The Relationship Between Number Skills and Language in Preschool Children

By

Jake Kaufman

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

October 31, 2020

Nashville, Tennessee

Approved:

Bethany Rittle-Johnson, Ph.D.

Gavin Price, Ph.D.

ACKNOWLEDGEMENTS

First off, I need to thank Dr. Bethany Rittle-Johnson, Dr. Gavin Price, Dr. Brianna Yamasaki, and Dr. James Booth for their support throughout this process. Without their advice and guidance, this project would not have been completed, and I would not be where I am today. In addition, I need to thank Marisa Lytle for her contributions to the project, as well as Minji Kim for her role in data collection. I also need to thank all the families and children who participated. Without their support, this, and all of our research, would not be possible. Finally, I need to thank my friends, and especially my family, for the unwavering support, and for always being in my corner.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Numbers are everywhere. Whether telling the time, paying for groceries, or checking the temperature, the ability to understand and manipulate symbolic numbers is critical. Not only is this skill crucial for everyday success, but it is well known that symbolic number skills, which for the purposes of this paper refer to both verbal number skills (e.g. verbal counting) and written (i.e. Arabic) number skills (e.g. Arabic number comparison), are important for later academic success. Early numeracy skills measured before or during kindergarten have been found to predict growth and performance on later math measures in the first few years of formal schooling (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Chiara Passolunghi, Vercelloni, & Schadee, 2007). In addition, prekindergarten math skills appear to be the single best predictor of later school achievement; better than early literacy, attention skills, socioemotional skills, family background measures, or IQ (Duncan et al., 2007).

Although the importance that symbolic number skills play in our lives is clear, we are still cultivating an understanding of how to develop a strong symbolic number system. One important domain that could impact symbolic number development is language. LeFevre et al. (2010) suggest that learning the rules of the number system is comparable to mastering the rules of any symbolic system, such as language. Therefore, it may be that the skills children relied on when learning language come into play when learning numbers; and that those with stronger language skills have an easier time learning about numbers. Previous work examining the relationship between language and symbolic number skills has mainly focused on two language skills: semantic knowledge and phonological awareness (Koponen, Aunola, Ahonen, & Nurmi, 2007; Krajewski & Schneider, 2009; Moll, Snowling, Göbel, & Hulme, 2015; Negen & Sarnecka, 2012). However, it appears that these language skills may be related to different components of symbolic number skill. Semantic knowledge involves mapping from a symbol to its meaning, a process required when working with numbers that require a quantitative understanding (e.g. Arabic numerals). This symbolic mapping skill, a term we will use to reflect number knowledge that requires a quantitative understanding, is necessary when completing an Arabic number comparison task (i.e. identifying the largest numeral out of at least two numerals) as one must be able to map from the number to the quantity that number represents. It is this symbolic mapping skill that is one potentially important skill involved when accessing the meaning associated with an item (e.g. what a word means, or what a number represents). On the other hand, phonological skill involves the recognition and manipulation of sounds, suggesting it plays a role in developing number skills that do not require an understanding of quantity (e.g. verbal number skills). A good example of this would be verbal counting ability; which reflects knowledge of number words and the count sequence. One can produce number words and the count sequence without an understanding of the meaning associated with each number word. See Figure 1 for an overview of these hypothesized relationships.

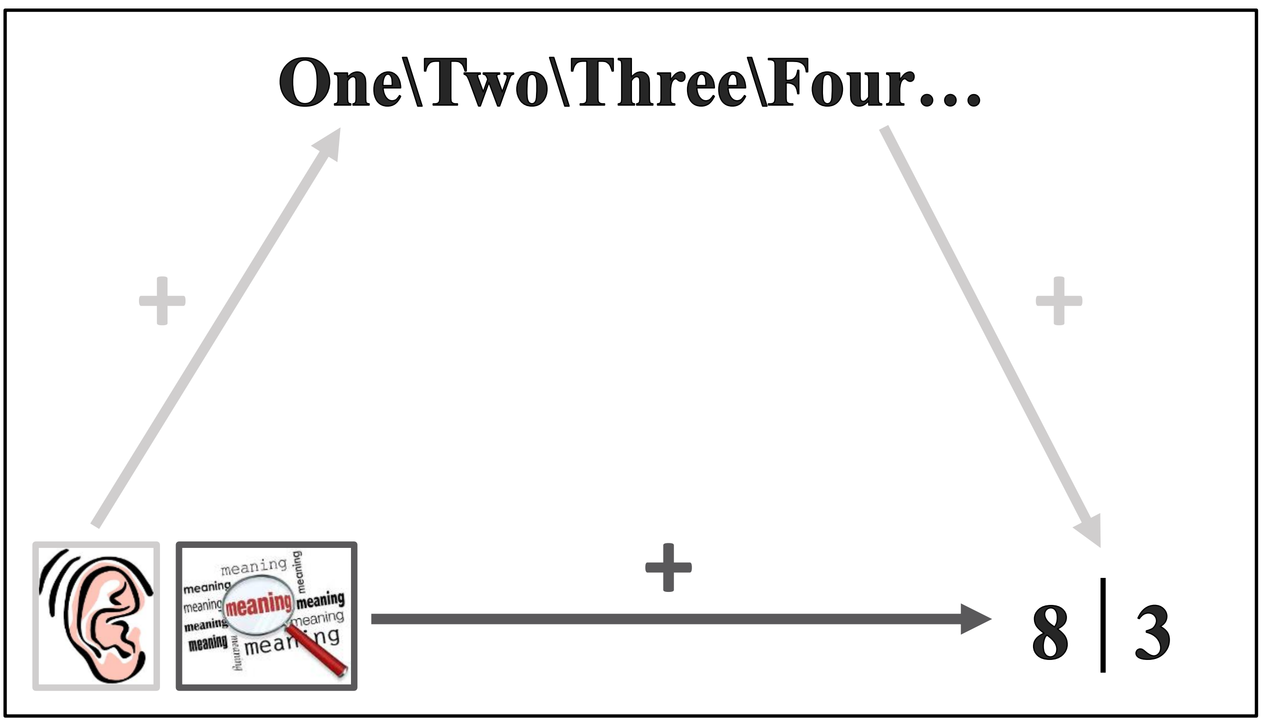

Figure 1. In light grey is our indirect phonological skill hypothesis. In dark grey is our direct semantic knowledge hypothesis.

Semantic Knowledge and Symbolic Mapping Skill

Semantic knowledge is our understanding of what words mean and represent. Previous work demonstrates a direct relationship between semantic knowledge and symbolic mapping skills. In a sample of 4- to 6-year-olds, Purpura and Ganley (2014) examined the relationship between vocabulary knowledge, a common measure of semantic knowledge, and a range of early numerical skills, including verbal counting and number comparison. They found that vocabulary knowledge was significantly related to number comparison ability; however, vocabulary knowledge did not significantly predict verbal counting skill. In addition, they found a significant relationship between verbal counting and number comparison ability, suggesting counting ability may play a role in the development of symbolic mapping skills. In an earlier study, Negen and Sarnecka (2012) examined the relationship between vocabulary knowledge and performance on the Give-N Task in a sample of 2- to 4-year-old children. To be successful on the Give-N task, an individual needs to identify the number being requested (via a verbal number name), and correctly associate it with the quantity it represents (i.e., cardinality knowledge). Negen and Sarnecka (2012) found that performance on the Give-N task was significantly related to vocabulary knowledge, independent of age. These results suggest that an ability to identify the meaning associated with an item (e.g. what a word means, or what a number represents) may be underlying the relationship between semantic knowledge and symbolic mapping skills.

