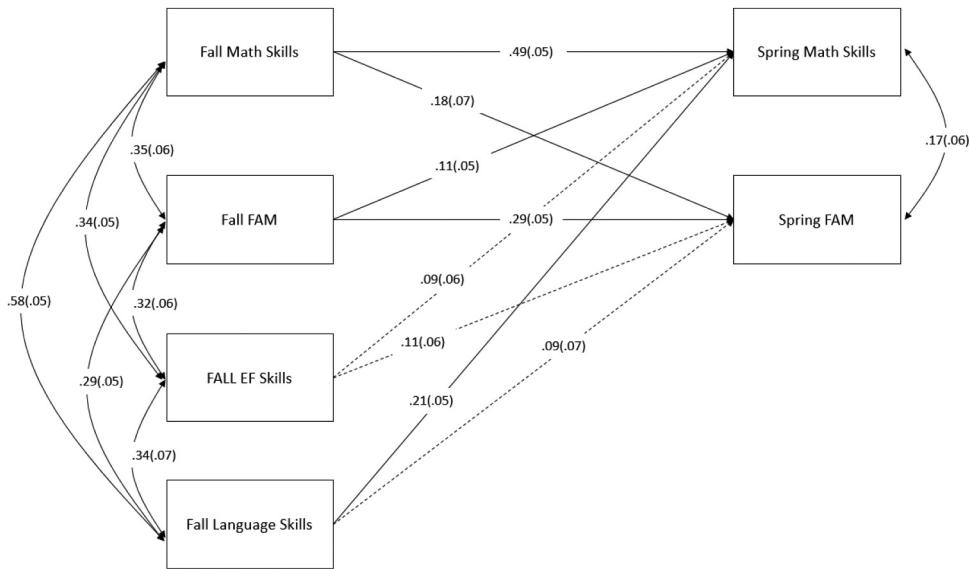


Figure 3. Study 1: Cross-lagged panel model. Note: Fall FAM = Fall FAM Number/Mixed Trials Accuracy. Fall Math Skills = Fall WJ-IV ECAD Number Sense subtest W score. Fall EF Skills = Fall MEFS Standard Score. Fall Language Skills = Fall WJ-IV ECAD Picture Vocabulary subtest W score. Standardized path coefficients (SEs) are shown. Solid lines indicate significant paths ($p < .05$), and dashed lines indicate non-significant paths. All fall measures were also regressed on child age, child gender, family income, and parental education (paths not shown).

Picture Vocabulary W scores. Fall child assessments were allowed to covary and were regressed on our covariates (age, gender, family income, and parent education). Fall FAM variables were regressed on Fall FAM Form. Family income and parent education were also allowed to covary. We accounted for the nesting of children within classrooms by including classroom as a cluster variable using the MPlus function TYPE = COMPLEX. The model was a good fit for the data, $\chi^2(19) = 14.80$, $p = .735$, RMSEA = .000, CFI = 1.00 SRMR = .030. We found evidence of a significant bidirectional association between children's FAM ability and their math skills, with fall FAM ability predicting children's spring math skills ($B = .11$, $SE = .05$, $p = .040$), and fall math skills predicting children's spring FAM ability ($B = .18$, $SE = .07$, $p = .016$).

Exploratory analyses

We explored four additional models using alternative scoring for fall FAM task performance (originally scored as FAM Number + Mixed trials accuracy). We were interested in testing potential alternative explanations of the path model results specific to the effect of fall FAM ability on spring math skills. In each of the exploratory models, we retained all variables and paths in our original model (see Figure 3) with the exception of replacing the fall FAM task performance variable with a new variable with alternative FAM task scoring.

First, we asked if children's FAM task performance was primarily driven by their ability to focus on numerical magnitudes when in conflict with spatial magnitudes, but was less driven by their ability to shift back and forth between numerical and spatial magnitudes in

the mixed trials. In other words, was the association between FAM ability and children's spring math skills driven primarily by the Number trials and not the Mixed trials? We explored this possibility by replacing the original fall FAM Number + Mixed combined accuracy variable that was used in our model depicted in [Figure 3](#) with two separate variables, Number trials accuracy and Mixed trials accuracy. The standardized path coefficients for each accuracy variable were similar in magnitude and neither was significant (Number trials, $B = .07$, $SE = .05$, $p = .191$; Mixed trials, $B = .05$, $SE = .06$, $p = .361$). We also explored this possibility by adding the FAM Size trials as an additional predictor in the model; we did not include Size trials in the original model as they did not involve shifting from one dimension to another either across or within trial levels because they were always administered first. Again, none of the standardized path coefficients from each type of fall FAM trial to spring math achievement were significant (Size trials, $B = .03$, $SE = .05$, $p = .569$; Number trials, $B = .07$, $SE = .05$, $p = .178$; Mixed trials, $B = .05$, $SE = .06$, $p = .414$).

One might also wonder if, within the FAM Mixed trials, the trials asking children to attend to number drove the link between fall FAM skills and spring math skills, rather than the trials requiring children to attend to size. We explored this possibility by re-running the path model described above, with both Number and Mixed trials accuracy included as separate variables, but further separating the Mixed trials into Mixed Number ($n = 6$) and Mixed Size ($n = 6$) trials. In this model, there were three fall FAM task performance predictors (Number, Mixed Number, and Mixed Size). Again, we found similar strengths of associations across these three variables in predicting children's spring math skills and again, none were significant (Number trials, $B = .06$, $SE = .06$, $p = .360$; Mixed Number trials, $B = .07$, $SE = .07$, $p = .322$; Mixed Size trials, $B = .04$, $SE = .06$, $p = .513$).

