Give Yourself a Hand: The Role of Gesture and Working Memory in Preschoolers' Numerical Knowledge

Raychel Gordon[1], Nicole R. Scalise[1,2], Geetha B. Ramani[1]

[1] Department of Human Development and Quantitative Methodology, University of Maryland, College Park, College Park, MD, 20742, USA

[2] Department of Cognitive Sciences, University of California, Irvine, Irvine, CA, 92617, USA

Final accepted version of manuscript February 2, 2021

In Press at *Journal of Experimental Child Psychology*

Abstract

Hand gestures can be beneficial in math contexts to reduce the user's cognitive load by supporting domain-general abilities, such as working memory. Although prior work has shown a strong relation between young children's early math performance and general cognitive abilities, it is important to consider how children's working memory ability may relate to their use of spontaneous gesture, as well as their math-specific abilities. The present study examined how preschool aged children's gesture use and working memory relates to their performance on an age-appropriate math task. Head Start preschoolers (n=81) were videotaped completing a modified version of the Give-N task to measure their cardinality understanding. Children also completed a forward Word Span task and a computerized Corsi-Block task to assess their working memory. The results showed that children's spontaneous gesture use and working memory was related to their performance on the cardinality task. However, children's gestures were not significantly related to working memory after controlling for age. Findings suggest that young children from low-income background use gestures during math contexts in similar ways to preschoolers from higher-income backgrounds.

Keywords: Gesture, Math, Cardinality, Head Start, Preschool

Research Highlights

- Preschool children's gestures, working memory, and cardinality ability were studied
- Children's gestures related to their cardinality knowledge
- Children's working memory related to their cardinality knowledge, but not gestures
- Low-income children use gestures in similar ways to higher-income children

Give Yourself a Hand: The Role of Gesture and Working Memory in Preschoolers' Numerical Knowledge

The factors that impact children's early mathematical learning are critical to understand, as these early abilities are strongly linked to later math achievement (Claessens & Engel, 2013; Geary, Hoard, Nugent, & Bailey, 2013; Watts, Duncan, Siegler, & Davis-Kean, 2014). One such factor is working memory (WM), which is a person's domain-general ability to hold information in their mind while simultaneously carrying out a mental process (Baddeley & Hitch, 1974). Prior work has shown a consistent link between children's domain-general WM abilities and early math success (see Clements, Sarama, & Germeroth, 2016 for a recent review).

Another factor that impacts children's math learning is their use of and exposure to nonverbal communication methods, such as hand gestures. Gestures play a role in children's acquisition of novel math concepts (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007). Specifically, gestures are thought to facilitate math learning through a reduction of cognitive demands (Cook, Yip, & Goldin-Meadow, 2012; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001) and direction of attention (Wakefield, Novack, Congdon, Franconeri, & Goldin-Meadow, 2018). Thus, when children use gestures during math instruction, they can recruit specific cognitive resources that may not otherwise be readily available. However, less is known about the combined relation between children's gesture use, WM, and mathematics performance.

The current work examines how preschool-aged children's gesture use and WM relate to their performance on a cardinality task. While prior work has focused on each of these relations independently (i.e., gesture and WM, WM and math, gesture and math separately), we sought to explore potential interdependencies between them. Additionally, we are interested in examining these skills with a sample of preschool-aged children given the concurrent and rapid development of both math-related knowledge and executive functions.

Gestures in Mathematical Contexts

Gestures are bodily movements involving the use of hands, which appear spontaneously alongside and in complement to our speech. Gestures are theorized to play a role in cognitive development because they are useful during communication for both the listener and the speaker (Goldin-Meadow & Beilock, 2010; Goldin-Meadow, Cook, & Mitchell, 2009). This has important implications when considering how and why a child may use gestures within a math environment. Critically, children's gestures can display implicit information that does not otherwise appear in speech. (Broaders & Goldin-Meadow, 2010; Goldin-Meadow, Wein, & Chang, 1992; Iverson & Goldin-Meadow, 2005; McNeill, 1992). Indeed, children who are instructed to gesture during difficult math tasks have shown evidence of new and correct problem-solving strategies in their gestures (Broaders et al., 2007). Thus, children's gestures have the capacity to both show and scaffold children's own math learning.

Prior research has focused on the role of gesture in mathematical contexts primarily for school-age children. These learning environments typically involve processing both visual and auditory information, which can be aided by gestures (Goldin-Meadow, Kim, & Singer, 1999). Gestures may also be important in early childhood when children are learning foundational numerical concepts, such as counting and cardinality. The cardinality principle is the rule that the last word in a correctly recited count-list is representative of the number of items in the set (Gelman & Gallistel, 1978). During this process of learning to count, children first learn the meaning of the numerosity "one", followed by "two", and so on (Le Corre, Van de Walle, Brannon, & Carey, 2006; Sarnecka & Gelman, 2004). Children's understanding of numbers can be assessed through how and when they use the cardinality principle within the Give-a-number task (e.g., Give-N; Wynn 1990, 1992). During this task, children are asked to create different set sizes of objects. The highest numeric set size that children are able to create reliably provides information regarding their current level of understanding for cardinality and counting principles.

Counting and cardinality tasks are a prime opportunity for examining children's gestures. For example, children's use of gesture during cardinality tasks, in conjunction with their speech, can be indicative of their readiness to learn new mathematical information (Gibson et al., 2019). In particular, children appear to use gestures during cardinality tasks when they are asked to create set sizes that are at or just above the quantity they are able to produce (Authors, 2019). In other words, children may be using gestures during counting to help ease their cognitive load by tracking objects they have already counted. Specifically, gestures provide an external link between the verbal count list and the counted objects, in a sequential manner. This practice has led to greater counting accuracy when used by children, as they individuate and tag each item as they count (Alibali & DiRusso, 1999). Furthermore, there is a positive relation between children's number knowledge during a counting task and their use of pointing and counting gestures, such that children with less knowledge about cardinality engage in less serial pointing compared to those with a better understanding of cardinality (Le-Corre & Carey, 2007).