Computational models have also been developed to try and identify how we complete tasks that require knowledge of number quantities. Tasks involving simple arithmetic (e.g. addition or subtraction) are often used to assess whether an individual has this knowledge, as successful performance relies on our ability to access and manipulate the quantities associated with written numerals. Therefore, it is critical for us to understand the mechanisms underlying our ability to perform simple arithmetic. For this reason, Stoianov, Zorzi, and Umilta (2003) set out to simulate a simple addition task (two operands only, with each operand between numbers 1 and 12) using computational models. They based their model off Zorzi and Butterworth's (1999) numerosity code, where a numbers magnitude is represented as the number of units activated, such that larger numbers include smaller numbers. Stoianov et al. (2003) tested a model that was trained on both symbolic and semantic inputs. Symbolic inputs involved activation of just the single unit corresponding to that number, while semantic inputs activated all units up to and including the given number. For example, if we were working with an operand of seven, the model with symbolic input only activated the single unit associated with seven, while the model with semantic input activated all units up to and including seven. They found that it was the semantic representations, based off the numerosity code, that led to the success of the model in making number judgments. This lends support to the notion that the relationship between vocabulary and symbolic mapping skills may involve an ability to access the underlying semantic representation of words or written (i.e. Arabic) numerals.

Phonological Skill and Verbal Number Skill

A consistent finding in the literature is that increased phonological skill is related to increased verbal counting ability (Koponen et al., 2007; Moll et al., 2015; Soto-calvo et al., 2019). In their isolated number words hypothesis, Krajewski & Schneider (2009) argue that phonological awareness should influence the development of basic numeral skills (e.g. verbal counting), since at this point the need to associate number words with quantities is not necessary. To support this claim, they examined the impact of phonological awareness on early numeracy skills in 4-6-year-old children. A composite phonological awareness score was created by combining scores from a phoneme synthesis task and rhyme judgment task. On the phoneme synthesis task participants heard a sequence of single phonemes and had to match the sound to the appropriate word, and then pick the corresponding picture out of a choice of four pictures. On the rhyme task, children listened to four words and then had to indicate which of the four words did not rhyme with the others. Basic numerical skills (i.e. number words isolated from quantity, and what they labelled as Level 1 number skills) were assessed by having children manipulate the count sequence. They were asked to count forwards, backwards, and to verbally identify individual Arabic numerals between 1 and 20 that were randomly presented to them. They also measured number skills that required quantity knowledge (what they labelled as Level 2 number skills). These tasks involved comparing which of two number words represented more or less, matching quantities (e.g. three and five dots) to the corresponding Arabic numbers, and matching Arabic numbers (e.g. 4 and 6) to their corresponding quantities. In support of their isolated number words hypothesis, they found that differences in phonological awareness substantially predicted individual differences in Level 1 number skills (i.e. number words not linked to quantities); but not Level 2 skills that involved an association between symbol and quantity. However, they did find that Level 1 number skills predicted Level 2 number skills. This suggests that phonological skill is important for developing verbal counting ability, and that verbal counting ability may serve as the foundation for developing more advanced symbolic mapping skills.

Further support for the relationship between phonological awareness and verbal counting comes from Simmons & Singleton's (2007) weak phonological representation hypothesis. They suggest that individuals who are unable to form strong phonological representations are likely to be poor counters, and that this will specifically manifest in verbal counting speed. Support for this was demonstrated in a study where dyslexic children (who struggle to form strong phonological representations), performed worse than their typically developing peers on a test of verbal counting speed (Simmons, 2002).

Verbal to Symbolic Mapping Skills: Counting to Comparison

Although there is support for the relationship between phonological and verbal number skills, as well as semantic knowledge and symbolic mapping skills, the question remains of how verbal and symbolic mapping skills are related. Krajewski & Schneider's (2009) findings suggest that basic verbal number skills (e.g. verbal counting) may serve as the foundation for more complex symbolic mapping skills (e.g. Arabic number comparison). In addition, Aunio and Niemivirta (2010) found a significant correlation between counting skills (both verbal and object counting) measured in kindergarten, and relational number skills measured one year later. Included in these relational skills was number comparison ability, indicating that those who are better counters will perform better on number comparisons. Furthermore, in a different study with kindergarten children, Purpura and Ganley (2014) found a significant correlation between verbal counting ability and number comparison performance. These findings suggest that verbal counting ability is important for the development of symbolic mapping skills.

Further support for this notion comes from the Symbolic Account of Arabic number development proposed by Hurst, Anderson, and Cordes (2017). The Symbolic Account suggests that children acquire an understanding of Arabic numerals through their knowledge of the count sequence and number words. In support of this claim, Hurst et al. (2017) found a significant indirect effect from quantity-word performance to quantity-numeral performance through word-numeral performance, in a sample of 3- and 4-year-old children. This suggests that children first learn the verbal count sequence, then develop an understanding of what number words represent, and ultimately leverage that knowledge to develop an understanding of Arabic numerals. Their results indicate that this final step is accomplished by matching number words to their corresponding Arabic numeral, and deducing that the number word and Arabic numeral represent the same quantity. These findings suggest that verbal number skills may serve as a foundation for the development of symbolic mapping skills; with phonological skill being important for helping first develop verbal number skills (see Figure 1).

The Present Study

Although the previous work examining the relationships between semantic knowledge, phonological awareness, and symbolic number knowledge provide us with some indication of their relatedness, this work is still lacking in two respects. First, these studies have not examined both phonological skill and semantic knowledge and their relationships with symbolic number skills. Rather, they have used either phonological skill or semantic knowledge as a proxy for language skill in general (e.g. Geary, van Marle, Chu, Hoard, & Nugent, 2019; Purpura & Ganley, 2014; Purpura & Reid, 2016). This makes it difficult to accurately assess how these two language skills may be related to different components of symbolic number skill (e.g., verbal vs. mapping skill) at this stage of development. Second, although studies have indicated a relationship between phonological awareness, semantic knowledge, and symbolic number skill, we are still developing an understanding of the mechanisms that are underlying these relationships (Purpura, Logan, Hassinger-Das, & Napoli, 2017). The goal of the present study is to identify how phonological skill and semantic knowledge differentially relate to symbolic number skills, by examining the mechanisms underlying these relationships. We are specifically interested in how these language skills relate to Arabic (i.e. written) number comparison ability because being able to identify the larger of two numbers is often how we judge if an individual understands numbers (McCloskey, 1992).

Hypotheses

We hypothesize an indirect relationship between phonological skill and Arabic number comparison ability that will be mediated by verbal counting ability. Greater phonological skill reflects an ability to segment and maintain sounds over time, a process that is involved in successful verbal counting. Furthermore, we expect that this ability to segment, and particularly maintain sounds, allows for better learning of Arabic numerals as individuals have more time to attach meaning (i.e. quantity) to numerals through matching with number words. Therefore, we predict that those who perform better on our phonological task will also be faster verbal counters (indicating greater proficiency with the count sequence); and faster verbal counters will perform better on the Arabic number comparison task (see Figure 1).

In relation to semantic knowledge and Arabic number comparison ability, we hypothesize a direct relationship. Semantic knowledge requires an understanding of the underlying meaning of a word. Similarly, knowledge of Arabic numerals requires an understanding of the underlying meaning (i.e. quantity) of that number. When making semantic judgments or accessing the quantity associated with an Arabic numeral, we must map from the symbol (e.g., picture or Arabic numeral) to what the symbol is representing (i.e. meaning). This shared underlying mechanism of mapping from symbol to meaning leads us to hypothesize that those who perform better on our semantic knowledge task, will also perform better on our Arabic number comparison task (see Figure 1).

Methods

Participants

All experimental procedures were approved by the Institutional Review Board at Vanderbilt University. Written assent and consent was obtained from the children and their parents or guardians. Participation occurred either at the Brain Development Lab at Vanderbilt University, or at two local preschools. Participants were given a developmental history questionnaire to assess eligibility for participation. According to the questionnaire, all participants were typically developing native English speakers. Exclusion criteria for participation were: (a) diagnosis of a psychiatric disorder including attention deficit hyperactivity disorder (ADHD), (b) neurological disease or epilepsy, (c) preterm birth before 34 weeks, (d) birth complications requiring admission into the neonatal intensive care unit, (e) head injury requiring emergency medical evaluation, (f) uncorrected visual impairment, (g) intellectual disability, (h) medication affecting central nervous system processing, (i) speak a language other than, or in addition to English, (j) hearing impairment, (k) enrollment in early intervention services (e.g. speech language therapy or auditory verbal therapy).