Finally, we explored whether flexibly shifting back and forth between trials would be particularly important for children's spring math skills, regardless of which magnitude dimension was involved. According to the FAM account, this would be the most likely explanation for a significant association between fall FAM ability and children's spring math skills. To test this, we repeated the same model procedure as prior models, but our FAM task variables now included FAM Number trials, FAM Switch trials (trials in which the prior item required focus on a different magnitude dimension), and FAM Non-Switch trials (trials in which the prior item required focus on the same magnitude dimension) rather than breaking down the FAM Mixed trials by size and number. Note that there were an equal number of Switch and Non-Switch trials for number and size in the Mixed condition ($n = 6$). Again, these were entered into the same model as a replacement for the original fall FAM task variable used in [Figure 3](#), and we found that FAM Switch trials were most closely, and significantly, associated with children's math achievement in the spring (Number trials, $B = .07$, $SE = .05$, $p = .161$; Mixed Switch trials, $B = .10$, $SE = .05$, $p = .023$; Mixed Non-Switch trials, $B = -.05$, $SE = .05$, $p = .329$).

Study 1 discussion

Results supported our hypothesis that children's FAM ability is a significant predictor of their math achievement across the preschool year, controlling for covariates and children's general EF skills. As this was the initial examination of a newly developed account and assessment of FAM ability, there are many questions that remain. One question that was not answered with the design of the FAM task is whether children's performance would have



differed if the numerical magnitude trials were presented first, prior to the switch to spatial magnitude trials. If children are more prone to attending to spatial magnitudes early in life (Mix, Huttenlocher, & Levine, 2002), then they may be more likely to develop a stronger prepotent response in the pre-switch trials in our task, more so than if they had been asked to switch from pre-switch number to post-switch size trials. In Study 2, we tested this by counterbalancing the pre-switch trials to start with either size or number trials.

Another important next step was to determine whether the association between children's FAM ability and their math achievement holds not only controlling for EF, but also when controlling for children's non-symbolic numerical magnitude discrimination skills. Non-symbolic numerical magnitude discrimination skills, at least on trials that are incongruent with respect to spatial and numerical magnitudes, may assess children's attention to numerical magnitudes when faced with a conflict in spatial magnitude, but do not assess children's flexible shifting between differing dimensions of magnitude. This could explain, at least in part, prior mixed findings with respect to the association between children's non-symbolic numerical magnitude discrimination skills and their math achievement. We also addressed this limitation in Study 2.

Study 2 method

Participants

We received written parental consent from 161 participants recruited from preschools in Philadelphia, PA, and Dayton, OH metropolitan areas. Of those, 157 completed at least one assessment (the four excluded children withdrew from the pre-k program [$n = 1$], were absent [$n = 1$] or did not assent [$n = 2$]). The final sample of 157 children were 36–68 months old (47.4% female; $M_{age} = 53.36$ months, $SD_{age} = 7.17$ months). Children gave verbal assent before participating.

Parents reported their annual family income and level of education on an optional demographic questionnaire. Parental report indicated that 40.1% of families were living with yearly incomes below 25,000, USD 3.2% within the 25,001 USD – 42,000 USD bracket, 1.9% in the 42,001 USD – 60,000 USD bracket, 3.2% in the 60,001 USD – 79,000 USD bracket, and 15.3% in the 79,001 USD and up bracket (36.3% did not report their family income). Parents reported their highest level of education; 4.5% did not complete high school, 22.3% had a high school diploma or GED, 28.7% had some college (up to an Associate's Degree) 14% had a Bachelor's degree, 1.3% had some graduate training, and 15.3% had a graduate degree (14% did not report their highest level of education). Parents also reported their child's race/ethnicity: 38.9% were Black or African-American, 26.8% White, 5.1% Hispanic or Latinx, 1.9% Asian or Asian American, and 9.5% multi-racial or other (17.8% did not report their child's race/ethnicity).

Stimuli and procedure

The majority of children completed the assessments in two sessions, 15 to 20 minutes per session. Sessions were on average 9.17 days apart ($SD = 11.63$ days). In the first session, the experimenter assessed children's EF skills first, then vocabulary, literacy skills (again included to maintain consistency with data collection for a larger ongoing study of

preschool practices at the first site), and math achievement. In the second session, children completed a measure of their FAM ability followed by a measure of non-symbolic numerical magnitude discrimination skills. Some children completed all the assessments in one session lasting 30 to 45 minutes ($n = 31$).

FAM task

The FAM task was identical to the task from Study 1, except that we counterbalanced whether number or size trials were presented in the first or second block. We also counterbalanced the pairing of background (black or white) with trial type (number or size), leading to four versions of the task. Children were randomly assigned to complete one of the four versions.

Specifically, in the first block (6 pre-switch trials), children played either the size game or the number game. In the second block (6 post-switch trials), they played either the size game or the number game, whichever game they did not play in the pre-switch trials. Finally, in the final block (12 mixed trials), children switched between the two games. For example, if they saw a black background they played the size game and if they saw a white background they played the number game. The items, demonstration trials, practice trials, and reminder prompts were the same as in Study 1.

In our main analyses, we examined children's accuracy on post-switch trials and mixed trials combined. Analyzing both trial types together was done to replicate prior analyses but also for the same theoretical reasons we used this scoring method in the first study. Specifically, we are interested in children's ability to shift from attending to one dimension to another (post-switch trials) as well as to flexibly shift between them within the same task (mixed trials).