Recent work has examined this trend with preschool children's performance on the standard, titrated Give-N task (Authors, 2019). Children's cardinality ability, but not their age, was positively related to their use of spontaneous gestures on this task. Furthermore, children in this study who had not yet mastered the cardinality principle used more gestures on specific parts of the Give-N tasks which involved creating sets objects that they had just learned or were in the process of learning. In other words, children who were in the process of learning how to create sets of four objects could reliably create sets of 2 and 3 objects without the help of gesture but tended to use more gestures to create a set of 4 objects. This suggests that the spontaneous gestures young children employ while learning about cardinality are dynamic and change with their underlying knowledge of the task. However, it is an open question as to why these patterns may emerge. In particular, this study did not include measures of more domain-general abilities, such as WM, which may account for additional variation in both the types of gestures, as well as the resulting score on the cardinality task.

Working Memory and Early Math Abilities

While prior work has shown that spontaneous gesture use relates to children's knowledge, it is critical to also consider how differences in children's domain-general abilities such as their WM may impact both their gesture use and math performance. Early math performance is highly related to general cognitive abilities, such as WM (see Bull & Espy, 2006 for a review). WM is limited in capacity (e.g., Cowan, 2001; Miller, 1956), and thus different tasks can be conceptualized as having differing amounts of cognitive "load", such that a task which requires more simultaneous memory and processing would have a higher load (Cognitive Load Theory; Sweller, 1988; Sweller, Van Merriënboer, & Paas, 1998). In addition to the variation in potential cognitive load, individuals vary in their personal WM capacities. In other words, the more WM capacity an individual has, the better their performance on cognitively difficult tasks (Engle, 2002), including mathematics tasks. For example, prior work shows that children with higher WM capacities have higher accuracy in solving arithmetic word problems (Grades K-3; LeBlanc & Weber-Russell, 1996). Furthermore, young children's WM strongly relates to their subsequent scores on standardized math tests (Monette, Bigras, & Guay, 2011) and is an important predictor of their early number skills (Bull & Lee, 2014; Kolkman, Hoijtink, Kroesbergen, & Leseman, 2013). Thus, children with lower WM abilities may have more difficulties learning about math, such as having difficulty remembering and carrying out instructions, monitoring their own progress in a task, or remembering to use particular strategies.

Gestures and Working Memory in Mathematical Contexts

Given the relation between gesture and math, as well as WM and math, it is important to consider how gestures and WM may interact. Gestures appear to both impact the demand on WM, as well as have differential benefits depending on the user's WM ability level (Alibali & DiRusso, 1999; Goldin-Meadow, 2001; Iverson & Goldin-Meadow, 1998). For example, gesture can lighten a speaker's cognitive load during problem solving and free up potential WM resources (Goldin-Meadow et al., 2001; Ping & Goldin-Meadow, 2010). Gestures also occur more frequently when task demands are high (Chu & Kita, 2011). While prior work has focused on the relation between gesture and WM, in the current study we are interested in how an individual's WM and their use of gesture may interact in a mathematical environment.

Gestures' positive impact on WM load has been found in both children and adults during math-related tasks (Goldin-Meadow et al, 2001) and a broader array of contexts (Wagner et al., 2004). Furthermore, the specific *type* of gesture that is used is relevant in its ability to reduce WM load. Cook and colleagues (2012) asked adult participants to solve math problems, remember a span, and either gesture (a meaningful movement), move both hands in circles (a meaningless movement), or not move their hands while explaining their solution to the math problems. Participants who used their natural gestures remembered more of the span than those who engaged in meaningless movements (moving their hands in circles). These findings indicate that meaningful gestures to the task can lighten the speaker's overall WM load during math tasks.

Evidence from studies with adults suggests that there are differential benefits from using gestures during a math task depending on an individual's WM ability. Specifically, Marstaller & Burianová (2013) showed adults a numerical equation on a screen and were asked to judge whether the solution to a mathematical equation was correct. After receiving feedback, they were shown a series of random letters before being asked to explain their prior judgment. While providing their explanation they were either asked to gesture to the screen or not. Lastly, they were asked to recall the letters. Results indicated that individual differences in participants' WM capacity was related to whether gestures would benefit their WM load. Specifically, the instruction to use gestures had a significant, beneficial effect on WM performance but only for those adults who had a low WM capacity. This suggests that gestures can assist and reduce WM load on a particular task, but that it is critical to understand how an individual's WM capacity may impact both their task performance and their use of gesture. However, less is known about the interplay between WM load and capacity for children, and whether these may interact with children's gesture use and math abilities.

A separate, but related literature suggests an interdependency between gestures, WM, and math performance exists for young children. Consider again the example of a child learning how to count and the underlying required knowledge; they need to hold information in their WM related to cardinality and ordinality (the ordered relation between each number), while simultaneously producing a verbal count list and tracking the objects visually to see if there are more objects that need to be counted. Gesture helps overcome some of the burden of these counting procedures by facilitating a more direct, external representation of the information. Gesture links the physical objects in space to their more abstract verbal count list (Alibali & DiRusso 1999). Indeed, prior empirical work supports the idea that representing information externally through gesture lightens the load on the individual's WM (Kirsh, 1995; Kirsh & Maglio, 1994). This in turn frees up WM resources to complete the task and produce the correct number of objects. In summary, WM has been shown to relate to math performance, gesture has been shown to assist with math performance, and gesture allows for a reduction of WM demand. However, prior empirical work has not examined this dynamic relation in young children in the context of early math understanding.

The Current Study

We begin to address this gap by investigating the relations between preschoolers' number knowledge, WM, and spontaneous gesture use. In the current study, we coded preschoolers' gestures during the Give-N task to explore how individual differences in gesture use relate to the link between their WM ability and their performance on the cardinality task. In particular, we sought to address two primary questions in relation to children's gesture use, WM, and their number knowledge.

First, we investigated the relations between children's WM ability, cardinality knowledge, and gesture use during the Give-N task. Specifically, we expected to find separate, positive relations between children's gesture use and Give-N score, WM and Give-N score, as well as WM and gesture use. In particular, we expected that children's point and count gestures, the most task-relevant and conceptually useful gesture, would be the crux of each gesture relation. Second, we examined whether children's gestures affected the relation between their WM and cardinality knowledge. Here, we predicted that children's spontaneous, unprompted gestures would moderate the relation between children's WM ability and their cardinality performance on the Give-N task. Lastly, we conducted an exploratory, post-hoc analysis to consider whether children's WM capacities may impact their use of spontaneous gesture and performance in the Give-N task. We discuss the implications of our findings, including a discussion of how the results from our low-income sample and modified Give-N methods may compare to previous work from a higher-income sample with a standard Give-N measure (Authors, 2019).

