Individual Differences in Executive Function and Learning: Role of Type of Knowledge and Instructional Approaches

A Dissertation
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Amanda Grenell

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Stephanie M. Carlson, Ph.D.

June 2020

Acknowledgements

First, I would like to thank all of the children, families, and childcare centers that participated in these dissertation studies. Without them, this research would not have been possible. Completing the data collection for these dissertation studies would have also not been possible without the dedication and hard work of my data collection team including Jasmine Ernst, Kerry Houlihan, Ryan Anderson, Caitlin Petersen, Katie Magnan, Krista Garrett, Maggie Ryan, Teddy Anderson, Katlyn Mccarty, Julia Cai, Audrey Kirschling, Natalie Carlson, and Samuel Kroening. This research was supported by two Institute of Child Development small grants.

To my advisor, Dr. Stephanie Carlson, thank you for allowing me to pursue questions that I was passionate about and for teaching me how to be a better researcher. Thank you also to my committee members: Dr. Melissa Koenig, Dr. Michèle Mazzocco, and Dr. Sashank Varma, for all your helpful feedback and insights throughout the dissertation process. I would also like to thank my lab- and cohort-mates for all their support and friendship over the last six years.

To my mom, thank you for teaching me the value of hard work and for always encouraging me never to give up. To my sisters, Rachel, Jessica, and Krista, thank you for always being just a phone call away and for always believing in me. Lastly, to my partner, Zachary, thank you for being willing to move to Minnesota with me so I could earn my Ph.D. and for loving and supporting me through all the highs and lows of graduate school.

Abstract

Executive function (EF) predicts children's academic achievement; however, less is known about the relation between EF and the actual learning process. Furthermore, more research is needed to better understand how different aspects of the learning environment interact with EF to influence learning. The current dissertation includes two studies to examine how two aspects of the learning environment (the type of knowledge and instructional approaches) influence the relation between EF and learning. Study 1 examined how aspects of the material to be learned—the type of information and the amount of conflict between the content to be learned and children's prior knowledge – influence the relation between individual differences in EF and learning. Typically developing 4-year-olds (*N* =61) completed a battery of EF tasks and several animal learning tasks that varied on the type of information being learned (factual vs. conceptual) and the amount of conflict with the learner's prior knowledge (no prior knowledge, no conflicting prior knowledge, conflicting prior knowledge). Individual differences in cool EF predicted children's overall learning, controlling for age, verbal IQ, and prior knowledge. Cool EF skills predicted children's conceptual learning, whereas motor inhibition skills predicted children's factual learning. Additionally, individual differences in EF mattered more for children's learning of information that conflicted with their prior knowledge. Study 2 extended the findings from Study 1 by examining how individual differences in EF predicted children's expression and construction of knowledge in a new science domain and whether EF moderated the effectiveness of different instructional approaches (direct instruction vs. discovery learning). Typically developing 4-and 5-year-olds (*N* =93) were randomly assigned to a

Direct Instruction, Discovery Learning, or Control Condition. A pre-post-test design was used to measure change in children's knowledge of sinking and floating before and after the instructional groups received three instructional sessions. The participants also completed a battery of EF tasks and standardized measures to assess their non-verbal and verbal IQ and their literacy and math achievement. Results showed EF was not a significant predictor of children's expression of their sinking and floating knowledge at pre-test, controlling for age, verbal and non-verbal IQ, and SES. EF was also not a significant predictor of children's construction of knowledge controlling for the covariates. However, exploratory analyses revealed there were trends for interactions between EF and SES to predict children's prior knowledge about sinking and floating and between EF and Prior Knowledge to predict children's learning when collapsed across instructional groups. We also did not find evidence that individual differences in EF moderated the effectiveness of the different instructional approaches for children's sinking and floating learning. Taken together, these findings suggest that individual differences in EF should be considered when creating personalized instructional materials and interventions to optimize preschoolers' learning.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

A large body of research shows that executive function (EF) — higher-order neurocognitive skills such as working memory, cognitive flexibility and inhibitory control — is related to and predicts school readiness and academic achievement (e.g., Allan, Hume, Allan, Farrington, & Lonigan, 2014; Best, Naglieri, & Miller, 2011; Blair & Razza, 2007; Jacob & Parkinson, 2015). Additionally, intervention research shows that EF can be trained and improved in certain contexts (e.g., Diamond & Ling, 2016). This has motivated researchers and educators to push for targeting EF as a way to improve young children's success in school and to help narrow the income-achievement gap. Despite this desire to use EF interventions to improve academic achievement, there is little support for a causal relation between EF and academic outcomes (Jacob & Parkinson, 2015) and more research is needed to determine *how* EF relates to learning. In particular, more research is needed to better understand how specific aspects of the information being learned and the instructional approaches used to present the information to the learner influence the relation between individual differences in EF and learning.

Executive Function

EF skills refer to higher order cognitive skills used during goal-directed behaviors (Carlson, Zelazo, & Faja, 2013). The development of EF coincides with the development of the prefrontal cortex (Casey, Giedd, & Thomas, 2000). There is rapid development in EF during the preschool period and EF continues to improve through adolescence and young adulthood before beginning to decline (Carlson et al., 2013). Research with older children and adults has shown there are three main components of EF: working memory, cognitive flexibility, and inhibitory control (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). However, research with younger children suggests that EF consists of one unitary construct in preschool (e.g., Wiebe et al., 2011) that differentiates by adulthood into two main factors, working memory and cognitive flexibility, with inhibitory control shared between the two (Miyake & Friedman, 2012). In addition to academic achievement, early EF predicts important outcomes such as wealth and health in adulthood (Casey et al., 2011; Moffitt et al., 2011).

EF and Academic Achievement

Most of the research on the role of EF in learning has focused on the relation between EF and standardized achievement measures in math, literacy, and science (e.g., Allan et al., 2014; Clements, Sarama, & Germeroth, 2016; Jacob & Parkinson, 2015; Raghubar, Barnes, & Hecht, 2010; Yeniad, Malda, Mesman, Van Ijzendoorn, & Pieper, 2013). The main finding from this body of literature is that higher EF skills are associated with and predictive of higher academic achievement, even when controlling for IQ. However, academic achievement is only one way to measure students' success in school (Neuenschwander, Rothlisberger, Cimeli, & Roebers, 2012). For example, grades and classroom behaviors that are optimal for learning also are indices of success in school. Additionally, it is important to consider how academic achievement as measured by standardized achievement tests may be different from learning or constructing new knowledge (Modrek, Kuhn, Conway, & Arvidsson, 2019; Neuenschwander et al., 2012). For example, to score highly on an academic achievement test, students need to express knowledge they have already acquired after the learning process has occurred. In a recent study, Modrek and colleagues (2019) measured middle school students' learning during a science inquiry task and collected their scores on English and math standardized tests given by their schools. They also measured different types of self-regulation such as EF and behavior regulation in the classroom. Although students' inquiry learning and academic achievement scores were positively correlated, different types of self-regulation predicted children's learning versus academic achievement: Children's EF predicted their learning on the inquiry task, but not their academic achievement scores, whereas children's behavior regulation did not predict their inquiry learning scores but did predict their performance on the standardized math test. This study suggests that learning and academic achievement may be different constructs and points to the importance of considering other ways in which EF might influence learning or success in school.