Numerical magnitude discrimination skill

We assessed children's non-symbolic numerical magnitude discrimination skills using the Panamath task (<http://panamath.org>; Halberda et al., 2008). Children were introduced to two puppets on the screen – Big Bird and Grover. Big Bird had a box of yellow dots on the left-hand side of the screen and Grover had a box of blue dots on the right-hand side of the screen. Children were asked to judge as quickly and accurately as possible which box had more dots and to respond by pressing the "A" or "L" keyboard key covered with correspondingly colored stickers. The number of dots in each box ranged from 5–21 dots. The experimenter initiated each trial by pressing the space bar and dot arrays were set to display for 2322 ms to prevent children from counting. Children completed one practice trial on paper and six practice trials on the computer. The experimenter gave feedback during the practice trials. The six practice trials on the computer were followed by 48 test trials. No feedback was given during the test trials. On $\frac{1}{3}$ of the test trials, the dots in the two boxes had the same average sizes so that the box with more numerous dots also had a larger cumulative surface area. On another $\frac{1}{3}$ of the test trials, the dots in the two boxes had the same cumulative surface area so that the box with more numerous dots had dots with a smaller average size. On the last $\frac{1}{3}$ of the test trials, the box with more numerous dots had dots with both a smaller average size and a smaller cumulative surface area. Percent of correct responses was calculated as our measure of children's numerical magnitude discrimination skills.

Math achievement

We assessed children's math achievement using the same measure as in Study 1, the Number Sense subtest of the Woodcock-Johnson IV Tests of Early Cognitive and Academic Development (ECAD) (Schrank et al., 2015). We used W scores in all analyses.

EF skills

We assessed EF skills using the same measure as in Study 1, the Minnesota Executive Function Scale (MEFS) via the MEFSTM App (Carlson & Zelazo, 2014). We used standar-dized scores in all analyses.

Vocabulary

We assessed language skills using the same measure as in Study 1, the Picture Vocabulary subtest of the Woodcock-Johnson IV ECAD (Schrank et al., 2015). We used W scores in all analyses.

Study 2 results

Descriptives and correlations

Descriptive statistics and sample sizes for all Study 2 variables are reported in Table 4. Participants' standardized scores indicated that they were similar to national norms on the MEFS ($M = 97.38$, $SD = 9.37$) and Picture Vocabulary ($M = 99.25$, $SD = 14.42$), and slightly below national norms in Number Sense ($M = 89.60$, $SD = 19.61$). Correlations between key study variables are reported in Table 5.

FAM task performance

In the FAM task, performance on Size trials, Number trials, and Mixed trials, by task order, is shown in Figure 4. We conducted a 2 (Order: number-first, size-first) x 3 (Block: pre-switch, post-switch, mixed) ANOVA on FAM task accuracy. We used the Greenhouse-Geisser correction for sphericity. There was a significant main effect of Block ($F(1.78, 260.3) = 57.95$, $p < .001$, $\eta_p^2 = .28$), a significant Block x Order interaction ($F(1.78, 260.3) = 19.50$, $p < .001$, $\eta_p^2 = .12$), and no significant main effect of Order ($F(1, 146) = 0.94$, $p = .333$, $\eta_p^2 = .01$).¹

To further understand the Block x Order interaction, we examined performance within each Order. As seen in Figure 4, children who saw size trials first significantly declined from pre-switch ($M = .94$, $SD = .16$) to post-switch ($M = .75$, $SD = .37$; $t(74) = 4.46$, $p < .001$, $d = .51$), and then got significantly worse again from post-switch to mixed trials ($M = .61$, $SD = .21$; $t(74) = 3.34$, $p = .001$, $d = .39$). In contrast, children who saw number trials first significantly *improved* from pre-switch ($M = .76$, $SD = .31$) to post-switch trials ($M = .89$, $SD = .23$; $t(72) = 3.13$, $p = .002$, $d = .37$), then got significantly worse from post-switch to

¹Adding Color (black background = size vs. white background = size) as a factor in the ANOVA did not change the significant main effects of Block or the significant Block x Order interaction. It did reveal a small and unexpected Color x Block interaction ($F(1.74, 250) = 3.40$, $p = .041$, $\eta_p^2 = .02$), such that when a white background was paired with size trials, participants did better in the pre-switch and post-switch blocks, and worse on the mixed block, than when the black background was paired with size trials. However, given the small and unexpected effect, we did not include Color in our main analyses.

Table 4. Study 2 descriptive statistics.

Variable	N	Min	Max	M	SD
MEFS (Standard Score)	157	61	131	97.38	9.37
Picture Vocabulary (W Score)	153	414	500	456.55	14.51
Number Sense (W Score)	149	362	462	413.06	22.43
Panamath Accuracy	140	0.35	1	.63	.16
FAM Post-switch/Mixed Trials Accuracy	148	0.28	1	.67	.19

Table 5. Study 2 correlations among variables.

Variable	1	2	3	4	5	6	7	8
1. MEFS Standard Score	-							
2. Picture Vocabulary W Score	.40***	-						
3. Fall Number Sense W Score	.48***	.72***	-					
4. Panamath Accuracy	.31***	.44***	.53***	-				
5. FAM Post-Switch/Mixed Trials Accuracy	.32***	.46***	.58***	.44***	-			
6. Age	-.11	.40***	.35***	.32***	.35***	-		
7. Gender	.11	.00	.01	-.03	.10	.10	-	
8. Family Income	.38***	.26*	.25*	.23*	.17	-.34***	-.11	-
9. Parent Education	.32***	.17	.19*	.15	.15	-.28**	-.02	.65***

* $p < .05$, ** $p < .01$, *** $p < .001$.

mixed trials ($M = .57$, $SD = .17$; $t(72) = 10.94$, $p < .001$, $d = 1.28$). This suggests that the initial rule switch from pre-switch to post-switch trials was less difficult when the switch was from number to size than from size to number. Mixed trials were the most difficult regardless of whether children saw size trials or number trials first.