Methods

Participants

Participants were 81 preschoolers, ranging in age from 3.40 years old to 5.67 years old (*M*age=4.75; 60% girls). One additional participant was recruited but was not included in the final sample as they were unable to complete any of the tasks because of limited language production. An a priori power analysis was performed to determine the appropriate sample size via G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). The analysis including 3 predictor variables with a medium effect size, $\alpha = .05$ and $1 - \beta$ power of 0.85 yielded a projected sample size of *N*=87. However, data collection had to be discontinued unexpectedly in March 2020 due to the global pandemic resulting in the final sample of $n = 81$. The majority of the participants ($n = 70$) were recruited as part of a larger study investigating children's early numerical and executive functioning abilities (Authors, Under Review). The remaining participants ($n = 11$) were recruited to participate in the present study only.

Recruitment took place at four Head Start centers in a mid-Atlantic metropolitan area, which is a federally funded program for families whose incomes are below the federal poverty guidelines (annual household income of $25,750 or less for a family of four in the year these data were collected). Eleven parents did not complete the parent survey; of the remaining 70 participants, the race of the sample was 54% African American, 10% Caucasian, 9% Asian or Pacific Islander, 1% Native American or Alaskan, and 10% biracial or multiracial, with 16% of survey respondents not providing a response to this question. The ethnicity of the responding sample was 26% Hispanic or Latino, 71% not Hispanic or Latino with 3% of survey respondents not providing a response to this question. The parental education of survey respondents was 3% completing less than high school, 4% completing some high school, 24% completing a high school diploma or GED, 24% completing some college coursework or vocational training, 10% completing 2-year college degree (Associates), 31% completing a 4-year college degree, and 3% completing a postgraduate or professional degree (MA, PhD, MD, JD).

Procedure

Participants were seen for one visit at the children's school for approximately 15 minutes. Participants sat either next to the experimenter (White female) on the floor or across a child-sized table. The visit was video recorded. The experimenter administered a battery of three tasks in English: two tasks to assess WM and one to assess cardinality. These tasks were administered in the same order, as described below.

Working Memory

Children completed a Forward Word span and a Forward Corsi-Block task as measures of their WM. These two tasks were collapsed into a composite WM score as detailed in the results section. Both WM tasks were scored live by the experimenter. Later, a separate research assistant verified the live scores by watching each video, then entering the score into the data sheet.

Forward Word Span. Children were read a sequence of color words at approximately a rate of one word per second and were asked to repeat those words back to the experimenter in the same order (adapted from Müller et al., 2012; forward span tasks presented with "good retest reliability" with preschoolers, ICC = .56, $p < 0.05$). The number of color words within a trial ranged from two to seven, with two trials at each span level (i.e., two trials of two, two trials of three). The first two trials were practice trials with a string of two color words (e.g. green, blue) with feedback. The experimenter ended the task if the participant incorrectly answered both trials of a particular span. The dependent measure was the highest span that they were able to recall correctly.

Figure 1 Still image from adapted Corsi-Block task

Forward Corsi-Block. Children were presented with a tablet version of the forward Corsi-Block task (Authors, 2019; Adapted from Corsi, 1972). Previous research that has used similar forward Corsi block tapping tasks with preschoolers reported high reliability (e.g., a non-tablet version shows a high test-retest reliability, using Pearson correlations, of r = .83 with preschool aged children; Alloway et al., 2006). The tablet (LENOVO® Tablet with a screen size of 25.7 cm, measured diagonally) displayed a picture of a pond, with an animated frog "jumping" onto different lily pads (Figure 1). Children were instructed to tap on the lily pads in the same order that the frog jumped on them. The task began with a two-span practice trial with corrective feedback (happy face or sad face). If the child got the first practice trial wrong, they were given another practice trial with additional corrective feedback. If the child got the practice trial right, they moved onto test trials with no feedback. The trials increased in span size each time the child got two of the same span size correct. For each trial, children received 1 point for every lily pad they accurately recalled (i.e., 2 points for a 2 span, 3 points for a 3 span, and so on). This allows for partial scores on longer trials that are more difficult for a child to remember the whole span, rather than using the less variable measure of the longest span they could remember with complete accuracy. Thus, the dependent measure was the total points they achieved.

Cardinality

Modified Give-N. Children were asked to place sets of ducks from a pile of 12 ducks into a blue basket or "pond" (adapted from Krajcsi, Fintor, & Hodossy, 2018[1]). The experimenter asked for sets between 1-8 in a predetermined randomized order generated from a random number generator prior to the start of data collection; 1, 5, 3, 8, 4, 7, 2, 6. Children were asked for each set size three times in the same randomized list order, for a total of 24 trials regardless of their response. The alpha coefficient for the 24 trials is .93, suggesting high internal consistency for this task. Accuracy for each trial could be assessed as a proportion of trials correct (either 0/3, ⅓, ⅔, or 3/3). The dependent measure was the child's knower level which was defined for the purposes of this study as when all trials including and below a particular set size is above chance (⅔ or 3/3), and the trial immediately after in the count list is below chance (either 0/3 or ⅓). This task was scored live, trial by trial, by the experimenter. Later, a separate research assistant verified the accuracy of each trial watching each video, then entering the score into the data sheet.

Transcription and Coding

All speech and gestures from the videotaped Modified Give-N task were transcribed by research assistants trained to transcribe reliably using the CHAT conventions of the Child

[1] The portion of our sample who were recruited for a larger study (n=70; Authors, Under Review) also completed the standard Give-N assessment in a separate visit. To compare the similarity between these measures, all children who received a Knower Level (KL) of 7 or 8 in the adapted version were re-coded as having a KL of 6, to maintain consistency in KL assignment for the standard measure. Next, Pearson's product-moment correlation was run between the re-coded adapted KL scores and standard KL scores, showing a strong positive relation, $r(68) = .828$, $p < .01$. This indicates that the adapted measure provides a consistent measure of cardinality knowledge compared to the standard measure, with the added benefit of controlling for the total number of trials children received.