Potential Roles that EF Plays in Learning

There are several different roles that EF could play in learning. One possibility is that EF influences learning indirectly by allowing students to engage in behaviors that are optimal for learning. Learning-related behaviors increase opportunities for children to be engaged in instructional activities and include participating, being able to work together successfully with teachers and peers, refraining from behavior that would disrupt learning activities, and paying attention (Fantuzzo, Perry, & McDermott, 2004; Nesbitt, Farran, & Fuhs, 2015; Razza, Martin, & Brooks-Gunn, 2015). Research shows that children's EF is related to their learning-related behaviors (Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009) and that EF mediates the association between children's learning-related behaviors and their academic achievement (Baptista, Osório, Martins, Verissimo, & Martins, 2016; Nesbitt et al., 2015; Neuenschwander et al., 2012; Portilla, Ballard, Adler, Boyce, & Obradovic, 2014; Sasser, Bierman, & Heinrichs, 2015; but see Vitiello & Greenfield, 2017). The idea that EF allows learners to be ready to learn and to be compliant in the classroom is often discussed in professional development and policy-related materials as an important way that EF helps with success in school (e.g., Ackerman & Friedman-Krauss, 2017).

Another potential role that EF plays in learning is that EF skills are needed for children to “show what they know.” For example, if students who have learned how to add and subtract are completing a math worksheet that has a mix of addition and subtraction problems, they will need to be able to flexibly switch between them to accurately express their knowledge of both operations. Another way EF might influence the expression of knowledge is by dictating the types of strategies children use. For example, there have been studies showing that children select different strategies depending on their EF while solving arithmetic problems (Barrouillet & Lepine, 2005) and that children with higher EF are more likely to choose the more optimal arithmetic strategy than children with lower EF (Lemaire & Lecacheur, 2011).

At an even deeper level, EF skills might be important for processing new information and constructing new knowledge. For instance, studies have shown that EF, especially, inhibitory control, predicts children’s ability to use newly learned strategies that are more effective than previously learned strategies when solving algebraic and arithmetic problems (Khng & Lee, 2009; Robinson & Dubé, 2013). There is also evidence showing that individual differences in EF moderate the effectiveness of an intervention for gains in academic knowledge, with children with higher EF benefitting the most (Bascandziev, Powell, Harris, & Carey, 2016; Bascandziev, Tardiff, Zaitchik, & Carey, 2018; Fuchs et al., 2005, 2013a, 2013b, 2014; Laski & Dulaney, 2015; Miller, Rittle-Johnson, Loehr, & Fyfe, 2016; Rhodes et al., 2014; Rhodes et al., 2016). EF also might be important for determining which instructional strategies or approaches are most helpful when individuals are constructing new knowledge. For example, a study with high school students found that students with lower EF showed better transfer of learning from a chemistry computer simulation that was designed with a guided instruction approach, whereas students with higher EF exhibited better transfer when the simulation had an exploratory learning approach (Homer & Plass, 2014). Although there are several different potential roles that EF might play in children's success in school, the current dissertation focuses on those related to the learning process itself such as expressing and constructing new knowledge.

Expression vs. Construction of Knowledge

There is previous research showing that EF is related to the expression and construction of new knowledge. Individual differences in children's EF are related to their expression of conceptual knowledge related to science (Baker, Gjersoe, Sibielska-Woch, Leslie, & Hood, 2011; Vosniadou et al., 2015; Zaitchik, Iqbal, & Carey, 2014) and math (McKenzie, Bull, & Gray, 2003; Prager, Sera, & Carlson, 2016; Xenidou-Dervou et al., 2014), wherein children with higher EF demonstrate more knowledge in these domains, even when controlling for IQ or verbal ability. There is also evidence that EF plays a role in the construction of new conceptual knowledge in science (Bascandziev et al., 2016; Bascandziev et al., 2018) and math (Fuchs et al., 2005; Fuchs et al., 2013a; Fuchs et al., 2013b; Geary et al., 2019; Khng & Lee, 2009), even after controlling for IQ or verbal ability.

A few studies have examined the role of EF in both the expression and construction of knowledge within the same domain and group of participants (Bascandziev et al., 2016; Bascandziev et al., 2018; Miller et al., 2016) and demonstrated EF plays a role in both expressing knowledge students already have learned and constructing new knowledge. For example, one study investigated how individual differences in all three EF components (working memory, cognitive flexibility, and inhibition) influenced the effectiveness of a math intervention for preschoolers (Miller et al., 2016). During the intervention, the experimenter helped children recognize patterns and then children had a chance to replicate patterns with new materials. The authors examined how children's EF measured before the intervention predicted their expression of pattern knowledge before the intervention and their construction of pattern knowledge after the intervention. They found children's working memory and cognitive flexibility, but not inhibitory control, predicted their expression of pattern knowledge before the intervention, controlling for age, relational knowledge, and the other EF components. However, working memory was the only EF skill that predicted children's learning of pattern information after the intervention using the same controls (Miller et al., 2016). These studies suggest there might be different relations between EF and learning depending on whether the task requires children to express or to construct knowledge and that working memory might be especially important for constructing new knowledge (Bascandziev et al., 2016; Miller et al., 2016). This research is informative given that most of the correlational studies examining the relation between EF and learning used standardized academic achievement tests, which appear to align more with children's ability to express their existing knowledge. It also points to the need for more research on the role that EF plays in the construction of knowledge to better understand how learning interventions could be altered to be more beneficial for students, depending on their level of EF.

Current Studies

The two studies included in this dissertation aimed to add to our understanding of the relation between individual differences in EF and young children's learning. Study 1 focused on aspects of the material that children are asked to learn, such as the type of knowledge and the amount of conflict with prior knowledge. Study 2 focused on how individual differences in EF may influence the effectiveness of different instructional approaches, such as a direct instruction versus a discovery learning approach. Better understanding how these different aspects of the learning environment relate to the relation between EF and learning can inform the creation of more personalized and effective instructional materials and approaches that consider a child's EF to maximize every child's learning.