We also examined whether performance differed across orders, within block. As seen in Figure 4, children scored significantly lower on pre-switch trials in the number-first order than the size-first order ($t(146) = 4.50$, $p < .001$, $d = .74$). Similarly, they performed significantly better on post-switch trials in the number-first than the size-first order ($t(146) = 2.89$, $p = .004$, $d = .48$). This shows evidence that number trials were significantly harder for children than size trials regardless of whether they saw number trials in the pre-switch or post-switch trials. However, there was no significant difference between the orders in the mixed trials ($t(146) = 1.40$, $p = .165$, $d = .23$).

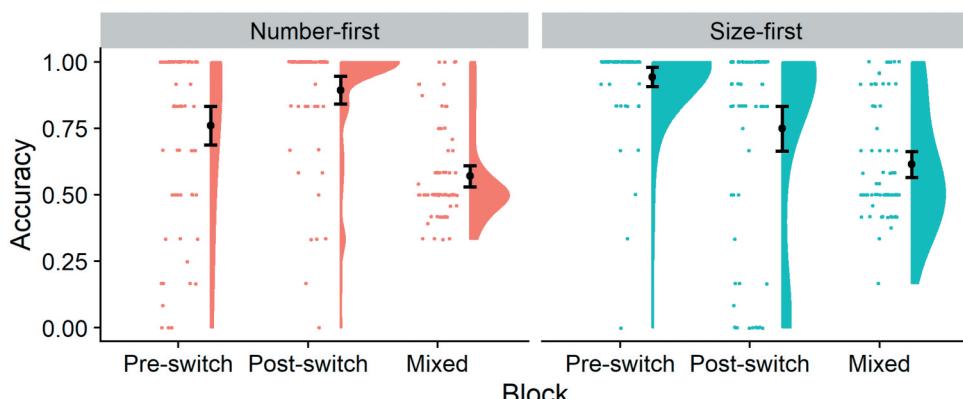


Figure 4. Study 2: children's accuracy on the FAM task by trial type and order. Note. Black bars show error bars ± 1 SE.

Table 6. Regression models predicting number sense W score.

	Model 1	Model 2	Model 3
	β	β	β
Age (months)	.14*	.12	.12
Gender	-.07	-.06	-.06
Picture Vocabulary W Score	.41***	.38***	.38***
MEFS Standard Score	.26***	.23***	.23***
FAM Post-switch/Mixed Trials Accuracy	.29***	.25***	.24*
Panamath Accuracy		.14*	.15*
FAM Order (Number vs. Size First)			-.05
FAM Order x FAM Accuracy			.01
R^2	.62	.64	.64
<i>F</i> -test	<i>F</i> (5, 124) = 40.60***	<i>F</i> (6, 123) = 35.67***	<i>F</i> (8, 121) = 26.43***

* $p < .05$, ** $p < .01$, *** $p < .001$.

We also ran the same models with income and education included as covariates. Due to list-wise deletion, this reduced the sample size to $n= 85$ participants. The pattern of results remained the same: FAM accuracy was a significant predictor of math achievement in Model 1 ($\beta = .25$, $p = .003$) and Model 2 ($\beta = .21$, $p = .011$), and the FAM Order x FAM Accuracy interaction in Model 3 was not significant ($p = .341$).

Regression analyses

For our main analyses, we ran a series of hierarchical regressions predicting children's math achievement, measured by the ECAD Number Sense (Table 6). In Model 1, we tested whether FAM performance, measured by post-switch/mixed trials accuracy, was associated with math achievement even after controlling for EF skills, vocabulary, age, and gender.² Analyzing post-switch trials and mixed trials accuracy together was done in order to measure children's ability to both shift from attending to one dimension to another as well as to flexibly shift between them. FAM performance was significantly associated with math achievement ($\beta = .29$, $p < .001$), even after controlling for vocabulary, age, and gender.

In Model 2, we tested whether FAM performance was still associated with math achievement even after adding Panamath accuracy as a control measure. Consistent with our hypothesis, even after controlling for Panamath accuracy, FAM post-switch/mixed trials accuracy significantly related to children's math achievement ($\beta = .25$, $p < .001$). Panamath accuracy also made a small but significant contribution in predicting math achievement ($\beta = .14$, $p = .030$).

In Model 3, we tested whether the significant relationship between FAM performance and math achievement would be moderated by FAM task order (i.e., whether children saw size trials or number trials first). We tested this by adding FAM Order and the interaction of FAM Order x FAM performance to the model. The interaction was not significantly different from zero ($\beta = .01$, $p = .949$). Therefore, whether children saw size trials or number trials first did not significantly change the relation between FAM performance and math achievement.

Study 2 discussion

In Study 2, we examined whether attending to numbers first, or sizes first, would impact children's performance on the FAM task or impact FAM's relationship to math

²Adding parents' education and income did not change the pattern of results, and they were not significant predictors of math achievement in any model.

achievement. Order did affect performance, such that attending to number was more challenging than attending to size, regardless of whether number trials were presented in the pre-switch or post-switch block. This suggests that in Study 1, the decline in performance from pre-switch to post-switch blocks could be attributable to the greater difficulty of number trials than size trials, rather than the difficulty of switching rules between blocks. However, performance on the mixed block of trials did not depend on the order of the first two blocks, and remained quite challenging for children. Importantly, we found that the relationship between FAM performance and math achievement was not moderated by block order. In other words, the relation between FAM post-switch/mixed block performance and children's math achievement was similar regardless of whether the post-switch block focused on size or on number. This suggests that the relation between FAM and math achievement was not driven by the number-focused block, but rather, by the challenge of switching focus between size and number. Finally, in Study 2, we showed that the relationship between FAM and math achievement held while controlling for acuity of the approximate number system.