Language Data Exchange System (CHILDES; MacWhinney, 2000). Transcription reliability was assessed by having a second reliable coder provided verification and agreement on the speech and gestures decisions.

Measures

Children's Gestures. Each child's transcript was divided into sections based on the numerical trial and then coded for specific behaviors from three categories. The primary behavior of interest was Math Specific Gestures (Table 1), divided into Pointing and Counting gestures, and Other Math Gestures specifically, Fingers Held Up, and Magnitude (adapted from Authors, 2019). All transcripts were coded for the behaviors of interest by one primary coder. A second, reliable coder coded 21% of the transcripts for the same behaviors of interest, with an inter-rater reliability of 82.35% for Point and Count gestures, and 100% for the other gestures.

Table 1

Math Specific Gesture Definitions and Behavioral Examples

Gesture Type	Definition	Example
Point and Count	Uses their finger(s) or hand to indicate objects while verbally producing a count list.	Experimenter: "Is that three ducks?" Child: "One, two, three." *Uses pointer finger to point to the 1st duck, then 2nd duck, and then the 3rd duck in the pond.*
Fingers Held Up	Any finger configuration on one or both hands that is meant to convey a number/quantity	Experimenter: Can you put two ducks in the pond? Child: "This many?" *Holds up pointer and middle finger to indicate the number two.*
Magnitude	Pointing and/or circling, waving, or any similar hand gesture referencing a grouping of objects while also talking about the magnitude, quantity, or total number of objects in the set	Experimenter: Can you put three ducks into the pond? Child: "…" *Picks up set of three ducks and dumps in pond.* "Three!" *Waves hand over pond with palm down to indicate the set.*

Pointing and Counting gestures were coded as a singular unit (e.g. pointing and counting to 6 items is equal to one instance of a point and count gesture). All point and Count gestures were coded as one unit of gesture, regardless of whether the gestures occurred with speech, whether the participant was touching the object as they counted, if the gesture itself was correct (i.e., correct if 1-to-1 correspondence with each object), and whether their verbal count list was correct. Other Math Gestures consisted of two primary gesture types. First, children would hold up a particular quantity of fingers in order to represent a specific quantity. Second, children would indicate information regarding magnitude by making a pointing or waving gesture to a set of objects and pairing that gesture with a verbal statement regarding the magnitude of a set. Critically, we did not differentiate between gestures that did or did not co-occur with speech. In other words, any instance of hand gestures that could be recognized under our coding scheme were included in our analyses.

Children's Speech. While all instances of gesture regardless of speech were coded, one measure of children's speech during the modified Give-N task were extracted from the transcripts for use as a control variable. Word tokens, or the total number of words the child said during the task, was extracted as a measure of the overall amount of speech (M= 142.99, SD= 129.80). Given the ample literature suggesting that children's verbal language and gestures are intertwined from infancy (e.g. Iverson & Goldin-Meadow, 2005), we included this variable as a control in our subsequent analyses in order to account for the possibility that children who are more talkative are naturally more likely to gesture.

Results

All analyses were performed using R (R Development Core Team, 2011). There were no significant correlations between children's gender or any other variables (Knower Level, Pointing and Counting, Age, and WM measures; all *p*-values > .05), and so all further analyses were collapsed across boys and girls.

Working Memory Measures

The mean score on the Word Span task was 3.31 (*SD* = 1.24), with a range of 0 to 6. The mean score for the Touch Base assessment was 6.68 (*SD* = 6.14) with a range of 0 to 30. These measures of WM were positively correlated, *r*(79) = 0.31, *p* = .005. Thus, the scores were combined into a composite measure of children's WM. A z-score was calculated for each measure, and then an overall composite was calculated by averaging the two z-scores.

Gestures by Type

Sixty-three percent of children (*n* = 51) used at least one math gesture during the Give-N task, for an overall total of 319 gestures used across all children. Out of this total, 275 of those gestures were Point and Count (86%), 25 gestures were Fingers Held Up (8%), and 20 gestures were Magnitude Related (6%). On average, children used 3.94 gestures (*SD* = 4.83) with a range of 0-20 gestures (Table 2).

Table 2

Descriptive Statistics by Gesture Type

	M	*SD*	Min	Max
Point and Count Gestures	3.40	4.59	0	18
Fingers Showing Numbers	0.31	0.89	0	5
Magnitude Gestures	0.25	0.60	0	2
Total Gestures	3.94	4.83	0	20

Additionally, we were interested in how each of these gestures were used by children of different Knower Levels. In particular, we wanted to know whether our findings using the adapted Give-N method produced similar types and rates of gesture seen within other version of the Give-N tasks, from previous literature (Authors, 2019; see Table 3 for a breakdown of Gesture type by Knower Level). The distribution of gestures in our study is consistent with prior literature, where Point and Count Gestures dominated the overall total of spontaneous gestures employed by children in a cardinality task. Given the low frequency of the second and third gesture types, and to maintain consistency with the prior literature, Fingers Showing Numbers and Magnitude Gestures were collapsed into one variable, *Other Gestures*.

Table 3

Number and Percent of Children who Used Gesture

		Point and Count		Fingers Showing Numbers		Magnitude	
Knower Level	*N*	*N*	%	*N*	%	*N*	%
0	2	1	50	0	0	0	0
1	9	0	0	2	22	0	0
2	18	6	33	4	22	6	33
3	5	0	0	0	0	0	0
4	1	1	100	0	0	1	100
5	3	1	33	0	0	0	0
6	5	3	60	0	0	0	0
7	1	1	100	0	0	0	0
8	37	30	81	6	16	6	16

Are Children's WM Ability, Cardinality Knowledge, and Gestures Related?

In order to investigate the relations between variables, Pearson's product-moment correlations were calculated between children’s Age, Word Tokens, Point and Count gestures, Other Math Gestures, their Knower Level and their WM composite (Table 4).

Table 4

Descriptive Statistics and Correlations between Age, Word Tokens, Gestures, WM, and Knower Level

Variable	*M*	*SD*	1	2	3	4	5
1. Age	4.75	0.59					
2. Word Tokens	142.99	129.80	0.10				
3. Point and Count	3.40	4.59	.28*	0.57**			
4. Other Math Gestures	0.56	1.05	-0.13	0.19	0.16		
5. WM Composite	0.00	0.81	0.56**	0.24*	0.24*	-0.07	
6. Knower Level	5.09	3.00	.66**	.26*	.45**	-0.19	0.62**

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. * indicates $p < .05$. ** indicates $p < .01$.