Chapter 2: Study 1

The goal of Study 1 was to add to our knowledge of how EF may be important for the construction of knowledge depending on (1) the type of information being learned and (2) the amount of conflict between the information to be learned and the learner's prior knowledge.

Type of Knowledge Being Learned

Although there is a growing body of evidence showing that individual differences in EF predict children's expression and construction of knowledge, the majority of this research has focused on learning conceptual information. This may be due in part to the fact that researchers began to explore whether EF was a domain-general skill that could be used to explain how conceptual change occurs (Bascandziev et al., 2018; Carlson & Moses, 2001; Carlson, Claxton, & Moses, 2015; Vosniadou, 2014; Tardiff et al., 2020; Zaitchik et al., 2016). Most conceptual change researchers posit that individuals start with a naïve theory that must be restructured to form more scientifically accurate theories (Carey, 2009; Shtulman & Valcarcel, 2012; Vosniadou, 2014; Wellman & Gelman, 1992). Children's naïve theories are formed based on their experiences in the world. When they start formal instruction in math and science, they are presented with information that conflicts with their naïve theories. When this occurs, children have to restructure their conceptual knowledge and categories to allow their initial theories to coexist with newer theories (Carey, 2009; Vosniadou, 2014). Some have proposed that children use Quinian bootstrapping to add knowledge to their original theories, change connections between different types of representations, and ultimately change their concepts (Carey, 2009; Zaitchik et al., 2016). According to this idea, EF might allow children to recognize conflicts between new knowledge and their own theories, inhibit old theories to use more accurate theories, and switch back and forth between different theories depending on the situation (Vosniadou, 2014; Vosniadou et al., 2015; Zaitchik et al., 2016).

Recently, researchers have begun to extend the study of the domain-general processes involved in conceptual change by making the distinction between different types of knowledge acquisition; some types of information might be easier to learn than others, thus relying on different cognitive skills. For example, the distinction between "knowledge enrichment" and "conceptual construction" has been used by some researchers (Bascandziev et al., 2018; Tardiff et al., 2020). Knowledge enrichment refers to instances where the information to be learned can be understood using concepts and words the learner already has, whereas conceptual construction refers to instances where the learner does not yet have the concepts or words to understand the content being learned. Evidence from a few studies looking at knowledge construction has suggested that EF is more strongly related to learning conceptual information than factual information (Bascandziev et al., 2018; Rhodes et al., 2014; Rhodes et al., 2016; Tardiff et al., 2020). For example, a vitalist biology intervention for early school-aged children was more beneficial for children with higher EF than lower EF when learning conceptual knowledge. However, individual differences in EF were not related to improvements in learning facts about animals. Instead, children's IQ and verbal ability were found to be related to their learning of factual information (Bascandziev et al., 2018). Moreover, a recent study investigated the expression of conceptual knowledge related to vitalist biology with 5-6-year-olds and included a standardized measure of factual knowledge. They found that EF related to children's vitalist biology conceptual knowledge, even after controlling for their factual knowledge scores (Tardiff et al., 2020).

There also have been a few training studies that included teaching and measuring adolescents' learning of both factual and conceptual knowledge in the domains of biology and chemistry (Rhodes et al., 2014; 2016). Rhodes and colleagues (2014) found students' factual learning about DNA was significantly correlated with their performance on a planning EF task (Tower of London). Their conceptual learning was significantly correlated with their performance on the planning task and a spatial working memory task (Rhodes et al., 2014). Contrary to the findings from Bascandziev and colleagues

(2018), the authors found that students' EF was related to both their conceptual and factual learning after the intervention, but they did not control for verbal ability or IQ (Rhodes et al., 2014). Another training study by Rhodes and colleagues (2016) investigated the relation between 12-13-year-olds' EF and factual and conceptual learning about chemistry. The researchers found a significant correlation between the students' conceptual learning and their performance on EF tasks that measured planning and spatial working memory, controlling for age and verbal ability. In this case, consistent with the study by Bascandziev and colleagues (2018), children's EF was not a significant predictor of their learning of factual information (controlling for age and verbal ability). However, Rhodes and colleagues (2016) did not account for children's prior knowledge so it is unclear whether they were measuring gains in knowledge due to the intervention or capturing differences in pre-existing knowledge.

These studies suggest that the relation between individual differences in EF and learning might depend on whether children are asked to learn factual or conceptual knowledge, with a stronger relation for the latter. However, more research is needed to see if these findings replicate, given the majority focused on vitalist biology specifically and other studies did not control for IQ (Rhodes et al., 2014) or students' prior knowledge related to the intervention (Rhodes et al., 2014; 2016). Therefore, the current study included both factual and conceptual learning tasks and controlled for prior knowledge and verbal IQ, to better understand how the role of EF might differ depending on the type of information children are expected to learn.

Conflict with Prior Knowledge

Lastly, we considered whether overcoming conflicts with prior knowledge could be a reason why EF skills are related to conceptual change. Theoretically, information processing skills are essential for progressing through the stages of conceptual development. EF allows individuals to engage in the type of information processing that is necessary to hold two ways of thinking about something in mind and to think about them in conflict with each other. There are at least three different relations between prior knowledge and new information that can occur during learning (Chi, 2008). One possibility is that the learner has no prior knowledge and the material being learned is completely novel. It is also possible the learner has some prior knowledge related to the material to be learned, but the new information does not conflict with their prior knowledge and instead adds to it. A third possibility is the new information conflicts with learners' prior knowledge and they must overcome or revise their prior knowledge for learning to occur. These three common conditions under which children learn new information provide a means to further explore the role that EF plays in learning, given they might rely on EF skills to different degrees.

Previous research has focused on two of these possibilities: no conflict with prior knowledge (knowledge enrichment) and conflict with prior knowledge (conceptual change or construction). However, studies have not included tasks in which children have no prior knowledge. Therefore, the current study created learning tasks that had a range of conflict with the learner's prior knowledge (no prior knowledge, no conflicting prior knowledge, and conflicting prior knowledge) to see if the role of EF in learning increased as the amount of conflict in the task increased. By manipulating how much conflict is present in the learning tasks, we were able to examine how increasing the information processing demands of the task by requiring children to overcome prior knowledge influences the relation between EF and learning. To more fully understand how individual differences in EF might be contributing to learning, we aimed to investigate aspects of the learning process that have not been studied extensively in previous research, such as characteristics of the content to be learned, and to do so in a critically important period of development just prior to entering formal schooling.