Due to the COVID-19 pandemic we unfortunately had to stop data collection. This resulted in a small and underpowered sample, making it difficult to properly identify relationships between our measures. A power analysis conducted through G* Power (Faul, Erdfelder, Buchner, & Lang, 2009) revealed that to see relationships similar to those observed in previous studies we would need a sample of 34 participants to assess our semantic knowledge hypothesis, and 44 participants to assess our phonological skill hypothesis (Purpura et al., 2017).

At the time of our data collection stoppage, parents of 31 children completed the study consent forms. Participants were excluded from analyses for: (a) averaging less than 55% accuracy on at least one experimental task, (b) not being able to identify numerals 1-9, as this is necessary to be able to complete the Arabic number comparison task, (c) response bias greater than 35% (i.e. favoring one response over the other) on at least one task, suggesting a misunderstanding of the task, (d) scoring greater than one standard deviation (SD) below the mean on at least one standardized measure, as this suggests a mild to moderate deficiency in the skill being assessed (although no participants were excluded solely for this reason). Three participants were excluded due to low accuracy (averaged less than 55%) on at least one task, a response bias greater than 35% on at least one task, and not being able to identify numerals 1-9. Two participants were excluded due to low accuracy on at least one task, and a response bias greater than 35% on at least one task. Four participants were excluded due to low accuracy on at least one task, and not being able to identify numerals 1-9. Lastly, five participants were excluded due to low accuracy on at least one task.

Of the remaining 17 participants included in analyses, 56.3% were female. Children were 50 – 60 months old (*M* = 4 years 8 months, *SD* = 3 months). Socioeconomic status was assessed through parental self-reports of household income. Eight participants listed an annual income greater than $100,000; two listed an income between $75,000 and $100,000; two listed an income between $50,000 and $75,000; and five chose not to report. Race and ethnicity were self-reported by the child's parent, with fifteen identifying as White/Caucasian, one identifying as Black/African American, and one identifying as more than one race/ethnicity.

Measures

Early Symbolic Number Skill. To assess knowledge of Arabic numerals, participants completed a symbolic number identification task. Participants were presented numerals 1-9 in the prespecified random order of: 2, 5, 1, 8, 3, 7, 9, 6, 4. Participants were asked to verbally label each number. This measure was used for exclusionary purposes and was not included in any analyses.

Verbal counting knowledge was assessed with two tasks that were modelled after previous verbal counting tasks (cf. Purpura & Ganley, 2014; Hornburg, Schmitt, & Purpura, 2018). The first task assessed knowledge of the count sequence by examining participants' counting accuracy. Participants were instructed to "count as high as you can starting from 1." If a participant reached 30, we would stop them, as the counting pattern begins to repeat itself after this point. Spontaneous self-corrections during counting were permitted, and one point was given for each number correctly counted up to 30. In the second verbal counting task, we were interested in counting speed. In this case, participants were instructed to "count from 1 to 10, as fast as you can." Spontaneous self-corrections during counting were permitted, however, if a participant did not count all numbers from 1-10 they were excluded from analyses involving this task. Both verbal counting tasks were used in our analyses.

Lastly, to examine participants' quantitative understanding of symbolic numbers (i.e. symbolic mapping skill) we used an Arabic number comparison task developed in E-prime (Schneider, Eschman, & Zuccolotto, 2002). Compared to previous research involving preschool and kindergarten children, we adjusted the task so it only involved visual presentation of two Arabic numerals. Previous work has asked children to determine the largest of four numbers, and involved both visual and verbal trials (Hornburg, Schmitt, & Purpura, 2018; Purpura & Ganley, 2014); however, we were primarily interested in the mapping from a visual symbol (e.g. Arabic numeral) to its meaning (i.e. quantity). In the Arabic number comparison task, participants had to determine which of two simultaneously presented numbers between 1-9 was more (e.g. 2 | 5). The numerical distance between the two numbers ranged from 1-6 (e.g. 2 | 5 is a numerical distance of 3), with one third of the trials consisting of a small distance (i.e. the distance between the numbers was 1-2), one third consisting of medium distance (.e. the distance between the numbers was 3-4), and one third consisting of a large distance (i.e. the distance between the numbers was 5-6). Numbers were presented in size 70 font, located in the middle of the screen, and were 9.5 cm apart. One trial consisted of a 1000 ms fixation period, followed by a 2000 ms presentation period where both numbers were presented simultaneously, and lastly, a 3000 ms response screen. However, a response could be made as soon as the numbers were presented, and up until presentation of the stimuli in the following trial. Therefore, the entire response interval lasted for 6000 ms (e.g. 2000 ms second stimuli presentation + 3000 ms response screen + 1000 ms fixation from following trial). There were six practice trials, with accuracy feedback provided after each trial. However, the practice trials included numerical distances of 7 and 8. This was for two reasons: (a) to not repeat any distances in the real trials, (b) to provide easier examples of what the participant had to do in the task with the hopes of increasing understanding. The task consisted of 60 experimental trials, split across 5 runs. This meant there were 12 trials per run, which allowed for each numerical distance to be tested twice per run. Data from this task was included in our analyses. However, we only analyzed performance on the small distance comparison trials. This decision was made because these trials are the most challenging. Therefore, we expected these trials to most accurately reflect, and be most sensitive to, a participants' ability to map from a symbol to that symbols meaning. As a result of this, we felt that analyzing the small distance comparison trials would provide us with the best measure of a participants' Arabic number knowledge and the underlying mechanism of symbol to meaning mapping.

Early Language Skill. Phonological awareness was measured using two tasks. The first was the word segment recognition test of A Developmental Neuropsychological Assessment, Second Edition (NEPSY-II; Brooks, Sherman, & Strauss, 2009). In this task, the experimenter presented the child with a page containing either three or four images. The experimenter then explained how he will name the images on the page, but will then just say a part of one of the images, and that it is the participant's job to then correctly identify the image the experimenter is referring to. Testing ended when the participant made 6 consecutive errors or completed all items. This task was administered for exclusionary purposes. Furthermore, phonological skill was assessed with a novel rhyme judgment task we developed in E-prime (Schneider, Eschman, & Zuccolotto, 2002). This task was based off a sound judgment task used in a study conducted by Weiss, Cweigenberg, and Booth (2018). Participants had to judge whether two sequentially presented auditory-visual stimuli pairings rhymed (see Figure 2 for an overview of one trial).

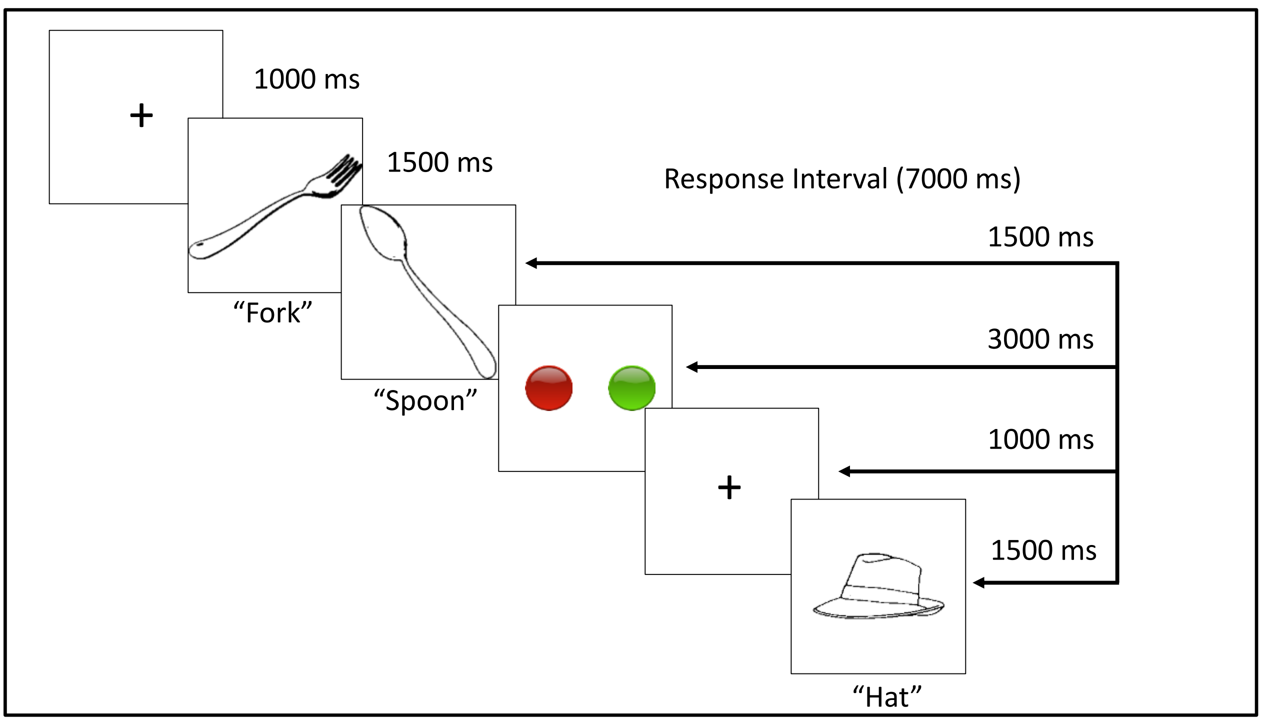

Figure 2. Overview of one semantic meaning judgment task trial. *Note*. The structure of one trial is the same for the rhyme judgment task.