We found child age was positively correlated with Knower Level and the WM composite. Children's Word Tokens, or amount of speech they used in the Give-N task, was positively correlated with their Point and Count Gestures, WM Composite, and their Knower Level. There was a strong, positive correlation between Knower Level and WM composite. Children's Point and Count gestures had a significant, positive relation with their age, Knower Level, and Working Memory Composite; however Other Math Gestures were not correlated with any variables of interest. Therefore, all subsequent analyses are run with Point and Count gesture as the sole gesture variable; however, each analysis was also run with a composite of all math gestures, including both Point and Count and Other math gestures combined, and the pattern of results are the same.

To better understand how children's Point and Count Gestures, as well as their WM Composite may impact their performance on the Give-N task, we conducted three partial correlations controlling for Age. Consistent with prior literature (Authors, 2019), Point and

Count gestures were positively correlated with Knower Level, $r(81) = 0.36$, $p < .01$. Furthermore, Knower Level was also positively correlated with WM Composite, $r(81) = 0.40$, $p < .001$. However, the correlation between Point and Count Gestures and the WM Composite (Table 4) was no longer significant when Age was controlled, $r(81) = 0.10$, $p = .385$. Thus, children's Knower Level was correlated with their Point and Count Gestures as well as their WM Composite after controlling for Age. However, Point and Count Gestures were no longer correlated with the WM Composite.

In order to account for the possibility that children who use more speech are more likely to gesture, we conducted three partial correlations controlling for Word Tokens. Even while controlling for amount of speech, Point and Count gestures were positively correlated with Knower Level, $r(81) = 0.38$, $p < 0.01$ as were Knower Level and the WM Composite, $r(81) = 0.59$, $p < 0.001$. The correlation between Point and Count Gestures and the WM Composite was not significant while controlling for Word Tokens, $r(81) = 0.13$, $p = 0.263$. Thus, children's Knower Level was correlated with their Point and Count Gestures as well as their WM Composite after controlling for the total amount of speech.

Do Gestures Moderate the Relation Between WM Ability and Cardinality?

Next, we tested the hypothesis that Point and Count gestures moderate the association between children's WM composite and their Knower Level, controlling for age, using multiple regression analysis. Step 1 included Age and the main effect of the WM composite, Step 2 included the main effect of Point and Count gestures, and Step 3 included the interaction between the WM composite and the Point and Count gestures. The overall model was significant, $F(4,76) = 27.95$, $p < .001$, accounting for 60% of the variance in children's Knower Levels. There were significant main effects for the WM composite ($\beta = 1.46$, $t = 4.01$, $p < .001$), Pointing and Counting gestures (β = 0.18, t = 3.56, p < .001), and Age (β = 2.04, t = 4.50, p < .001). However, the interaction term between WM composite and Pointing and Counting was not significant (β = -0.14, t = -1.49, p = .14).

Does Children's WM Impact Their Gestures and Performance in the Give-N Task?

Although our regression analysis found that the WM composite did not significantly moderate the relation between gestures and knower level in our sample, we conducted a series of post-hoc exploratory analyses to lay the groundwork for hypothesis building for future studies with more power to detect significant interaction effects. In order to gain a descriptive, exploratory understanding of the relations between gestures and working memory in early mathematical contexts, we followed Marstaller & Burianová's (2013) protocol to explore patterns of gesture use within groups of individuals with different WM capacities. Thus, we used the median WM composite score (-0.02) to divide the participants into two groups: low WM ability children (n = 36) and high WM ability children (n = 45). Children with a low WM used on average 1.89 gestures (SD = 3.83), whereas High WM children used 4.60 gestures on average (SD = 4.83). We then looked at the relation between children's pointing and counting gestures and their subsequent Knower Level in the two WM groups separately. Thus, two post-hoc multiple regression analyses were used to test if children's gestures significantly predicted their Knower level controlling for age. The first was run solely with the low WM children; the overall regression was significant and t the two predictors explained 52.31% of the variance (F(2,33) = 18.10, p < .001). We found that use of Pointing and Counting gestures significantly predicted low-WM children's Knower Level (β = 0.35, p < .001), as did Age (β = 2.11, p < .01). The second regression analysis included only the high WM children; the overall regression was significant with the two predictors explaining 31.12% of the variance (F(2,42) = 9.49, p < .001).

However, high WM children's use of Pointing and Counting was not a significant predictor of their Knower Level ($\beta = 0.09$, $p = .192$), though age was significant ($\beta = 2.86$, $p < .001$). Figure 2 shows a scatterplot of the relation between Pointing and Counting gestures and knower level by WM group (low and high).

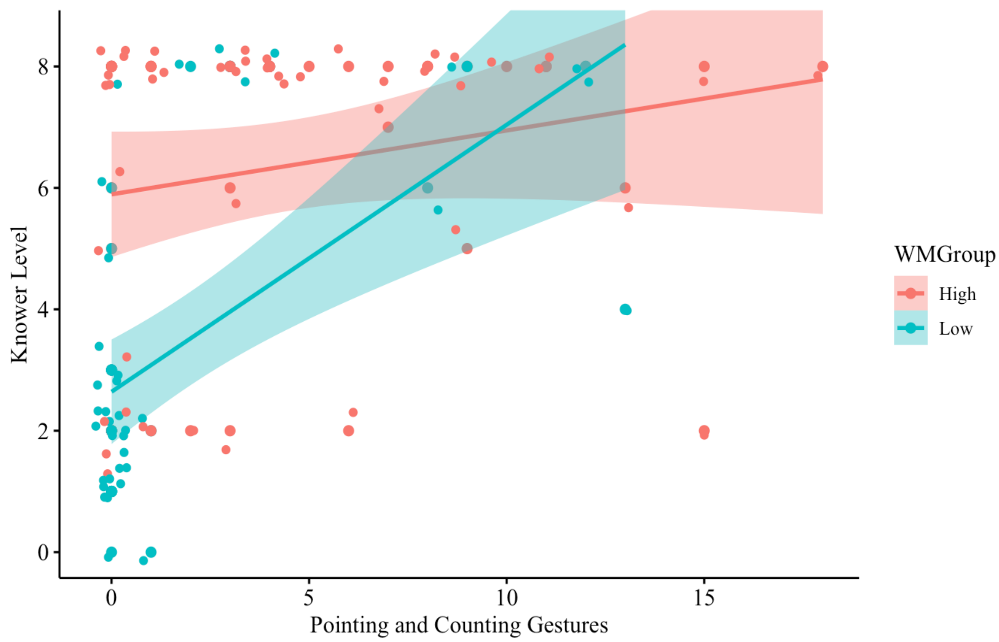

Figure 2. Jittered scatter plot with Pointing and Counting Gestures predicting Knower Level by median split WM group.