General discussion

We assessed children's flexible attention to numerical and spatial magnitudes (FAM) using a newly developed task in two distinct studies to test three hypotheses. The first hypothesis, that flexibly shifting between *both* numerical *and* spatial magnitudes is a unique challenge in early childhood math, was supported. Specifically, we found that children's performance significantly declined when asked to flexibly shift back and forth between spatial and numerical magnitudes. Our second hypothesis was that children's FAM task performance would be related to both their EF skills and their math achievement. This hypothesis was supported in both zero-order correlations across both studies as well as within-time-point path analyses controlling for demographic variables in Study 1. The correlations between FAM ability, EF, and math achievement were similar to one another in magnitude (r_s between .32 and .40). Finally, our hypothesis that children's FAM ability would significantly predict their math achievement after controlling for EF skills, non-symbolic numerical discrimination acuity, and other covariates was supported.

Across two studies, preschoolers' flexible attention to numerical and spatial magnitudes (FAM) was consistently related to both their EF skills and math achievement. Importantly, FAM significantly predicted later math achievement, over and above prior math achievement and EF skills (Study 1). FAM was also related to children's contemporaneous math achievement even after controlling for both EF and approximate number system acuity (Study 2). These results are consistent with at least two other accounts of young children's numerical magnitude skills. First, the competing processes account (Clayton & Gilmore, 2015) suggests that the association between young children's non-symbolic numerical magnitude comparison task performance and their math achievement can, at least in part, be explained by their inhibitory control as it relates to attending to numerical magnitudes when in conflict with spatial magnitudes. Second, recent work by Wilkey et al. (2020) suggests that number-specific EF is a key factor explaining children's math difficulties, above and beyond domain-general EF skills. Our account is unique in that we argue that it is not just children's ability to attend to numerical quantity when conflicting

with spatial magnitudes, but also their ability to flexibly shift attention between numerical and spatial magnitudes, that is a key driver of math achievement in preschool. This hypothesis was supported by our exploratory follow-up analyses showing that it was the switch trials driving the effect, rather than only number trials.

Many recent studies in early math have been conducted in the context of non-symbolic numerical magnitude comparison tasks, which tap into the ability to estimate and compare large quantities quickly without exact counting. We specifically designed the FAM task to test children's flexible attention to numerical and spatial magnitudes outside of the limited scope of these commonly-used tasks, as we think that children's FAM ability may be applicable to a variety of early math tasks that require flexible attention to both numerical and spatial magnitudes. Therefore, we designed our task to emulate everyday math experiences. The FAM task includes set sizes both within and outside children's subitizing range, does not prohibit children from using counting as a strategy, and includes set sizes (1–10 items) typical of common number books. Our findings in Study 2 indicated that this FAM ability was a stronger predictor of children's math achievement than non-symbolic numerical discrimination acuity tasks (i.e., Panamath). If children's non-symbolic numerical discrimination acuity captures both EF and math skills, as some have argued (e.g., Clayton & Gilmore, 2015; Fuhs & McNeil, 2013; Szucs et al., 2013), including it as a covariate gives us additional evidence that there is a unique association between FAM task performance and math that is not otherwise captured by general EF abilities.

Both studies also revealed evidence that attention to number alone could not explain the relation of FAM to math achievement. In Study 1, exploratory analyses of the mixed block indicated that switch trials (trials on which the prior trial focused on the other dimension of magnitude) were a better predictor of math achievement than number-focused trials. These effects could not be explained by switch trials simply being more difficult than the non-switch trials (switch trials = 56% accuracy; non-switch trials = 56% accuracy). In Study 2, we showed that whether the post-switch trials focused on number or size, the relation of FAM performance (post-switch and mixed trials) to math achievement was similar in magnitude. Together, these findings converge to support our initial hypotheses, that flexibly switching between number and size is a key challenge and driver of math achievement at this age. These results suggest that more research is warranted to understand how and why children's ability to flexibly shift between magnitudes develops and what role it plays in specific math skills.

Interestingly, we found a bidirectional association between children's FAM ability and their math achievement in Study 1. Children's math skills in the fall of preschool also predicted their growth in FAM ability, with a standardized path coefficient larger than the path from fall FAM ability to children's growth in math achievement. We did not have a specific hypothesis guiding this exploration, but there are likely a few possible explanations. Formal math learning in preschool may increase children's attention to numerical magnitudes through activities involving counting and cardinality. Further, knowledge of relevant math language, including terms like "more", "fewer", "bigger", "smaller", "tall", and "wide" may increase children's ability to focus on spatial and numeric dimensions that may conflict with one another, increasing performance on the FAM task.

Relatedly, one might wonder if there were items on the WJ-IV ECAD Number Sense assessment that were very similar to those on the FAM task, which might explain the bidirectional findings as well as the non-significant association between EF skills and

math once FAM was entered into the longitudinal model in Study 1. However, the WJ-ECAD Number Sense does not contain any numerical or spatial magnitude comparison items similar to the FAM task. There were a few items that asked children to point to the tallest, shortest, largest, and smallest items among a single group of items. These items are designed to assess children's spatial math language and may also involve switching attention more generally; this could partly account for the association between children's fall math achievement and their growth in FAM task performance across the school year. It should be noted that the FAM task correlations across the time points in Study 1 was $r = .43$, which is somewhat small but consistent with temporal stability statistics on cognitive self-regulation measures (Lipsey et al., 2017). It will be important for future longitudinal studies to investigate methods for reducing noise in the measurement of FAM ability.