Participants responded by pressing a green button if the words rhymed, and a red button if the words did not rhyme. Headphones were used to listen to the words and to limit distracting noise. There were 24 practice trials (12 rhyme pairs and 12 non-rhyme pairs), which included accuracy feedback after each trial, and 48 experimental trials (24 rhyme pairs and 24 non-rhyme pairs). The 48 experimental trials were split across two runs, each consisting of 12 rhyme pairs and 12 non-rhyme pairs. Data from this task was included in our analyses.

Semantic knowledge was measured using two tasks. The first was the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007). In this task, participants were shown a page with 4 colored pictures. The experimenter then asked the child to point to one of the four images on the page. Testing continued until the participant made 8 or more errors in a given set (sets are comprised of 12 trials). This task was administered for exclusionary purposes. In addition, semantic knowledge was assessed using a novel meaning judgment task we developed in E-prime (Schneider, Eschman, & Zuccolotto, 2002). Similar to our rhyme judgment task, this task was based off a meaning judgment task that was also used in the study conducted by Weiss et al. (2018). Participants had to judge whether two sequentially presented auditory-visual stimuli pairings were related. The presentation and structure of one trial (see Figure 2), and the number of trials in the semantic meaning judgment task were the exact same as trials in the phonological rhyme judgment task. The only difference being that in this task participants judged whether the two sequentially presented auditory-visual stimuli pairings were related in meaning (24 related pairs and 24 unrelated pairs). Data from this task was included in our analyses.

General Cognitive Ability. General cognitive ability was measured using the Reynolds Intellectual Screening Test, Second Edition (RIST-2; Reynolds & Kamphaus, 2015). The RIST-

2 is comprised of two subtests: a verbal “guess what” subtest, and a nonverbal “odd-item out” subtest. In the verbal “guess what” subtest, participants were asked a question to which they needed to provide a response (e.g. “What is round, bounces, and is often thrown or kicked?”). In the nonverbal “odd-item out” subtest, participants were presented with a page filled with images, and had to identify the item that did not belong with the rest of the set (e.g. identifying that one circle presented with five squares is the “odd-item out”). Testing stopped when participants were incorrect on 4 consecutive trials. This task was administered for exclusionary purposes.

Language Tasks Stimuli Development. For the phonological rhyme judgment task, and the semantic meaning judgment task, all images used were from the International Picture Naming Project (IPNP; Bates et al., 2000).

Images included in the phonological awareness rhyme judgment task were selected based off the following criteria: (a) words were sorted by shared rhyme ending, (b) only single syllable words beginning with a consonant were then kept, (c) the words had to have a percent name agreement of .80 or higher (i.e. participants produced the target name 80% or more after being presented with the picture), (d) there could be no repeated words within the task, or between the phonological rhyme judgment task and the semantic meaning judgment task.

Images included in the semantic meaning judgment task were selected based off the following criteria: (a) forward cue-to-target strength (FSG) above .10, and backward cue-to-target strength (BSG) of .05 as indicated by the USF Free Association Database (Nelson, McEvoy, & Schreiber, 1998), (b) only single syllable words beginning with a consonant were then kept, (c) the words had to have a percent name agreement of .80 or higher, (d) there could be no repeated words within the task, or between the phonological rhyme judgment task and the semantic meaning judgment task.

All images were rated for visual complexity. All images used were black and white. The black component of the images were lines that provided the outline for each image. Visual complexity ratings were based off the amount of black lines within a given image, with less complex images having fewer black lines (i.e. less detail needed to portray the image), and more complex images having greater amounts of black lines (i.e. greater detail needed to portray the image). Images were given a “1” for little complexity, “2” for medium complexity, and “3” for high complexity. Four independent raters went through the images and provided ratings. If an image did not have a majority rating (e.g. 2 raters gave the image a “1” and 2 raters gave it a “2”), the image was discussed, and a unanimous decision was agreed upon. Visual complexity was assessed in order to control for this within and between tasks.

The procedure for developing the audio stimuli was the same for both tasks. Phase one involved the recording of the words. Recording took place in a sound booth, and all words were recorded by a native American English female speaker. Each word was recorded twice, and the order of the words varied between the recordings. Words were recorded in a natural, consistent, and rising tone. Phase two involved editing the words. All editing took place in Praat, a system for doing phonetics by computer (Boersma & Weenink, 2019). Editing involved the following steps: (a) a native American English speaker listened to both words and selected the better option in terms of natural speech, clarity, and rising tone, (b) the beginning and ending silence or noise from each single-word auditory file was cut, (c) words were sorted by length (shortest to longest) and binned to 500, 600, and 700 ms using Praat’s change duration function, (d) all files were normalized for volume and to an amplitude of 76 decibels (dB) using Praat’s scale intensity function, (e) four native American English speakers listened to, and identified all single-word auditory files, (f) any words identified incorrectly or as problematic were re-recorded and edited through steps (a) – (e).

Lastly, condition x task ANOVA's revealed no significant differences between or within conditions for picture naming agreement, written word frequency (Balota et al., 2007), phonotactic frequency (Vitevitch & Luce, 2004; only for the rhyme judgment task), visual complexity of the pictures, amplitude of the audio stimuli, and duration of the audio stimuli (all *p*s > .12).

Procedure

Assessment procedure. Children were assessed on all tasks in the fall of 2019. Assessments were conducted by individuals who either completed or were working towards a bachelor's degree in psychology or neuroscience. Participation took place at either the Brain Development Lab at Vanderbilt University, or at local preschools at times identified by the schools in a room designated by the individual school directors or teachers. Individuals who participated at the Brain Development Lab completed all measures in one 60- to 75-minute session. Individuals who participated at their preschool completed all measures over two 30-minute sessions; with one session focused on the standardized measures and the other session focused on the experimental tasks.

Analytic procedure. The primary research questions were addressed using mediation analyses. Given the limited sample size, bootstrapping (i.e. random sampling with replacement) was employed to help answer our questions and examine the mechanistic relationships between early language and number skills. Bootstrapping is a technique where we form a resample, by randomly drawing observations with replacement, from our original sample. This technique can be carried out many times (typically in the thousands), and with it, we are able to increase the number of samples of our data – making it an extremely useful tool when you are working with a small original sample. This process of repeatedly taking samples with replacement from our original sample provides a $100(1-\alpha)\%$ confidence interval (CI) for the effect being bootstrapped. The effect being bootstrapped, typically the indirect effect with mediation analyses, is considered statistically different from zero if the CI does not contain zero (Hayes & Preacher, 2014).

Our analyses focused on our experimental language and number measures (phonological rhyme judgment task, semantic meaning judgment task, verbal counting tasks, and Arabic number comparison task). The standardized measures we administered were used only for exclusionary purposes and were not included in any of our analyses. The reason for this was because our primary interest was to examine reaction time relationships between the tasks, and we did not have reaction time data for the standardized measures. Furthermore, we wanted to keep our analyses consistent in terms of the measures analyzed, so we decided to examine accuracy only for the tasks we also had reaction time data for (our experimental measures). Our focus was on reaction time because we were interested in examining the speed with which individuals are able to segment and maintain sounds, as well as the speed with which individuals are able to map from a symbol to its meaning. All analyses were conducting in R (R Core Team, 2013).