Current Study

In Study 1, 61 4-year-olds completed a series of EF, learning, and IQ tasks. A within-person design was used such that all children received all the learning tasks to examine how individual differences in EF related to their learning of information about animals that differed in our two main variables of interest: (1) type of knowledge and (2) amount of conflict with prior knowledge. The current study had three aims to better understand the relation between individual differences in young children's EF and learning. The first was to determine if individual differences in children's EF predicted their overall learning performance. We hypothesized that children's EF skills would be a significant predictor of their overall learning, with children with higher EF skills having higher learning scores. The second aim was to examine how individual differences in EF predicted children's learning of factual versus conceptual learning. Based on previous studies that measured both factual and conceptual learning in older children and adolescents (Bascandziev et al., 2018; Rhodes et al., 2014; Rhodes et al., 2016; Tardiff et al., 2020), we predicted that individual differences in EF would predict conceptual learning, but not factual learning. The third aim of the current study was to investigate whether the amount of conflict between the information to be learned and the learner's prior knowledge would interact with individual differences in EF to predict learning. To study this aim, we created three conceptual learning tasks that theoretically differed in the amount of conflict with the learner's prior knowledge (no prior knowledge, prior knowledge but not conflicting, and conflicting prior knowledge). Based on previous research (Bascandziev et al., 2018; Tardiff et al., 2020) and theoretical connections between conceptual change and EF, we hypothesized that EF would be a stronger predictor of children's conceptual learning when there was conflict with the learner's prior knowledge. If individual differences in EF have different relations to learning depending on the type of information being learned and the extent to which it conflicts with prior knowledge, this would suggest there may be certain learning contexts in which is it most useful to target children's EF to improve academic outcomes.

Method

Participants

Families were recruited from the Institute of Child Development Participant Pool (IPP). Institutional Review Board approval was obtained from the University of Minnesota. Participants included 61 typically developing 4-year-olds (29 girls, M_{age} = 53.72 months, *SD* = 2.71, Range = 49 to 59 months). The sample was mostly Caucasian (77%), but some participants identified as Asian (1.6%), African-American (1.6%), White-Hispanic (11.5%), and bi-racial (8.2%). The majority of the sample was also upper-middle class (median family income = $125,000-$149,000) with 3.3% earning $50,000 or less, 24.6% earning $50,000-$100,000, 37.7% earning $100,000-$150,000, and 34.4% earning more than $150,000 annually. The sample was also well educated with 49.2% of primary caregivers having a graduate or professional degree, 3.3% with some graduate school, 34.4% with a Bachelor's degree, 9.8% with some college, and 3.3% with a high school diploma. Five additional children participated but were excluded due to prematurity (*n* =1), speech and motor delays (*n* =1), not enough proficiency with English (*n* =1), or non-compliance/refusal to complete all of the learning tasks (*n* =2).

Procedure

Children participated in the lab individually for one 75-minute videotaped session. Parents completed a demographics questionnaire. Children completed an animal naming task, a series of learning tasks about animals (counterbalanced), three EF tasks, and a verbal IQ measure. Tasks were administered by a trained female graduate student. Parents received a $10 gift card and children received a T-shirt and a small toy for participating.

Measures

Learning

Animal Naming Task (5 min)

To measure children's familiarity with animals, they were shown 17 pictures of animals that varied in how familiar they are to children and adults. For example, a low-difficulty animal would be a pig, a medium-difficulty animal would be an ostrich, and a high-difficulty animal would be an axolotl. The experimenter presented each picture and asked children to name each animal. They were told they might not have seen some of the animals before, so they could say, "I don't know." Children received 1 point for each correctly labeled animal (0-17). Videos were double-coded and reliability was acceptable (*ICC* =.87).

No prior knowledge, factual learning task (10 min)

Children were told a story with pictures about pangolins (novel animal) including facts about where they live, what they like to eat, what they do when they are scared, whether they are nocturnal, etc. A pangolin was included in the animal naming task to see if children were familiar with and could name the pangolin. After the story, children were asked two memory check questions to ensure they were paying attention (e.g., "What was the Pangolin's name in the story?"), four comprehension questions (e.g., "What do pangolins do when they are scared?"), and two application questions (e.g., "Point to all the things that pangolins would like to eat"). Children received a memory check score out of 2, a comprehension score out of 4, and an application score out of 2. Inter-rater reliability was excellent for each of the three scores (*ICC*s = .90-.98).

No prior knowledge, conceptual learning task (10 min)

Children were taught about a new animal group (fantastical animals created in Piekny & Maehler, 2013) and features that make an animal part of this new animal category (e.g., red fur, wings, etc.). Children were first given a practice block of trials in which they were taught how to play the game. Here they were shown a picture of an animal and told it was called a "mufi" because it had long legs. Children were then shown three more pictures of animals one at a time and asked if they were a mufi or not and why. The animals varied in whether they had long legs but also had other characteristics of the first picture shown (e.g., fur). Then children were given three blocks of three test trials, with each block consisting of a new fantastical animal with a defining characteristic. Children received 1 point for every correct answer on test trials (0-9). Their reasons also were coded and children received 1 point if they gave a correct reason for why an animal was or was not the target novel animal (0-9). Therefore, children received an accuracy score and a reasoning score. Inter-rater reliability was excellent (*ICC*s =.94-.97).

No conflicting prior knowledge, conceptual learning task (10 min)

Children were taught about the category of mammals and the features that can be used to categorize an animal as a mammal (e.g., warm blooded, has hair/fur, bone in their backs, etc.). They were then shown two different animals (one mammal and one non-mammal) and were told characteristics about the animal that are needed to determine if it is a mammal. Then children were asked if the animal was a mammal or not and why. During these two practice trials, they were given feedback. Next, children were introduced to a puppet described as an animal expert named Mr. Hippo. They were then told they would be shown pictures of animals and their job was to decide if each one was a mammal or not. If they were not sure, they were told they could ask Mr. Hippo any questions to help them figure it out. The experimenter then demonstrated asking Mr. Hippo a question to determine if an elephant was a mammal or not (e.g., "Mr. Hippo, do elephants have warm blood?"), and invited children to ask Mr. Hippo a question. After the experimenter and/or child asked three questions (warm-blooded, bone in back, and has hair/fur), children were asked if they thought an elephant was a mammal or not and why (with corrective feedback). During the test trials, children were shown 7 pictures of animals individually and asked whether they thought each was a mammal or not and why. Any questions they asked Mr. Hippo were recorded. Lastly, children were asked how they know if an animal is a mammal. Children's responses were coded for accuracy (0-7). Children's reasons for why an animal was a mammal or not also were coded and they received 1 point if they correctly reasoned that an animal did or did not have all three of the characteristics they were taught (0-7). The number of questions children asked Mr. Hippo was also coded. Inter-rater reliability was excellent for each score (*ICC*s = .94-.99).