Limitations and future directions

Several important questions remain concerning FAM task performance and its relation to both EF skills and math achievement. The first question relates to the FAM task design. The current design cannot definitely rule out the possibility that children adopted a strategy where they flexibly shifted between choosing the larger items and smaller items, because the task was set up such that the more numerous object set always had smaller total surface area. There are two reasons we believe this strategy was unlikely to be adopted by children. First, children would need to first notice the incongruity between the dimensions, and then translate the instructions from the experimenter to choose which box has more items or which box has bigger items into an alternative strategy. This would require a sophisticated understanding of the relationship between the perceptual parameters of the items and the instructions. Although a child could learn such rules implicitly based on feedback across multiple trials, our design included feedback on a small number of practice trials. Second, this strategy was tested as a possible explanation for children's improvement across training in a prior experimental study of children's approximate number system acuity, which focused on identifying numerical magnitude comparisons while ignoring conflicting spatial information on some trials (Fuhs, Hornburg, et al., 2016). In that study, researchers found that children improved the most on trials that were incongruent with respect to spatial and numerical magnitudes. To rule out the possibility that the improvement was due to children learning a strategy to pick the box with smaller items and ignoring numerosity altogether, they included check trials at posttest in which the side with more numerous items also had greater spatial magnitude at an equal ratio (e.g., 1:2 ratio for numerosity would mean that the objects on the more numerous side would be larger in size than the less numerous box by a ratio of 2:1). If children used a "pick the side with smaller items" strategy, they would have responded incorrectly on these trials, whereas if they used a correct strategy, the answer to these items would be very obvious. The results indicated no effect of condition on the check trials, indicating that using this inaccurate strategy was not a likely explanation for the findings. Further, children in that study had participated in a multi-week intervention that included eight sessions of exposure to non-symbolic numerical magnitude discrimination prior to the posttest check trials, but they still did not show evidence of adopting the strategy to choose the smaller side on posttest trials. Thus, we argue that it is even more unlikely in our one-time assessment at both fall and spring that children would learn and

implement this strategy in 18 trials. Nonetheless, it will be important in future studies to test this possibility by incorporating check trials into the task.

The second question for future research is whether FAM task performance can be thought of as a separate construct from domain-general EF skills or as a domain-specific EF task. In other words, we do not yet have empirical evidence that FAM task performance either does or does not load onto a common EF factor. We found moderate correlations between the FAM task and our EF skills task, but because we used a single measure of EF skills, we cannot test if FAM task performance loads onto a separate latent factor from general EF skills. It may be that FAM task performance can be thought of as number-specific EF (similar to Wilkey et al., 2020), such that FAM task performance and domain-general EF skills could potentially load onto a common latent factor. Even if FAM can be considered a type of number-specific EF, the results of this study still suggest potential benefits in studying FAM task performance, specifically its development and malleability. If EF skills generally are a strong predictor of math achievement, and FAM task performance even more so, it suggests a potentially powerful target for intervention that could also affect children's EF skills, which we discuss below.

Third, given the limitations of correlational research, it will be important in future studies to test the causal nature of these predictions using tightly-controlled experiments. Researchers have tested interventions aimed at increasing children's non-symbolic numerical magnitude discrimination skills as well as their EF skills as means for increasing children's math achievement. Experimental evidence is mixed for the impact of training in non-symbolic numerical magnitude discrimination skills, as short-term non-symbolic approximate arithmetic training yielded significant improvements in math skills (Hyde, Khanum, & Spelke, 2014) while a longer-term pure non-symbolic magnitude discrimination training study did not (O'Rear, Fuhs, McNeil, & Silla, 2015). Fuhs and colleagues (Fuhs, McNeil, Kelley, O'Rear, & Villano, 2016) found that children who completed a non-symbolic magnitude discrimination training intervention only improved their performance on training trials in which there was incongruity between numerical and spatial magnitudes. This suggests that the specific ability to resolve incongruencies between numerical and spatial magnitudes is malleable.

Many early EF skills interventions have also been evaluated as a means for increasing children's math achievement. The results of these studies have also been mixed, showing little transfer or long-term effects. For example, computerized working memory training produced short-term boosts in cognitive skills that did not transfer to other untrained skills or academic achievement (Melby-Lervåg & Hulme, 2013), although a recent study did find evidence for transfer from working memory training to numerical skills (Ramani, Jaeggi, Daubert, & Buschkuhl, 2017). EF skills training might be most effective in promoting math achievement if trained in a targeted way within the domain of early math skills rather than training these skills in an abstract way outside of specific math content. One way to do this could be to focus on interventions that help children develop their FAM ability.

Conclusion

The aim of these two studies was to test specific hypotheses stemming from the FAM account of young children's understanding of numerical and spatial magnitudes. As predicted, preschool-aged children tended to experience difficulty in shifting attention between

these two types of magnitudes. Children's FAM ability related to their EF, non-symbolic numerical discrimination skills, and math skills, and FAM ability was a significant predictor of children's math achievement above and beyond other covariates. These results suggest that early math activities that incorporate FAM could potentially be beneficial to young children's development of math skills, though experimental studies are needed to establish the malleability of children's FAM ability and to test if it is causally related to math achievement.

Acknowledgments

The authors wish to thank the children and preschools who participated as well as a team of dedicated research assistants without whom this research would not be possible.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by a STEM Catalyst grant from the University of Dayton to Fuhs and NSF CAREER DRL-1452000 and James S. McDonnell Foundation Award #220020546 to Gunderson.

Data Availability Statement

Datasets, syntax, output, and stimuli for these studies are available on the Open Science Framework at <https://osf.io/fxms2/>.