The first research question examined the mechanistic relationship between phonological skill and Arabic number knowledge. It was hypothesized that there would be an indirect relationship between phonological skill and Arabic number knowledge with verbal counting skill underlying this relationship. Therefore, we conducted a mediation analysis with bootstrapping, with verbal counting skill as the mediating variable. Specifically, we bootstrapped (i.e.

resampled) the indirect effect from performance on the rhyme judgment task to performance on the Arabic number comparison task, through the mediating variable of counting skill. Five-thousand resamples were conducted to obtain a 95% CI for the indirect effect.

The second research question examined the mechanistic relationship between semantic and Arabic number knowledge. To parallel our approach for examining our first research question, we conducted a bootstrapped mediation analysis with verbal counting skill as the mediating variable. However, this time we did not expect to find an indirect relationship with counting skill underlying the relationship between semantic and Arabic number knowledge. Rather, we hypothesized that there would be a direct relationship between semantic and Arabic number knowledge. Therefore, we decided to bootstrap (i.e. resample) the direct effect from performance on the meaning judgment task to performance on the Arabic number comparison task. Five-thousand resamples were conducted to obtain a 95% CI for the direct effect.

Results

Primary Analyses: Reaction Time

Phonological Skill Hypothesis. Our primary analysis examining our hypothesized indirect effect focused on task reaction times, with the count to 10 (i.e. counting speed) task serving as our measure of verbal counting skill. Descriptive statistics and correlations for our phonological skill analysis are presented in Table 1. A representation of the phonological skill

Table 1

Primary Phonological Skill Analysis Descriptive Statistics and Task Correlations

	N	M	SD	Range			
Age (Months; 7 Female)	14	56.82	2.43	52.14–59.53			
Task					1	2	3
1. Rhyme Judgment	14	1998	462	1403–3060	—		
2. Small Distance Arabic Number Comparison	14	2030	363	1440–2685	.35	—	
3. Count to 10	14	3722	2034	1623–7073	-.30	.16	—

Note. Numbers represent task reaction times in milliseconds. No significant correlations.

hypothesis model is presented in Figure 3A. No significant paths were found in this model, and

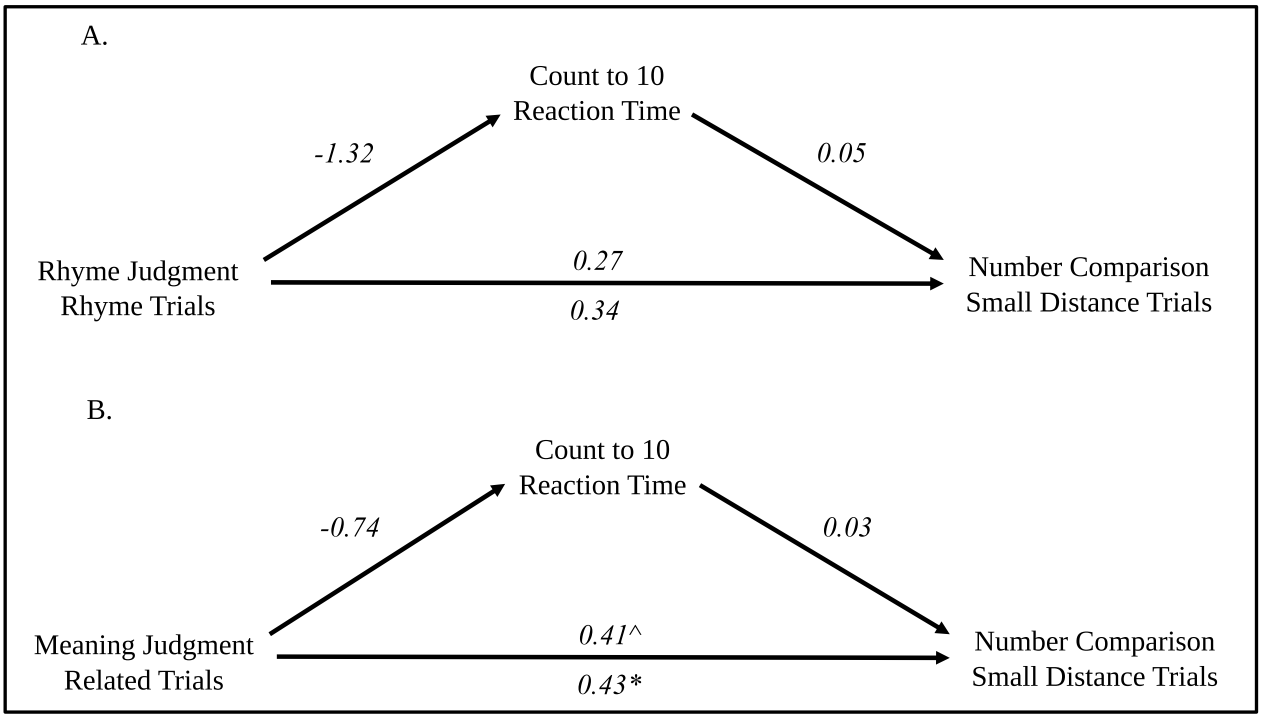

Figure 3. A. Phonological skill hypothesis mediation model using task reaction time. B. Semantic knowledge hypothesis mediation model using task reaction time. Values represent slope estimates. ^ p = .05, * p < .05. *Note*. Of the two numbers associated with the bottom line in the model, the number above the line represents the total effect (C path), while the number below the line represents the direct effect (C' path).

critically, our bootstrapped analysis – where we conducted 5,000 samples of the indirect effect – did not provide support for a mediation (Indirect effect = -0.07, bootstrapped 95% CI [-0.32, 0.16]).

Semantic Knowledge Hypothesis. Our primary analysis examining our hypothesized direct relationship between semantic knowledge and Arabic number knowledge focused on reaction time. Descriptive statistics and correlations for our semantic knowledge analysis are presented in Table 2. A representation of the semantic knowledge hypothesis model is presented

Table 2

Primary Semantic Knowledge Analysis Descriptive Statistics and Task Correlations

	N	M	SD	Range			
Age (Months; 9 Female)	15	56.91	2.46	52.14–59.53			
Task					1	2	3
1. Related Judgment	15	2024	431	1152–2832	—		
2. Small Distance Arabic Number Comparison	15	1949	346	1434–2685	.51^	—	
3. Count to 10	15	3607	1945	1623–7073	-.17	.10	—

Note. Numbers represent task reaction times in milliseconds. ^ $p = .053$.

in Figure 3B. The total effect from average reaction time on related trials to average reaction time on small distance Arabic number comparison trials was trending towards significance (B = 0.41, $t(13) = 2.13$, $p = .053$). However, when controlling for count to 10 reaction time, a significant direct effect was found from average reaction time on related trials to average reaction time on small distance Arabic number comparison trials (B = 0.43, $t(12) = 2.18$, $p < .05$). All other paths in the model were not significant. Lastly, consistent with our direct effect trend ($p < .05$), our bootstrapped analysis of the direct effect supported a direct relationship between average reaction time on related meaning judgment trials and average reaction time on small distance Arabic number comparison trials (Direct effect = 0.43, bootstrapped 95% CI [0.10, 0.78]).

Secondary Analyses: Accuracy

Although our primary interest was in the speed with which individuals were able to complete our tasks, previous work examining the relationships between phonological skill, semantic knowledge, and symbolic number skills has focused primarily on task accuracy (Krajewski & Schneider, 2009; Purpura & Ganley, 2014). To add to this literature, as well as add to our findings from our primary analyses, our secondary analyses examined our hypotheses in relation to task accuracy, with the count to 30 task (i.e. counting accuracy) serving as our measure of verbal counting skill.