Discussion

There were two primary goals of this paper. First, we sought to investigate how children's gesture use, WM, and cardinality knowledge related to each other. Second, we tested our hypothesis that children's gestures moderated the relation between children's WM ability and their performance on the Give-N cardinality task. In addition to these main objectives, we report exploratory post-hoc analyses related to how children of different WM levels may use gestures different on cardinality tasks.

Our first goal was to provide an assessment of the relations between children's gesture use, WM, and cardinality knowledge. In the present study, we found that over half of the children used gestures while completing the cardinality task and the majority of these gestures used were pointing and counting the objects. We also found a positive relation between children's performance on the cardinality task and their use of these gestures, which is consistent with prior literature (Authors, 2019). We also found that these relations cannot be explained by children's age or how much they talked during the task. Our findings extend prior work in three novel ways. First, by the inclusion of children's WM, we were able to investigate how each of these variables may be interrelated in early childhood. Specifically preschool aged children's Knower Level was related to both their Point and Count gestures, as well as their WM, while controlling for age. However, these children's Point and Count gestures were not significantly related to WM while controlling for age. Second, our results with a low-income sample showed the same positive, significant relation between preschool aged child's gestures and their Knower Level that were reported with higher-income children by Authors (2019). This finding is of particular interest, given the previous findings showing that low-income students perform below their mid-income peers on mathematical tasks, but these trends do not persist in nonverbal numerical tasks (Jordan, Huttenlocher, & Levine, 1992, 1994). While previous studies in the domain of mathematics have shown that young children's performance on nonverbal numerical tasks is equivalent regardless of income background (Ginsburg & Russell, 1981; Jordan et al., 1992, 1994), our study suggests that similar expectations could be held for children's use of gesture within numerical tasks as well. Finally, our study extends previous work in an important way by using a different, non-titration version of the Give-N task (adapted from Krajcsi, Fintor, & Hodossy, 2018). This is critical as it shows that neither children's gesture use nor their explicit task knowledge is directly tied to the n+1 titration format; rather, their performance on the task and their gesture use is instead based on their implicit knowledge and the level of difficulty of each individual trial.

Our second goal tested the hypothesis that children's gestures moderated the relation between their WM ability and cardinality performance on the Give-N task. Our hypothesis was based on the ample literature suggesting that children's WM is related to their math ability and that gestures can reduce WM load (Goldin-Meadow, 2011). However, our analyses showed that there was not a significant interaction between a child's WM and their gesture use on their knower-level in our sample. There are several reasons why this may be the case. First, it is possible that the interplay between gesture and WM was not actually captured in the cardinality task used in the study. Our study examined only children's spontaneous, unprompted gestures. In order for children to use these gestures, they would first have to have enough WM resources to remember to employ gestures in the first place. Even if children with lower WM used gesture, their efficacy and cognitive benefit in the math context is still in question. In our sample specifically, many of the children with lower WM did not use gesture at all, which makes assessing any potential relationship between gesture and KL difficult to tease out. Thus, future work with a larger sample would likely have more variability in lower WM kids' gesture use and would in turn give more power to detect the hypothesized (small) interaction effect.

Furthermore, while it is possible that a child could overcome some of the limits of their WM using gestures, here we are expecting them to gesture spontaneously, without any instruction. Thus, future research could consider the possibility that providing young children direction to use a specific gestural strategy, such as pointing and counting, may also change the relation between a child's current WM load, capacity, and their math ability. Finally, our results could be limited by our sample size. As noted in the methods, our proposed sample size was determined using an a priori power analysis including 3 predictor variables. However, in our final analyses we chose to include a fourth variable as a covariate (age), and our data collection was stopped unexpectedly before reaching our predetermined sample size. It is possible that the interaction between gesture and WM was occluded as our study may have been under powered. Future studies considering gesture as a moderator should consider a larger sample size in order to uncover any potential interaction effects.

Lastly, we reported exploratory post-hoc analyses to consider patterns between Point and Count Gesture and Knower Level within high and low WM groups. Based on the regressions and the visualization within the scatterplot, we would hypothesize that future research with a larger sample of children might see a significant relation between gesture and knower level for lower WM children, with less of a relation between gesture and KL for higher WM children. If this pattern of results was upheld in future research, it would imply that children who have higher WM may not necessarily need to rely on an external strategy such as gesture to lower the cognitive demand during the task, as they already have enough mental resources to solve the problem. However, further empirical research must be conducted before any conclusions are drawn related to this complex relation.

In sum, children from a low-income background use gestures in similar types and rates in a non-titrated version of the Give-N task compared to children from higher-income background in prior work using a standard Give-N task. Furthermore, children's use of gesture is positively related to their performance on this task (knower level), above and beyond the impacts of their age and amount they talked during the task. Children's knower level also positively related to their WM while controlling for age. However, no significant relation between children's gesture use and WM was found when controlling for age. While we did not find evidence for the hypothesized moderation model, we did find that both children's gesture use and WM were significant predictors of children's performance on the cardinality task suggesting that both play an important role in children's early numerical knowledge. Given prior work highlighting that quantitative abilities are predictive of later mathematical abilities (Chu, vanMarle, & Geary, 2015; Feigenson, Libertus, & Halberda, 2013; Geary & Vanmarle, 2016; Starr, Libertus, & Brannon, 2013), it is of particular importance to understand how both domain-specific and domain-general factors impact children's early math learning. Here, we provide additional evidence that there are nuanced relations between children's early math learning, their WM, and their use of gesture strategies.