Open scholarship



This article has earned the Center for Open Science badges for Open Data and Open Materials through Open Practices Disclosure. The data and materials are openly accessible at <https://doi.org/10.17605/OSF.IO/ZD3B9>.

References

- Bonny, J. W., & Lourenco, S. F. (2015). Individual differences in children's approximations of area correlate with competence in basic geometry. *Learning and Individual Differences*, 44, 16–24. doi:[10.1016/j.lindif.2015.11.001](https://doi.org/10.1016/j.lindif.2015.11.001)
- Boyer, T. W., & Levine, S. C. (2015). Prompting children to reason proportionally: Processing discrete units as continuous amounts. *Developmental Psychology*, 51(5), 615–620. doi:[10.1037/a0039010](https://doi.org/10.1037/a0039010)
- Carlson, S. M., & Zelazo, P. D. (2014). *Minnesota executive function scale: Test manual*. St. Paul, MN: Reflection Sciences, Inc.
- Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, 148, 163–172. doi:[10.1016/j.actpsy.2014.01.016](https://doi.org/10.1016/j.actpsy.2014.01.016)

- Clayton, S., & Gilmore, C. (2015). Inhibition in dot comparison tasks. *ZDM*, 47(5), 759–770. doi:[10.1007/s11858-014-0655-2](https://doi.org/10.1007/s11858-014-0655-2)
- Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, 10(5), 408–411. doi:[10.1111/1467-9280.00177](https://doi.org/10.1111/1467-9280.00177)
- Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, 3(2), 63–68. doi:[10.1016/j.tine.2013.12.001](https://doi.org/10.1016/j.tine.2013.12.001)
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168. doi:[10.1146/annurev-psych-113011-143750](https://doi.org/10.1146/annurev-psych-113011-143750)
- Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, 123, 53–72. doi:[10.1016/j.jecp.2014.01.013](https://doi.org/10.1016/j.jecp.2014.01.013)
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8(7), 307–314. doi:[10.1016/j.tics.2004.05.002](https://doi.org/10.1016/j.tics.2004.05.002)
- 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. doi:[10.1037/dev0000145](https://doi.org/10.1037/dev0000145)
- Fuhs, M. W., & McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: Contributions of inhibitory control. *Developmental Science*, 16(1), 136–148. doi:[10.1111/desc.12013](https://doi.org/10.1111/desc.12013)
- Fuhs, M. W., McNeil, N. M., Kelley, K., O'Rear, C., & Villano, M. (2016). The role of non-numerical stimulus features in approximate number system training in preschoolers from low-income homes. *Journal of Cognition and Development*, 17(5), 737–764. doi:[10.1080/15248372.2015.1105228](https://doi.org/10.1080/15248372.2015.1105228)
- 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), 16981709. doi:[10.1037/a0036633](https://doi.org/10.1037/a0036633)
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, 134, 31–60. doi:[10.1037/0033-2909.134.1.31](https://doi.org/10.1037/0033-2909.134.1.31)
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., ... Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *Plos One*, 8(6), e67374. doi:[10.1371/journal.pone.0067374](https://doi.org/10.1371/journal.pone.0067374)
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, 44(5), 1457–1465. doi:[10.1037/a0012682](https://doi.org/10.1037/a0012682)
- Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, 455(7213), 665–669. doi:[10.1038/nature07246](https://doi.org/10.1038/nature07246)
- Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition*, 131(1), 92–107. doi:[10.1016/j.cognition.2013.12.007](https://doi.org/10.1016/j.cognition.2013.12.007)
- 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. doi:[10.1002/ddrr.46](https://doi.org/10.1002/ddrr.46)
- Lee, K., Bull, R., & Ho, R. H. H. (2013). Developmental changes in executive functioning. *Child Development*, 84, 1933–1953. doi:[10.1111/cdev.12096](https://doi.org/10.1111/cdev.12096)
- Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From "sense of number" to "sense of magnitude": The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, 40. doi:[10.1017/s0140525x16000960](https://doi.org/10.1017/s0140525x16000960)
- Levine, S. C., Suriyakham, L. W., Rowe, M. L., Huttenlocher, J., & Gunderson, E. A. (2010). What counts in the development of young children's number knowledge? *Developmental Psychology*, 46(5), 1309–1319. doi:[10.1037/a0019671](https://doi.org/10.1037/a0019671)
- Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, 14(6), 1292–1300. doi:[10.1111/j.1467-7687.2011.01080.x](https://doi.org/10.1111/j.1467-7687.2011.01080.x)