Phonological Skill Hypothesis. Descriptive statistics and correlations are presented in Table 3. A representation of the model is presented in Figure 4A. We found a

Table 3

Secondary Phonological Skill Analysis Descriptive Statistics and Task Correlations

	N	M	SD	Range			
Age (Months; 7 Female)	15	56.34	2.99	49.64–59.53			
Task					1	2	3
1. Rhyme Judgment	15	85.67	11.90	54–100	—		
2. Small Distance Arabic Number Comparison	15	76.00	10.04	55–95	.13	—	
3. Count to 30	15	96.44	5.41	80–100	.69**	.07	—

Note. Numbers represent task accuracy out of 100%. ** $p < .01$.

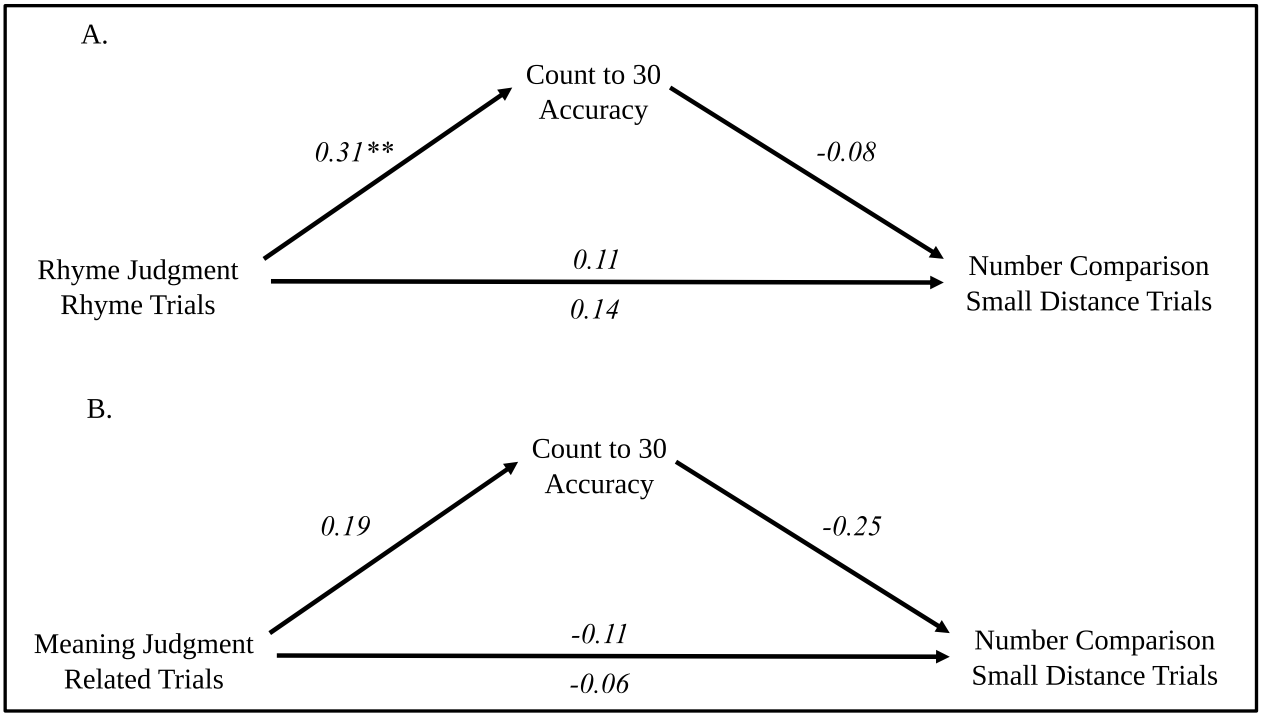

Figure 4. A. Phonological skill hypothesis mediation model using task accuracy. B. Semantic knowledge hypothesis mediation model using task accuracy. Values represent slope estimates. ** p < .01. *Note*. Of the two numbers associated with the bottom line in the model, the number above the line represents the total effect (C path), while the number below the line represents the direct effect (C' path).

significant relationship from accuracy on rhyme trials to count to 30 accuracy (B = 0.31, $t(13) = 3.43$, $p < .01$); however, all other paths were not significant. In addition, our bootstrapped analysis of the indirect effect indicated no mediation (Indirect effect = -0.03, bootstrapped 95% CI [-0.67, 0.61]).

Semantic Knowledge Hypothesis. Descriptive statistics and correlations are presented in Table 4. A representation of the model is presented in Figure 4B. No significant paths were

Table 4

Secondary Semantic Knowledge Analysis Descriptive Statistics and Task Correlations

	N	M	SD	Range			
Age (Months; 7 Female)	16	56.49	2.99	49.64–59.53			
Task					1	2	3
1. Related Judgment	16	77.88	11.00	54.00–100	—		
2. Small Distance Arabic Number Comparison	16	74.69	11.18	50.00–95.00	-.10	—	
3. Count to 30	16	96.88	5.09	80.00–100	.41	-.14	—

Note. Numbers represent task accuracy out of 100%. No significant correlations.

found, and our bootstrapped analysis of the direct effect did not provide support for a direct relationship between accuracy on related meaning judgment trials and accuracy on small distance Arabic number comparison trials (Direct effect = -0.06, bootstrapped 95% CI [-0.82, 0.87]).

Discussion

The goal of this study was to identify how phonological skill and semantic knowledge differentially relate to symbolic number skills by examining the mechanisms underlying these relationships. We found no evidence supporting our phonological skill hypothesis; which was that counting ability, particularly counting speed, would mediate the relationship between phonological skill and Arabic number comparison skill. However, we did find evidence in support of our semantic knowledge hypothesis; which was that there would be a direct relationship between semantic knowledge and Arabic number comparison skill.

Semantic Knowledge and Symbolic Mapping Skill

When making a semantic judgment, we must map from the symbol to that symbols meaning. When identifying and comparing Arabic numerals, a similar process is involved; we must map from the symbol (i.e. Arabic numeral) to its meaning (i.e. quantity). This similar shared underlying mechanism led us to predict a direct relationship between semantic and Arabic number knowledge. Previous research has found a positive relationship between vocabulary knowledge and tasks involving a quantitative understanding of numbers (Negen & Sarnecka, 2012; Purpura & Ganley, 2014), leading us to hypothesize that those who perform better on our semantic meaning judgment task would perform better on our Arabic number comparison task.

Our results provide support for this relationship and proposed mechanism. When examining the direct effect between reaction time on related semantic meaning judgment trials and small distance Arabic number comparison trials, our confidence interval did not contain zero. This indicates that over repeated sampling we will find a significant, positive relationship between reaction time on semantic related judgments and small distance Arabic number comparison judgments. This finding suggests that it may be the speed and efficiency with which individuals can map from symbol to meaning that is underlying the relationship between semantic knowledge and Arabic number skills. Furthermore, as our knowledge for words precedes our knowledge of Arabic numbers, it may be that the mapping skills developed in relation to words could be playing a role in the development of our Arabic number knowledge. However, since our design was cross-sectional it is difficult to make any directional claims about these relationships.

Phonological Skill and Verbal Number Skill

Increased phonological skill involves a greater ability to segment and maintain sounds over time – skills that are also important in successful counting (Koponen et al., 2007; Krajewski & Schneider, 2009; Moll, Snowling, Göbel, et al., 2015; Simmons & Singleton, 2007; Soto-calvo et al., 2019). Furthermore, increased counting ability has been found to be related to increased number comparison ability (Aunio & Niemivirta, 2010; Krajewski & Schneider, 2009; Purpura & Ganley, 2014). Therefore, we predicted an indirect relationship between phonological skill and Arabic number comparison skill. We hypothesized that those with greater phonological skill will be more efficient in the segmentation and maintenance of sounds over time, and that this increased efficiency will transfer to faster verbal counting; and, that increased verbal counting ability will be related to greater Arabic number comparison performance.

Our results did not support this proposed indirect relationship. Our primary analysis examining the relationship between reaction time on rhyme judgments, counting to 10 (i.e. counting speed), and small distance Arabic number comparisons, revealed no indirect effect with count to 10 reaction time as our mediating variable. In addition, we found no significant relationships between any variables in our analysis. Similarly, our secondary analysis examining the relationships between accuracy on rhyme judgments, counting to 30 (i.e. counting accuracy), and small distance Arabic number comparisons revealed no indirect effect with count to 30 accuracy as our mediating variable. However, consistent with previous research (Koponen et al., 2007; Krajewski & Schneider, 2009; Soto-calvo et al., 2019), we did find a significant positive relationship between accuracy on rhyme judgments and count to 30 accuracy.