Conflicting prior knowledge, conceptual learning task (10 min)

Children were taught about instances of familiar animals that were in conflict with their prior knowledge. We adapted stimuli from Keil's (1992) study of children's concepts of inheritance and insides/essentialism. Children were told stories about animals that look like one kind of animal on the outside (e.g., racoon) but after being studied by scientists, it was discovered that they had the insides (blood, bones, brain, etc.) of another animal (e.g., skunk) and their moms and dads were a different type of animal (e.g., skunk). Children were then asked whether the animal was really a [skunk] or a [raccoon] and why. They were given one pre-test trial to assess whether they already understood the concepts of essentialism and inheritance. Children were then shown another example and taught that insides are important regardless of what the animal looks like on the outside and that the parents' identities are important for determining what kind of animal it is (learning trial). Two test trials followed. For each, they were asked what they thought the animal really was and why they thought that. Children's responses were coded and they received a pre-test score (0-1), a learning trial score (0-1), and a test trial score (0-2). Additionally, children's reasons on test trials were coded and they received a score of 1 if they mentioned the animals' insides or parents/babies or a score of 0 if they gave no response or another reason (0-2). Inter-rater reliability was excellent for each of the scores (*ICC*s =.98-1.00).

Executive Function

Minnesota Executive Function Scale (MEFS; Carlson & Zelazo, 2014; 5 min)

This EF task is administered on an iPad and has 7 levels of difficulty. The task is reliable and valid and can be used with children as young as 2 years old. The MEFS has been normed on a U.S. sample of 32,800 typically developing children ages 2-18 years. Children were asked to sort virtual cards based on different rules or dimensions (e.g., sort by color, sort by shape) with some levels requiring children to switch rapidly between rules. Children's starting level was determined by their age and then they either went up or down in levels depending on their performance until a basal and ceiling were reached. The MEFS software algorithm calculates a total raw score (0-100) based on children's accuracy and response times.

Forward/Backward Word Span (FWS/ BWS; Carlson, Moses, & Breton, 2002; 5 min)

This task measures children's working memory. Children were asked to repeat a list of words read out loud by the experimenter either exactly as stated (Forward Word Span) or in reverse order (Backward Word Span). Before test trials, children were given one practice trial for each task and given feedback if they answered incorrectly. Children were given up to 4 attempts to get the practice trials correct (and then proceeded regardless). On test trials, lists began with two words and then levels increased by one word if the child responded correctly. Children had three chances to correctly repeat lists at each level. For both the FWS and BWS tasks, children received scores corresponding to the highest number of words correctly recalled with possible scores ranging from 1-7.

Statue Task (Korkman, Kirk, & Kemp, 1998; 2 min)

This subtest of the NEPSY (A Developmental NEuroPSYchological Assessment) standardized battery assesses children's ability to stay still and to inhibit impulses and motor responses. It was included as a measure of behavioral inhibition. Children were asked to stand like a statue holding a pretend flag in one hand at a 90-degree angle with their eyes closed for 75-sec. Meanwhile, the experimenter produced sounds and distractions (e.g., coughing or dropping a pencil on the table) at specified times. Body movements, vocalizations, and eye openings were coded as errors during 5-sec intervals. Involuntary coughing, silent smiling, small movements of the fingers, and subtle movements due to balance issues were not coded as errors. For each interval, children received a 2 (no errors), 1 (1 error), or 0 (2 or more errors). If children said they were done pretending to be a statue and wanted to quit the task, they received 0 for the remaining intervals. Scores could range from 0-30.

Verbal IQ

Stanford-Binet Intelligence Scales for Early Childhood (5th ed.; Roid, 2005; 10 min). We used the verbal knowledge subtest of the Stanford-Binet to control for children's verbal IQ in analyses. The starting level for 4-year-olds consisted of showing them different toys and figurines (e.g., ball, cat) and asking them to label them. Then children were shown pictures in which people are doing different actions (e.g., drinking, running) and asked to verbally state what the person is doing. Finally, they were asked to define words and definitions were scored on a scale of 0-2 using test manual procedures, until they reached a ceiling of four consecutive zeros. Standard scores with a mean of 10 and standard deviation of 3 were used in analyses.

Results

Missing Data

We had missing data from 3.28% of the Stanford Binet Verbal Standard Scores and from 9.84 % of participants on the Statue Task. The latter was due mainly to children declining to try the task or not wanting to close their eyes during the task, rendering their data invalid. Given the Stanford Binet Verbal subtest was the last task in the study procedure, missing data was due to children declining to participate, or not having sufficient time in the session to complete the task. We used multiple imputation to impute missing data values using the mice package in R (van Buuren & Groothuis-Oudshoorn, 2011), which resulted in 20 imputations that were used in the main analyses.

Preliminary Results

As shown in Table 1, there was sufficient range on each measure of interest, with no floor or ceiling effects. Correlations for the raw data are shown in Table 2. Due to significant correlations among BWS, FWS, and the MEFS (*r*s = .31-.42), a composite was created, referred to as "cool" EF because all three measures are relatively unemotional (Carlson, 2005). Standardized BWS and FWS scores were averaged and this word span composite score was then averaged with the MEFS Total standardized scores. Children's Statue scores were not included in this cool EF composite because it was not significantly correlated with the other three EF measures. Additionally, an overall learning composite score was created by standardizing and then averaging children's scores on all four learning tasks. An overall conceptual learning composite score was created by averaging children's scores on the three conceptual scores (No Prior Knowledge, No Conflicting Prior Knowledge, and Conflicting Prior Knowledge).

Given the main goal of the current study was to measure children's learning of new information, we controlled for children's prior knowledge, with the exception of the task based on fantastical animals that do not exist (conceptual, no prior knowledge task). A prior knowledge score was created by averaging scores on the three learning tasks that included prior knowledge and this score was used as a covariate in all analyses. Without controlling for children's prior knowledge, it would be challenging to interpret whether children's EF skills were predictive of their learning or whether children with higher EF already had sufficient prior knowledge. Interestingly, prior knowledge scores were significantly correlated with cool EF composite scores, *r* (61) = .31, *p* =.02, but not with Statue scores (*p* =.89). The correlation between prior knowledge and cool EF remained significant even after controlling for age in months, *r* (61) =.28, *p* =.03. Preliminary analyses also indicated there were no significant effects of gender or learning task order on children's learning, so these variables were not included in the following models.