While the present study did not test for a causal relations between gesture, math, and WM, the results indicate that future work on children's mathematical abilities and learning should take each child's WM ability and gestures into consideration. Consistent with prior work suggesting that gestures can help to alleviate demands on WM (Cook et al., 2012), future work should take into consideration the dynamic relation of these variables. In particular, new experimental studies could consider how individual differences in children's WM and math ability relate to their use of gestures, and how this may relate to gesture as an effective tool for learning.

A further line of inquiry may consider how these relations change across time. For example, as children grow older and learn new mathematical concepts, their WM resources and how they are applied to different settings also change (Case 1985; Case, Kurland & Goldberg, 1982; Halford, 1993). The current study considered only one domain of mathematics, within a small age range in early childhood. Thus, the relations found in the current study may change as a child's WM ability changes, and their knowledge of cardinality grows.

Furthermore, little is known about external factors which may increase or decrease their use of gesture in a math context. The current study investigated only the internal factors that may impact children employing these gestures spontaneously (i.e., math ability and WM), but did not provide any assessment of what types and rates of cardinality related gestures are currently being modeled to children by their teachers, parents, or even peers.

Lastly, while our study focused on younger children completing a math task that they were either in the process of learning, or who likely recently reached mastery in cardinality, there are a number of open questions related to how gestures are used in math contexts where mastery of the content has already been reached. While a child could be labeled a cardinal principle knower based on their success within the Give-N task used within this study, there is still room for future research considering whether children may still use gesture, and if so whether these gestures relate to domain-general functions such as WM.

Overall, we find that children's use of gestures in a cardinality task, number knowledge, and WM abilities are interrelated. The overall patterns of relations between gesture and number knowledge with low-income preschoolers is comparable to those found in previous literature with higher-income preschoolers. The descriptive patterns imply future research could consider potential moderations between children's WM and gesture use on their math knowledge. Thus, children's domain-specific and domain-general abilities may be intertwined with their use of gestures.

References

Authors (2019)

Authors (2019)

Authors (Under Review)

Aldugom, M., Fenn, K., & Cook, S. W. (2020). Gesture during math instruction specifically benefits learners with high visuospatial working memory capacity. *Cognitive Research: Principles and Implications*, *5*(1), 1-12. https://doi.org/10.1186/s41235-020-00215-8

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development*, *14*(1), 37-56. https://doi.org/10.1016/S0885-2014(99)80017-3

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science*, *10*(4), 327-333. https://doi.org/10.1111/1467-9280.00163

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, *30*(3), 221-256. https://doi.org/10.1016/S0079-7421(08)60422-3

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable?. *Child development*, *77*(6), 1698-1716.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* Vol. 8 (pp. 47–89). New York: Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In

A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (p. 28–61). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.005

Broaders, S. C., & Goldin-Meadow, S. (2010). Truth is at hand: How gesture adds information during investigative interviews. *Psychological Science*, *21*(5), 623-628. https://doi.org/10.1177/0956797610366082

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, *136*(4), 539. https://doi.org/10.1037/0096-3445.136.4.539

Bull, R., & Espy, K. A. (2006). Working memory, executive functioning, and children's mathematics. In: Pickering SJ, editor. *Working Memory and Education*. Academic Press; Burlington, MA: 2006. pp. 94–123.

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

Case, R. (1985). *Intellectual Development: Birth to Adulthood*. New York: Academic Press.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*(3), 386-404. https://doi.org/10.1016/0022-0965(82)90054-6

Chu, F. W., & 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, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving.

Journal of Experimental Psychology: General, 140(1), 102. https://doi.org/10.1037/a0021790

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23*(1), 43-71. https://doi.org/10.1016/0010-0277(86)90053-3

Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning?. *International Journal of Bilingual Education and Bilingualism*, *7*(4), 303-319. https://doi.org/10.1080/13670050408667815

Claesens, A., & Engel, M. (2013). How important is where you start? Early mathematics knowledge and later school success. *Teachers College Record*, 115(6), 1–29.

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

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, *84*(6), 1863-1871. https://doi.org/10.1111/cdev.12097

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes*, 27(4), 594-610. https://doi.org/10.1080/01690965.2011.567074

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87-114. https://doi.org/10.1017/S0140525X01003922

Crollen, V., Mahe, R., Collignon, O., & Seron, X. (2011). The role of vision in the development of finger–number interactions: Finger-counting and finger-montring in blind children. *Journal of Experimental Child Psychology*, 109(4), 525–539. https://doi.org/10.1016/j.jecp.2011.03.011

Development Core Team, R. (2011). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Engle, R. W. (2002). Working memory capacity as executive attention. C*urrent Directions in Psychological Science*, *11*(1), 19-23. https://doi.org/10.1111/1467-8721.00160

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*(1), 465-486. https://doi.org/10.1207/s15326942dn2601_6

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191. https://doi.org/10.3758/BF03193146

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives*, *7*(2), 74-79. https://doi.org/10.1111/cdep.12019

Gathercole, S. E., Lamont, E., & Alloway, T. P. (2006). Working memory in the classroom. In *Working Memory and Education* (pp. 219-240). Academic Press. https://doi.org/10.1016/B978-012554465-8/50010-7

Geary, D. C., & Vanmarle, K. (2016). Young children's core symbolic and nonsymbolic quantitative knowledge in the prediction of later mathematics achievement. *Developmental Psychology*, *52*(12), 2130. https://doi.org/10.1037/dev0000214

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PloS one*, 8(1), e54651. https://doi.org/10.1371/journal.pone.0054651

Gelman, R., & Gallistel, C. R. (1978). *The child's concept of number.* Cambridge, MA: Harvard.

Gelman, R., & Gallistel, C. R. (1986). *The child's understanding of number*. Harvard University Press.

Gibson, D. J., Gunderson, E. A., Spaepen, E., Levine, S. C., & Goldin-Meadow, S. (2019). Number gestures predict learning of number words. Developmental Science, 22(3), e12791.