The findings from our phonological skill analyses did not support our hypothesis, which was based on previous evidence showing relationships between phonological skill and counting, as well as counting and symbolic number comparison ability. However, this is likely due to the fact that we were underpowered given our limited sample size. Another possible explanation for the lack of a mediation could be our experimental design, specifically as it relates to examining phonological skill. For both language tasks, participants were presented with pairs of audio-visual stimuli. Previous studies that have found a relationship between phonological skill and symbolic number skills have used both auditory and visual stimuli, however, unlike our study, the auditory and visual stimuli were not presented simultaneously (Koponen et al., 2007; Krajewski & Schneider, 2009). Therefore, when making rhyme judgments during our task, participants may not be attending to the auditory information. Rather, they could have been focusing on, or distracted by, the associated visual stimuli. Although previous work has found a preference for auditory stimuli in children as young as 4-years-old (Sloutsky & Napolitano, 2003), this conclusion has been challenged due to the imbalance in complexity and familiarity between auditory and visual stimuli. Napolitano & Sloutsky (2004) found that when familiarity of stimuli was considered, 4-year-olds preference for auditory stimuli weakened. In fact, familiarity with stimuli drove the preference for that modality (i.e. children more familiar with visual stimuli preferred visual stimuli), suggesting that preference for stimuli may be based solely on familiarity. Further research more closely equating complexity and familiarity of auditory and visual stimuli found that preschool-aged children exhibited a strong preference for visual information (Noles & Gelman, 2011). Based off these previous findings, our current results would make more sense. Children may be attending to the visual stimuli, and thus not engaging the sound segmentation and maintenance mechanisms we proposed to be reflective of phonological skill and underlying the relationship with Arabic number comparison ability.

Verbal to Symbolic Mapping Skills: Counting to Comparison

Although we did not find a significant relationship between reaction time on rhyme judgments and count to 10 reaction time, we did find a significant relationship between accuracy on rhyme judgments and count to 30 accuracy. This suggests that at this age there is a relationship between phonological skill and verbal counting, but that this relationship may not involve the speed with which individuals are able to segment and maintain sounds over time. Therefore, the lack of a mediation in our phonological skill hypothesis may have more to do with the lack of a relationship between verbal counting ability and Arabic number comparison performance, which surprisingly was missing across all analyses.

The lack of a relationship between verbal counting speed and average reaction time on small distance comparison trials is a little more understandable as previous work has focused primarily on accuracy (Krajewski & Schneider, 2009; Purpura & Ganley, 2014); however, the lack of a relationship between verbal counting accuracy and accuracy on the small distance Arabic number comparison trials was unexpected. One reason for these null relationships could be due to the fact we focused on the most challenging trials within the Arabic number comparison task (i.e. small distance comparisons when the numbers differ by only one or two). We felt that analyzing the small distance comparison trials would provide us with the best measure of participants' Arabic number knowledge and the underlying mechanism of symbol to meaning mapping. However, previous work with children in this age range has not appeared to limit their analyses to include only the small distance trials (Aunio & Niemivirta, 2010; Purpura & Ganley, 2014).

An alternative explanation could be our choice of counting task. Although commonly used to assess counting ability, verbal counting may not be the most appropriate counting measure in relation to an Arabic number comparison task which requires a quantitative understanding of numbers, and the ability to map from a symbol to its associated quantity. Rather, verbal counting may be more appropriate when dealing with a number word comparison task, as both of these tasks rely more heavily on memory for number words. A study conducted by Jiménez Lira, Carver, Douglas, and LeFevre (2017) with preschool-aged children found a significant positive relationship between verbal counting ability and number word comparison performance, but not verbal counting and Arabic number comparison performance. A more appropriate counting measure when examining the relationship with Arabic number knowledge may be a task requiring an understanding of how numbers in the count sequence relate to one another – a skill more closely related to making Arabic number comparison judgments. Previous work has used measures of cardinality and counting subsets to assess counting ability as it relates to understanding sets of items associated with the count sequence (Moll, Snowling, Göbel, et al., 2015; Purpura & Ganley, 2014; Purpura & Lonigan, 2014); while others have used counting forwards or backwards from a given number, or asking participants what number they would get if they counted a certain amount of numbers forward from another number (e.g. "What number would you get if you counted four numbers forward from two?"; Koponen et al., 2007; Purpura, Hume, Sims, & Lonigan, 2011). Opposed to verbal counting where participants are told to begin at one and count as high as they can, these tasks require a greater understanding of the relational properties of numbers in the count sequence. Therefore, these counting tasks may be more appropriate when examining the relationship with Arabic number comparison performance – a task that requires relational and quantitative understandings of numbers.

Limitations

Contrary to the language tasks where the audio-visual stimuli pairings were presented sequentially, in the Arabic number comparison task, numbers were presented simultaneously. The reasoning for this was to decrease task demands place on the children; however, this meant that for some trials, the number on the left side of the screen was larger. In these instances, the correct (or "yes") response would have been for participants to press the red button on our response device, as the red button aligned with the left side of the screen. Unlike the language tasks where the green button was always the "yes" response (i.e. the items rhymed or were related), having a mix of red and green "yes" (i.e. the larger number) responses for the Arabic comparison task could have introduced confusion and potentially skewed our reaction time data.

Participants may have needed more time to determine the correct response, especially in cases where the “yes” response would have been to press the red button.

A second limitation involves how our language tasks present both auditory and visual stimuli. Given the young age range of our sample, we wanted to provide as much information to the participants in order to aid in their rhyme and meaning judgments. However, this limits our ability to determine what stimuli the participants were attending to when making their judgments. This makes it challenging for us to conclude that the mechanisms we proposed to be underlying these relationships are in fact the mechanisms the children are tapping into when completing our language tasks.

Lastly, as noted in the Methods section, due to the COVID-19 we needed to stop data collection, resulting in an underpowered sample. Given the low number of participants, this made it difficult to detect any possible relationships in our analyses, and limits our confidence in the results we did find.

Future Directions

There are a few ways to build upon this research moving forward. First, we could restructure the Arabic number comparison task to mirror the response process involved in the language tasks. This would involve making it so the “yes” response will always be associated with the green button. Previous work has employed a design whereby individuals need to indicate whether the second of two sequentially presented numbers was larger (Ansari, Lyons, Van Eimeren, & Xu, 2007). We could employ a similar design and have children press the green button if the second number presented was larger, and press the red button if the second number presented was smaller.

Second, we could redesign each language task to include only the stimuli we believe is involved in completing the tasks according to our proposed underlying mechanisms. Having multiple sources of stimuli make it difficult for us to be sure what children are attending to when making their judgments. By removing the visual stimuli from the rhyme judgment task, we can isolate the mechanisms – sound segmentation and maintenance – we believe to be underlying the relationship between phonological skill and symbolic number skills. Similarly, by removing the auditory stimuli from our meaning judgment task we will be isolating the symbol to meaning mapping mechanism we propose to be underlying the relationship between semantic knowledge and Arabic number skills. These changes will be specifically important for examining our phonological skill hypothesis as we currently do not have support for our proposed mediation.

Thirdly, we could include counting tasks that involve manipulation of the count sequence, object counting, or cardinality. These tasks reflect a relational understanding of the numbers in the count sequence, a skill that more closely resembles the type of quantitative understanding necessary to complete our Arabic number comparison task.

Finally, it would be important to examine these relationships longitudinally. Currently we have evidence in support of our semantic knowledge hypothesis, however what remains to be seen is how these specific skills are related over time. Previous work has examined the longitudinal relationships between phonological awareness, vocabulary knowledge, and numeracy skills as a whole (Purpura et al., 2011, 2017), but it is unclear how these language skills relate longitudinally to specific aspects of numeracy such as Arabic number comparison ability – a skill that has been argued to be the basis on which we determine if an individual understands numbers (McCloskey, 1992). By examining these relationships longitudinally, we will further our understanding of how these skills are related, with a specific interest on how early language skills may be impacting the development of Arabic number knowledge.