Main Analyses

Overall Learning

To address our first aim of whether individual differences in EF relate to children's overall learning of new information, we created an overall learning composite score. This score was created by standardizing and averaging children's total scores (accuracy plus reasoning scores) on each of the four learning tasks. We used hierarchical linear regressions predicting overall learning scores. In block 1, we included control variables of age in months and verbal IQ. In block 2, we included children's prior knowledge scores. In block 3, we added our EF variables (cool EF composite and Statue). Given these analyses used the multiply imputed data, we report the pooled parameter estimates for all results. As shown in Table 3, cool EF was a significant predictor of children's overall learning, controlling for age in months, verbal IQ, prior knowledge, and Statue task performance. However, Statue scores were only a marginally significant predictor of overall learning controlling for age in months, verbal IQ, prior knowledge, and cool EF.

Type of Knowledge

We next examined whether individual differences in EF predicted learning conceptual versus factual information. To address this second aim, we created an overall conceptual learning composite by averaging the standardized scores on each of the three conceptual tasks. We then used a series of hierarchical linear regressions predicting children's overall conceptual learning and factual learning separately. For each model, we included age in months and verbal IQ in block 1, prior knowledge in block 2, and the two EF measures in block 3. In the model predicting children's overall conceptual learning, we found that cool EF was the only EF measure that was a significant predictor controlling for age in months, verbal IQ, and Statue scores (Table 4). We next examined the model predicting children's factual learning and found that performance on the Statue task was a significant predictor controlling for age in months, verbal IQ, and cool EF (Table 5).

Amount of Conflict with Prior Knowledge

To address our final aim of investigating whether individual differences in EF interact with the amount of conflict there is between the child's prior knowledge and information to be learned, we used Hierarchical Linear Modeling to account for children's within-person variation in their performance on the three conceptual learning tasks that varied in the amount of conflict with children's prior knowledge (no prior knowledge, no conflicting prior knowledge, or conflicting prior knowledge). We ran two analyses, one looking at the interaction between task conflict and cool EF scores, and the other looking at the interaction between task conflict and Statue scores. We found a significant interaction between cool EF and the amount of conflict in the learning task, controlling for children's age in months, verbal IQ, prior knowledge, and Statue scores (See Table 6). As illustrated in Figure 1, children with lower EF performed worse compared to children with higher EF on the conflicting prior knowledge conceptual task (*B* =.17, *p* =.00), whereas the difference in learning performance between children with lower and higher EF for the other two tasks was not as prominent. Post-hoc simple slopes analyses revealed that only the slope for the conflicting prior knowledge task was significantly different from zero (*B* =.17, *p* =.00). In contrast, we did not find a significant interaction between task conflict and Statue scores, controlling for children's age in months, verbal IQ, prior knowledge, and cool EF scores.

Discussion

The current study further expanded our knowledge of the relation between individual differences in EF and young children's learning by examining different aspects of knowledge that children are asked to learn. By focusing on the type of knowledge (factual vs. conceptual) and the amount of conflict present between the child's prior knowledge and the information to be learned, the current study replicated and extended previous findings on the relation between EF and learning.

As we predicted, individual differences in EF predicted children's learning across tasks (factual and conceptual), even after controlling for age in months, verbal IQ, and children's prior knowledge. This finding is significant because it demonstrates that individual differences in EF uniquely predict children's overall learning after accounting for other variables that have been found to be related to both EF and learning. This finding adds to previous literature on the relation between individual differences in EF and children's *construction* of knowledge instead of focusing on children's ability to express knowledge they have already acquired.

The second main finding was that different types of EF measures predicted children's learning of different types of information. Individual differences in cool EF predicted children's conceptual learning, whereas their performance on a motor inhibition task with relatively low demands on working memory and cognitive flexibility (Statue task) predicted their learning of factual information. This was contrary to our hypothesis based on previous research, which found that verbal ability or measures of IQ, but not EF, were predictive of learning factual information (Bascandziev et al., 2018; Tardiff et al., 2020). Those authors interpreted their findings by stating that learning conceptual information involves overcoming conflicts, which theoretically would require EF skills, whereas there is no conflict in learning facts about a certain domain. One reason our findings were different might be due to the types of EF tasks used in the current study compared to previous studies. The Statue task used in the current study is a measure of motor inhibition that might serve as a proxy for being able to sit still while learning and represent an indirect effect of EF on learning. If this is the case, it is possible that performance on the Statue task predicted learning of factual information because learning facts still requires that the child sit still and pay attention. However, being able to sit still might not be enough to learn conceptual information because, in that case, children must rely on their working memory and cognitive flexibility skills to overcome previous conflicting knowledge to learn about a new concept. Therefore, if children can stand still and perform well on the Statue task, but do not yet have the cool EF skills needed to surmount the conflict in conceptual learning tasks, they would be likely to do well on factual learning but have trouble learning new conceptual information. Previous studies that found EF only predicted learning of conceptual information used shifting, inhibition, and working memory tasks that did not require behavioral inhibition and were similar to the tasks included in our cool EF composite (Bascandziev et al., 2018; Tardiff et al., 2020). The differences among EF tasks (cognitive regulation, behavioral regulation, hot EF, cool EF, etc.) need to be further examined in future research to understand better the relations between EF and learning of factual versus conceptual information.

The third main finding of the current study is that the amount of conflict between the learner's prior knowledge and the information to be learned significantly interacted with individual differences in EF to predict children's conceptual learning. Using a continuum of conflict (no prior knowledge and no conflict, prior knowledge but no conflict, prior knowledge and conflict), we were able to sort out the conditions under which EF plays a role in learning in relation to the amount of prior knowledge the child has and whether the information to be learned conflicts with any prior knowledge. Our findings extend previous research by suggesting that the conflict between the learner's prior knowledge and the information to be learned might be crucial for EF to play a role in children's construction of new knowledge. More specifically, interaction effects showed that cool EF was especially predictive of children's performance on the conflicting prior knowledge conceptual task, such that children with higher EF tended to have higher learning scores. Therefore, the amount of conflict with one's prior knowledge does seem to be one reason why there is a relation between EF and learning conceptual information.