Ginsburg, H. P., & Russell, R. L. (1981). Social class and racial influences on early mathematical thinking. *Monographs of the Society for Research in Child Development*, 46(6, Serial No. 69). https://doi.org/10.2307/1165946

Goldin-Meadow, S. (2011). Learning through gesture. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(6), 595-607. https://doi.org/10.1002/wcs.132

Goldin-Meadow, S., & Beilock, S. L. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science*, *5*(6), 664-674. https://doi.org/10.1177/1745691610388764

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, 20(3), 267-272. https://doi.org/10.1111/j.1467-9280.2009.02297.x

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology*, *91*(4), 720. https://doi.org/10.1037/0022-0663.91.4.720

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*(6), 516-522. https://doi.org/10.1111/1467-9280.00395

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction*, *9*(3), 201-219. https://doi.org/10.1207/s1532690xci0903_2

Griffin, S., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 25 – 49). Cambridge, MA: MIT Press.

Gunderson, E. A., Spaepen, E., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2015). Gesture as a window onto children's number knowledge. *Cognition*, 144, 14-28. https://doi.org/10.1016/j.cognition.2015.07.008

Halford, G. S. (2014). *Children's understanding: The development of mental models*. Psychology Press.

Imbo, I., & Vandierendonck, A. (2007). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology*, *96*(4), 284-309. https://doi.org/10.1016/j.jecp.2006.09.001

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, *396*(6708), 228-228. https://doi.org/ 10.1038/24300

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367-371. https://doi.org/10.1111/j.0956-7976.2005.01542.x

Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, 77, 153 – 175. https://doi.org/10.1111/j.1467-8624.2006.00862.x

Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology*, 28, 644 – 653. https://doi.org/10.1037/0012-1649.28.4.644

Jordan, N. C., Levine, S. C., & Huttenlocher, J. (1994). Development of calculation abilities in middle- and low income children after formal instruction in school. *Journal of Applied Developmental Psychology*, 15, 223 - 240.https://doi.org/10.1016/0193-3973(94)90014-0

Kendon, A. (1994). Do gestures communicate? A review. *Research on language and social interaction*, 27(3), 175-200. https://doi.org/10.1207/s15327973rlsi2703_2

Kirsh, D. (1995). Complementary strategies: Why we use our hands when we think. In *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 212-217).

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4), 513-549. https://doi.org/10.1016/0364-0213(94)90007-8

Kolkman, M. E., Hoijtink, H. J., Kroesbergen, E. H., & Leseman, P. P. (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

Krajcsi, A., Fintor, E., & Hodossy, L. (2018). A refined description of preschoolers' initial symbolic number learning. *OSF Preprint*. https://doi.org/ 10.31219/osf.io/2kh9s

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*(3), 677-692. https://doi.org/10.1016/j.jecp.2010.11.001

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395–438. https://doi.org/10.1016/j.cognition.2006.10.005

Le Corre, M., Van de Walle, G. A., Brannon, E., & Carey, S. (2006). Re-visiting the performance/competence debate in the acquisition of counting as a representation of the positive integers. *Cognitive Psychology*, *52*(2), 130–169. https://doi.org/10.1016/j.cogpsych.2005.07.002

LeBlanc, M. D., & Weber-Russell, S. (1996). Text integration and mathematical connections: A computer model of arithmetic word problem solving. *Cognitive Science*, 20, 357-407. https://doi.org/10.1016/S0364-0213(99)80010-X

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. transcription format and programs (Vol. 1)*. Psychology Press.

Marstaller, L., & Burianová, H. (2013). Individual differences in the gesture effect on working memory. *Psychonomic Bulletin & Review*, *20*(3), 496-500. https://doi.org/10.3758/s13423-012-0365-0

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81. https://doi.org/10.1037/h0043158

Monette, S., Bigras, M., & Guay, M. C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology*, *109*(2), 158-173. https://doi.org/10.1016/j.jecp.2011.01.008

Müller, U., Kerns, K. A., & Konkin, K. (2012). Test–retest reliability and practice effects of executive function tasks in preschool children. *The Clinical Neuropsychologist*, *26*(2), 271-287.

Nicolaou, E., Quach, J., Lum, J., Roberts, G., Spencer-Smith, M., Gathercole, S., Anderson, P.J., Mensah, F.K. & Wake, M., (2018). Changes in verbal and visuospatial working memory from grade 1 to grade 3 of primary school: Population longitudinal study. *Child: Care, Health and Development*, 44(3), 392-400. https://doi.org/10.1111/cch.12543

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, 3, 359–400. http://dx.doi.org/10.1016/0885-2014(88)90021-4

Perry, M., Woolley, J., & Lfcher, J. (1995). Adults' abilities to detect children's readiness to learn. *International Journal of Behavioral Development*, 18(2), 365-381. https://doi.org/10.1177/016502549501800211

Pickering, S. & Gathercole, S. (2001). *Working memory test battery for children*. London: Psychological Corporation Europe.

Ping, R., & Goldin-Meadow, S. (2010). Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive Science*, 34(4), 602-619. https://doi.org/10.1111/j.1551-6709.2010.01102.x

Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition*, 92(3), 329-352. https://doi.org/10.1016/j.cognition.2003.10.001

Saxe, G. B., Guberman, S. R., & Gearhart, M. (1987). Social processes in early number development. *Monographs of the Society for Research in Child Development*, 52(2, Serial No. 216). https://doi.org/10.2307/1166071

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*, 19, 99 – 120. https://doi.org/10.1016/j.ecresq.2004.01.002

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, 110(45), 18116-18120. https://doi.org/10.1073/pnas.1302751110

Stipek, D. J., & Ryan, R. H. (1997). Economically disadvantaged preschoolers: Ready to learn but further to go. *Developmental Psychology*, 33, 711-723. https://doi.org/10.1037/0012-1649.33.4.711

Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology*, 79(3), 294-321. https://doi.org/10.1006/jecp.2000.2587

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296. https://doi.org/10.1023/A:1022193728205

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28(2), 187-204.

https://doi.org/10.1016/S0361-476X(02)00007-3

Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language*, 50(4), 395-407. https://doi.org/10.1016/j.jml.2004.01.002

Wakefield, E., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science*, *21*(6), e12664. https://doi.org/10.1111/desc.12664

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/10.3102/0013189X14553660

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193. https://doi.org/10.1016/0010-0277(90)90003-3

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358(6389), 749. https://doi.org/10.1038/358749a0