Conclusions

We did not find support for our phonological skill hypothesis. The relationship between phonological skill and Arabic number knowledge may not be mediated by verbal counting skill and rely on the ability to segment and maintain sounds over time. However, there are some limitations to this study that could have prevented us from appropriately measuring these relationships. On the other hand, we did find support for our semantic knowledge hypothesis. It appears that the relationship between semantic and Arabic number knowledge may have to do with the ability to efficiently map a symbol to that symbols meaning.

References

Ansari, D., Lyons, I. M., Van Eimeren, L., & Xu, F. (2007). Linking Visual Attention and Number Processing in the Brain: The Role of the Temporo-parietal Junction in Small and Large Symbolic and Nonsymbolic Number Comparison. *Journal of Cognitive Neuroscience*, *19*(11), 1845–1853.

Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*. https://doi.org/10.1016/j.lindif.2010.06.003

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educational Psychology*, *96*(4), 699–713. https://doi.org/10.1037/0022-0663.96.4.699

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. Behavior Research Methods, 39(3), 445–459.

Bates, E., Andonova, E., D'Amico, S., Jacobsen, T., Kohnert, K., Lu, C., ... & Iyer, G. (2000). Introducing the CRL international picture-naming project (CRL-IPNP). *Center for Research in Language Newsletter*, *12*(1), 1-14.

Boersma, P. & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.08, retrieved 7 January 2019 from http://www.praat.org/

Brooks, B. L., Sherman, E. M., & Strauss, E. (2009). NEPSY-II: A developmental neuropsychological assessment. *Child Neuropsychology*, *16*(1), 80-101.

Chiara Passolunghi, M., Vercelloni, B., & Schadee, H. (2007). The precursors of mathematics learning: Working memory, phonological ability and numerical competence. *Cognitive Development*, *22*, 165–184. https://doi.org/10.1016/j.cogdev.2006.09.001

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

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*. Pearson Assessments.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160.

Geary, D. C., VanMarle, K., Chu, F. W., Hoard, M. K., & Nugent, L. (2019). Predicting age of becoming a cardinal principle knower. *Journal of Educational Psychology*, *111*(2), 256–267. https://doi.org/10.1037/edu0000277

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British journal of mathematical and statistical psychology*, *67*(3), 451-470.

Hornburg, C. B., Schmitt, S. A., & Purpura, D. J. (2018). Relations between preschoolers' mathematical language understanding and specific numeracy skills. *Journal of Experimental Child Psychology*, *176*, 84–100. https://doi.org/10.1016/j.jecp.2018.07.005

Hurst, M., Anderson, U., Cordes, S., Hurst, M., Anderson, U., & Cordes, S. (2017). Mapping Among Number Words , Numerals , and Nonsymbolic Quantities in Preschoolers Mapping Among Number Words , Numerals , and Nonsymbolic Quantities in Preschoolers. *Journal of Cognition and Development*, *18*(1), 41–62. https://doi.org/10.1080/15248372.2016.1228653

Jiménez Lira, C., Carver, M., Douglas, H., & LeFevre, J.-A. (2017). The integration of symbolic and non-symbolic representations of exact quantity in preschool children. *Cognition*, *166*, 382–397. https://doi.org/10.1016/j.cognition.2017.05.033

Koponen, T., Aunola, K., Ahonen, T., & Nurmi, J. E. (2007). Cognitive predictors of single-digit and procedural calculation skills and their covariation with reading skill. *Journal of Experimental Child Psychology*. https://doi.org/10.1016/j.jecp.2007.03.001

Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual-spatial working memory, and preschool quantity-number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology*, *103*, 516–531. https://doi.org/10.1016/j.jecp.2009.03.009

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

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, *44*(1–2), 107–157. https://doi.org/10.1016/0010-0277(92)90052-J

Moll, K., Snowling, M. J., G, S. M., & Hulme, C. (2015). Early language and executive skills predict variations in number and arithmetic skills in children at family-risk of dyslexia and typically developing controls. *Learning and Instruction*, *38*, 53–62. https://doi.org/10.1016/j.learninstruc.2015.03.004

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development*, *75*(6), 1850–1870. https://doi.org/10.1111/j.1467-8624.2004.00821.x

Negen, J., & Sarnecka, B. W. (2012). Number-Concept Acquisition and General Vocabulary

Development. *Child Development*, *83*(6), 2019–2027. https://doi.org/10.1111/j.1467-8624.2012.01815.x

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. 1998.

Noles, N. S., & Gelman, S. A. (2011). Preschool children and adults flexibly shift their preferences for auditory versus visual modalities, but do not exhibit auditory dominance. *Journal of Experimental Child Psychology*, *112*(3), 338–350. https://doi.org/10.1016/j.jecp.2011.12.002

Purpura, D. J., & Ganley, C. M. (2014). Working memory and language: Skill-specific or domain-general relations to mathematics? *Journal of Experimental Child Psychology*, *122*, 104–121. https://doi.org/10.1016/j.jecp.2013.12.009

Purpura, D. J., Hume, L. E., Sims, D. M., & Lonigan, C. J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. *Journal of Experimental Child Psychology*, *110*(4), 647–658. https://doi.org/10.1016/J.JECP.2011.07.004

Purpura, D. J., Logan, J. A. R., Hassinger-Das, B., & Napoli, A. R. (2017). Why do early mathematics skills predict later reading? The role of mathematical language. *Developmental Psychology*, *53*(9), 1633–1642. Retrieved from https://psycnet.apa.org/buy/2017-32731-001

Purpura, D. J., & Lonigan, C. J. (2014). Early Numeracy Assessment: The Development of the Preschool Early Numeracy Scales. *Early Education and Development*, *26*(2), 286–313. https://doi.org/10.1080/10409289.2015.991084

Purpura, D. J., & Reid, E. E. (2016). Mathematics and language: Individual and group differences in mathematical language skills in young children. *Early Childhood Research*

Quarterly, *36*, 259–268. https://doi.org/10.1016/J.ECRESQ.2015.12.020

Reynolds, C. R., & Kamphaus, R. W. (2015). *RIAS-2 (Reynolds Intellectual Assessment Scales) and the RIST-2 (Reynolds Intellectual Screening Test)*. Psychological Assessment Resources.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*. Psychology Software Incorporated.

Simmons, F. R. (2002). *Relationships between Number Skills and Cognitive Abilities in people with Specific Arithmetic Difficulties and people with Dyslexia*. (Doctoral dissertation, University of Hull).

Simmons, F. R., & Singleton, C. (2007). Do Weak Phonological Representations Impact on Arithmetic Development? A Review of Research into Arithmetic and Dyslexia. *Dyslexia*, (14), 77–94. https://doi.org/10.1002/dys.341

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a Picture Worth a Thousand Words? Preference for Auditory Modality in Young Children. *Child Development*, *74*(3), 822–833. https://doi.org/10.1111/1467-8624.00570

Soto-calvo, E., Simmons, F. R., Adams, A., Francis, H. N., Giofre, D., Simmons, F. R., … Francis, H. N. (2019). Pre-Schoolers ’ Home Numeracy and Home Literacy Experiences and Their Relationships with Early Number Skills : Evidence from a UK Study Pre-Schoolers ’ Home Numeracy and Home Literacy Experiences and Their Relationships with Early Number Skills : Evidence. *Early Education and Development*, *0*(0), 1–24. https://doi.org/10.1080/10409289.2019.1617012

Stoianov, I. P., Zorzi, M., & Umilta, C. (2003). A Connectionist Model of Simple Mental Arithmetic. *Proceedings of EuroCogSci 03*, (May 2014), 313–318.

https://doi.org/10.4324/9781315782362-62

Team, R. C. (2013). R: A language and environment for statistical computing.

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 481-487.

Weiss, Y., Cweigenberg, H. G., & Booth, J. R. (2018). *Neural specialization of phonological and semantic processing in young children*. https://doi.org/10.1002/hbm.24274

Zorzi, M. & Butterworth, B. (1999). A computational model of number comparison. In: M. Hahn & S.C. Stoness (Eds.), Proceedings of the Twenty First Annual Conference of the Cognitive Science Society (p. 778- 783). Mahwah (NJ): Erlbaum.