Limitations and future directions

In addition to its strengths, the current study had a number of limitations. One limitation is that the study only included 4-year-olds. This was intended to focus on the year prior to formal school entry so as to assess their "readiness" to learn, and to study a narrow age range to maximize variability in learning (constraining floor and ceiling effects), but we were unable to look at developmental change and how relations between EF and learning might change with age. For example, as children develop, different aspects of EF may be more important for their learning of different types of information. Prior research with older children (5- and 6-year-olds) found that EF predicted their learning of conceptual but not factual information (Bascandziev et al., 2018; Tardiff et al., 2020). However, in our study, we found that motor inhibition (Statue Task) predicted young children's learning of factual information. Given that younger children are still developing their motor inhibition skills, a child's ability to sit still while learning might be more relevant for learning during the preschool period. In contrast, as children get older and acquire more knowledge, there could be more interference from prior knowledge to be overcome in the process of revising and constructing new theories, wherein cool EF is most relevant. Future research should include longitudinal studies that can examine how the relation between EF and learning changes over time.

Another limitation was that there might have been some conflict inherent in conditions that we theoretically thought would have no conflict with children's prior knowledge. For example, during the mammals task (no conflicting prior knowledge task), some children expressed a conflict with their prior knowledge when they were resistant to having an animal also be called a mammal (e.g., a bear could not be a mammal because it is a bear). This task was designed to teach children that certain animals could also be part of a group called mammals (i.e., enrichment of prior knowledge), rather than to challenge their existing concepts, but it also might have challenged their emerging understanding of class inclusion (e.g., Markman & Callanan, 1983). Our study also was limited in that we developed learning tasks that could be completed and post-tested in one 75-minute visit to the lab. Future research should explore how the relations between EF and learning might differ when looking at retention of learning by having participants return to the lab at a later date and assessing their knowledge of the topics taught, or by conducting the study in a more ecologically valid classroom setting. Finally, although the current study illustrated the importance of EF for different types of learning, it was correlational. It provides a foundation for future experimental training and intervention studies to more directly connect the dots between EF and learning.

Conclusion

Overall, we found evidence that individual differences in cool EF predict 4-year-olds' learning of facts and concepts about animals above and beyond children's age in months, verbal IQ, and prior knowledge. We also found that children's cool EF skills are important for their learning of conceptual information whereas their gross motor self-regulation was related to their ability to learn factual knowledge. Lastly, our study suggested that the amount of conflict with the learners' prior knowledge interacts with children's cool EF skills to predict learning, such that higher EF is more beneficial during tasks where there is more conflicting old information to be reflected upon and reconciled with new information. These findings have important implications for designing effective and individualized instructional strategies and materials. The findings suggest that the type of information being taught should be considered when designing instructional materials and that academic interventions could tailor learning materials to students depending on their level of EF.

Chapter 3: Study 2

Rationale for Study 2

In Study 1, we found that individual differences in cool EF predicted children's overall learning of factual and conceptual information about animals above and beyond verbal IQ, age, and prior knowledge. We also found that the characteristics of the information children were asked to learn (i.e., type of knowledge and amount of conflict with prior knowledge) was important for how EF related to learning. For example, we found that cool EF predicted children's learning of conceptual information, whereas motor inhibition (measured using the Statue task) predicted their learning of factual information. We also found that there were greater differences in learning for children with lower and higher EF when the learning task had the most conflict with their prior knowledge. The findings from Study 1 suggested a child's EF level should be considered when designing instructional materials and approaches. Therefore, Study 2 extended the results from Study 1 by examining whether two different instructional approaches were more or less effective depending on the child's EF level.

Study 2 allowed the findings of Study 1 to be extended in several additional ways. In Study 2, preschoolers were taught about sinking and floating. Sinking and floating is a STEM area that has not previously been used to study how individual differences in EF influence learning in preschoolers. The majority of the previous research looking at the relation between preschoolers' individual differences in EF and learning have focused on vitalism (Bascandziev et al., 2018; Tardiff et al., 2020; Zaitchik et al., 2013). Additionally, Study 2 allowed us to create a more ecologically valid learning context. Whereas in Study 1, children were presented the to-be-learned information using pictures printed on paper or displayed on a laptop screen, the sinking and floating lessons in Study 2 were more hands-on and engaging and represented a more ecologically valid pedagogical situation. Study 2 also was conducted in a preschool setting instead of in the lab like Study 1. Additionally, having multiple sessions in Study 2 enabled us to collect measures of children's early literacy and mathematics achievement. These measures allowed us to examine how the relations between EF and learning and EF and academic achievement compared in preschoolers, which was not possible in Study 1 due to the time constraints of a single study session.

Role of EF in Learning from Interventions

There have been a handful of academic intervention studies in math and science with young children that have shown that individual differences in EF moderate the effectiveness of the intervention, with children with high EF benefitting the most (e.g., Bascandziev et al., 2018; Fuchs et al., 2005, 2013a, 2013b, 2014; Laski et al., 2015; Miller et al., 2016). These studies show that certain interventions have been successful in increasing children's academic knowledge and that individual differences in EF are important to consider when thinking about who will benefit most from an intervention. However, there has been less research conducted to determine which parts of those interventions were responsible for the gains in academic knowledge and how those interventions could be altered to be effective for children who did not benefit as much from the intervention, such as children with lower EF.

There are potentially several different aspects of an intervention that are important to consider. A critical one is how the material is presented to the learner. Some examples of ways in which materials can be presented to the learner are the physical presentation of the learning stimuli, the use of pedagogical tools such as diagrams or worked examples, or the instructional strategies or approaches that are used to teach the material. There have been some studies that have examined how the physical presentation of the learning stimuli interacts with EF to influence learning. For instance, one study found that how simple arithmetic problems were presented (e.g., horizontally or vertically aligned) was associated with different types of working memory (Caviola,Mammarella, Cornoldi, & Lucangeli, 2012). There has also been some research looking at the effectiveness of using pedagogical tools such as diagrams for children's learning depending on their EF skills. For example, Reuter (2017) found that preschool children with lower EF benefitted more than children with higher EF from using a chart to display their predictions and observations while completing a science activity where they had to predict which objects would roll down a ramp and which ones would slide down a ramp. However, less attention has been given to how different instructional approaches may have different effects on children's learning, depending on the children's level of EF skills.