- Lipsey, M. W., Nesbitt, K. T., Farran, D. C., Dong, N., Fuhs, M. W., & Wilson, S. J. (2017). Learning-related cognitive self-regulation measures for prekindergarten children: A comparative evaluation of the educational relevance of selected measures. *Journal of Educational Psychology*, 109, 1084–1102. doi:[10.1037/edu0000203](https://doi.org/10.1037/edu0000203)
- Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences of the United States of America*, 109(46), 18737–18742. doi:[10.1073/pnas.1207212109](https://doi.org/10.1073/pnas.1207212109)
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49(2), 270–291. doi:[10.1037/a0028228](https://doi.org/10.1037/a0028228)
- Miller, M. R., Rittle-Johnson, B., Loehr, A. M., & Fyfe, E. R. (2016). The influence of relational knowledge and executive function on preschoolers' repeating pattern knowledge. *Journal of Cognition and Development*, 17, 85–104. doi:[10.1080/15248372.2015.1023307](https://doi.org/10.1080/15248372.2015.1023307)
- Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin*, 128(2), 278. doi:[10.1037/0033-2909.128.2.278](https://doi.org/10.1037/0033-2909.128.2.278)
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21, 8–14. doi:[10.1177/0963721411429458](https://doi.org/10.1177/0963721411429458)
- Negen, J., & Sarnecka, B. W. (2015). Is there really a link between exact-number knowledge and approximate number system acuity in young children? *British Journal of Developmental Psychology*, 33(1), 92–105. doi:[10.1111/bjdp.12071](https://doi.org/10.1111/bjdp.12071)
- 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. doi:[10.1016/j.ecresq.2018.02.003](https://doi.org/10.1016/j.ecresq.2018.02.003)
- Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(6), 491–505. doi:[10.1002/wcs.1369](https://doi.org/10.1002/wcs.1369)
- O'Rear, C. D., Fuhs, M. W., McNeil, N. M., & Silla, E. (2015). *Approximate number system training in children from low-income homes*. Paper presented at the Biennial Meeting of the Cognitive Development Society, Columbus, OH.
- Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, 41, 1–20. doi:[10.1016/S0896-6273\(04\)00107-2](https://doi.org/10.1016/S0896-6273(04)00107-2)
- Prager, E. O., Sera, M. D., & Carlson, S. M. (2016). Executive function and magnitude skills in preschool children. *Journal of Experimental Child Psychology*, 147, 126–139. doi:[10.1016/j.jecp.2016.01.002](https://doi.org/10.1016/j.jecp.2016.01.002)
- Ramani, G. B., Jaeggi, S. M., Daubert, E. N., & Buschkuhl, M. (2017). Domain-specific and domain-general training to improve kindergarten children's mathematics. *Journal of Numerical Cognition*, 3(2), 468–495. doi:[10.5964/jnc.v3i2.31](https://doi.org/10.5964/jnc.v3i2.31)
- 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. doi:[10.1111/j.1467-8624.2007.01131.x](https://doi.org/10.1111/j.1467-8624.2007.01131.x)
- Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology*, 30(2), 344–357. doi:[10.1111/j.2044-835x.2011.02048.x](https://doi.org/10.1111/j.2044-835x.2011.02048.x)
- Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergarteners. *Quarterly Journal of Experimental Psychology*, 67(2), 271–280. doi:[10.1080/17470218.2013.803581](https://doi.org/10.1080/17470218.2013.803581)
- Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R., & McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology*, 109, 1120–1140. doi:[10.1037/edu0000193](https://doi.org/10.1037/edu0000193)

- Schmitt, S. A., McClelland, M. M., Tominey, S. L., & Acock, A. C. (2015). Strengthening school readiness for Head Start children: Evaluation of a self-regulation intervention. *Early Childhood Research Quarterly*, 30, 20–31. doi:10.1016/j.ecresq.2014.08.001
- 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, e12372. doi:10.1111/desc.12372
- Schrank, F. A., & Dailey, D. (2014, 2015). *Woodcock-Johnson online scoring and reporting system* [Online format]. Rolling Meadows, IL: Riverside.
- Schrank, F. A., McGrew, K. S., & Mather, N. (2015). *Woodcock-Johnson IV tests of early cognitive and academic development*. Rolling Meadows, IL: Riverside.
- Sokolowski, H. M., Fias, W., Ononye, C. B., & Ansari, D. (2017). Are numbers grounded in a general magnitude processing system? A functional neuroimaging meta-analysis. *Neuropsychologia*, 105, 50–69. doi:10.1016/j.neuropsychologia.2017.01.019
- Solomon, T. L., Vasilyeva, M., Huttenlocher, J., & Levine, S. C. (2015). Minding the gap: Children's difficulty conceptualizing spatial intervals as linear measurement units. *Developmental Psychology*, 51(11), 1564–1573. doi:10.1037/a0039707
- Starr, A., DeWind, N. K., & Brannon, E. M. (2017). The contributions of numerical acuity and non-numerical stimulus features to the development of the number sense and symbolic math achievement. *Cognition*, 168, 222–233. doi:10.1016/j.cognition.2017.07.004
- Szucs, D., Nobes, A., Devine, A., Gabriel, F. C., & Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Frontiers in Psychology*, 4, 444. doi:10.3389/fpsyg.2013.00444
- Tudusciuc, O., & Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proceedings of the National Academy of Sciences*, 104(36), 14513–14518. doi:10.1073/pnas.0705495104
- Verdine, B. N., Irwin, C. M., Golinkoff, R. M., & Hirsch-Pasek, K. (2014). Contributions of executive function and spatial skills to preschool mathematics achievement. *Journal of Experimental Child Psychology*, 126, 37–51. doi:10.1016/j.jecp.2014.02.012
- Ward, J. M., Mazzocco, M. M., Bock, A. M., & Prokes, N. A. (2017). Are content and structural features of counting books aligned with research on numeracy development? *Early Childhood Research Quarterly*, 39, 47–63. doi:10.1016/j.ecresq.2016.10.002
- Wendling, B. J., Mather, N., LaForte, E. M., McGrew, K. S., & Schrank, F. A. (2015). *Comprehensive manual: Woodcock-Johnson IV tests of early cognitive and academic development*. Rolling Meadows, IL: Riverside.
- Wilkey, E. D., Pollack, C., & Price, G. R. (2020). Dyscalculia and typical math achievement are associated with individual differences in number-specific executive function. *Child Development*, 91, 596–619. doi:10.1111/cdev.13194
- Wilkey, E. D., & Price, G. R. (2019). Attention to number: The convergence of numerical magnitude processing, attention, and mathematics in the inferior frontal gyrus. *Human Brain Mapping*, 40(3), 928–943. doi:10.1002/hbm.24422
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1(1), 297–301. doi:10.1038/nprot.2006.46