Instructional Approaches: Direct Instruction vs. Discovery Learning

A longstanding debate in education focuses on the best way to teach children information. This debate is often discussed using the dichotomy between direct instruction and discovery learning. Direct instruction refers to instruction in which the adult explicitly provides the information that he or she wants the learner to learn (e.g., Kirschner, Sweller, & Clark, 2006). However, the definition of discovery learning or instruction is less agreed upon. Despite various ways of defining discovery instruction, the majority of the definitions include that “the target information must be discovered by the learner within the confines of the task and its material” (Alfieri, Brooks, Aldrich, & Tennebaum, 2011, p. 2). Another component of discovery learning that can vary is the amount of guidance that is given to the learner. Some researchers have made the distinction between “pure” discovery learning approaches where the learner is given no help or guidance whatsoever and guided discovery approaches (Mayer, 2004; Alfieri et al., 2011). Guided discovery approaches include some level of guidance, such as providing feedback or using scaffolding techniques (Alfieri et al., 2011; Honomichl & Chen, 2012). Alfieri and colleagues (2011) conducted two meta-analyses on 164 studies with children (12 years and younger), adolescents (13 to 17 years), and adults (18 years and older) and compared the effectiveness of direct instruction, “assisted discovery learning,” and “unassisted discovery learning.” The main results were that learning outcomes were better for the direct instruction group compared to the unassisted discovery learning group. However, when they compared all three instruction types, they found that the assisted discovery learning group had better learning outcomes than the unassisted discovery and direct instruction groups. They found these results held for all age groups but were moderated by age. The effect sizes were larger for adolescents than adults in the meta-analysis comparing direct instruction to unassisted discovery learning and the effect sizes were larger for adults than children in the meta-analysis comparing all three instruction types (Alfieri et al., 2011).

Despite the results of the meta-analysis, there is evidence showing that direct instruction leads to more effective learning than discovery learning in some studies (Kirschner et al., 2006; Klahr & Nigam, 2004; Mayer, 2004; Tuovinen & Sweller, 1999). For example, in a study by Klahr and Nigam (2004), they taught third and fourth graders about the control-of-variables strategy which is a strategy used "… to design unconfounded experiments from which valid causal inferences can be made" in either a direct instruction or discovery learning condition (p. 662). The students were taught the control-of-variables strategy using ramps that had variables that could be varied: steepness, surface type, length, and type of ball. Children were told to figure out how the different variables influenced how far the balls could roll after going down the ramp. Children in the discovery learning condition were only told the goal of the activity but were given no other guidance or support. Children in the direct instruction condition were taught about the control-of-variables technique explicitly from a teacher who led the instructional session. They tested children directly after the instructional session and asked them to transfer what they learned to evaluate science fair posters of students a week later. They found that children in the direct instruction group learned more than the discovery group right after instruction. They also found that more children in the direct condition than in the discovery condition ended up being classified as having mastered the concept of control-of-variables strategy based on their performance on the initial post-test learning assessment. However, when they analyzed the transfer task measured a week later, they found that students in both instructional conditions who had mastered the control-of-variables strategy outperformed the students in both conditions who did not master the concept during the instructional session. The authors interpret their findings by stating that the way one learns the target information does not seem to matter for transferring the learned material. However, they still highlight the fact that more children in the direct instruction condition mastered the content than children in the discovery condition (Klhar & Nigam, 2004). Proponents of the superiority of direct instruction often discuss how direct instruction allows the learner to be focused on the relevant content, lessens the burden on the student's working memory capacity, and is more efficient and faster than discovery learning (Kirschner et al., 2006; Klahr & Nigam, 2004; Mayer, 2004).

The discovery learning side of the debate relies on support from ideas from constructivism (Piaget, 1970) and educational theorists such as Bruner (1961), who suggested that discovery may improve learning by allowing the child to be an active participant in the learning process and discover information on their own. Individuals in support of discovery learning often suggest that discovery learning is better because it allows the student to be active in their learning, construct their own knowledge, and to engage in deeper learning with a higher chance of transferring what they learned to other contexts (see Alfieri et al., 2011). While there is less evidence showing that pure discovery learning conditions lead to better learning, there is evidence that guided discovery learning conditions produce better learning than direct instruction (see Alfieri et al., 2011 for meta-analysis). Theoretically, guided discovery approaches deal with some of the concerns about pure discovery learning such as directing the learners' attention to what parts of the learning space are important and lessening the cognitive load (although not as much as direct instruction) (Mayer, 2004; Toub, Rajan, Golinkoff,

& Hirsh-Pasek, 2016; Weisberg, Hirsh-Pasek, & Gonlinkoff, 2013; Weisberg, Hirsh-Pasek, Golinkoff, Kittredge, & Klahr, 2016).

In the developmental literature, there is a similar set of instructional approaches that are compared and contrasted: learning, guided play, and free play. Learning is considered direct instruction, guided play is considered guided or assisted discovery, and free play is considered pure discovery. The same pros and cons mentioned earlier are discussed using these different terms, but there is a concern that pure or unassisted discovery approaches are especially not effective for young children. For example, Klahr and Nigam (2004) argue that “children in discovery situations are more likely than those receiving direct instruction to encounter inconsistent or misleading feedback, to make encoding errors and causal misattributions, and to experience inadequate practice and elaboration” (p. 661). Guided play and discovery involve concepts from Vygotsky’s sociocultural theory, such as following the child’s lead and scaffolding the child’s learning based on their current skill level (Vygotsky, 1978). Guided play occurs when “adults initiate the learning process, constrain the learning goals, and are responsible for maintaining focus on these goals even as the child guides his or her own discovery” (Weisberg et al., 2013, p. 105). Therefore, a balance between child autonomy and adult guidance is necessary for guided play to be most effective (Weisberg et al., 2016). There is evidence that teaching children through guided play leads to better learning of content such as math concepts (Fisher, Hirsh-Pasek, Newcombe, & Golinkoff, 2013), vocabulary (Han, Moore, Vukelich, & Buell, 2010) and how novel toys work (Bonawitz, Shafto, Gweon, Goodman, Spelke, & Schulz, 2011; Yu, Landrum, Bonawitz, & Shafto, 2018). For example, Fisher and colleagues (2013) taught 4-6-year-olds about different shapes in either a direct instruction, guided play, or free play condition. In the guided play condition, the child and the experimenter pretended to be detectives to figure out the shapes' secrets (i.e., each shape's attributes). Children in the direct instruction condition watched the experimenter figure out the shapes' secrets by observing the experimenter interact with the shapes and narrate what she was discovering about each shape. In the free play condition, children were allowed to play with the shape materials used in the other conditions for an amount of time equal to how long the sessions in the other two conditions lasted. They found that children in the guided play condition showed a greater change in their knowledge of shapes than the other groups and that this condition effect was still observed when the children were tested again one week later (Fisher et al., 2013). Some of the developmental studies highlight the idea that while direct instruction seems to lead to better learning of certain information, it also makes children less likely to explore or learn information beyond what they were explicitly taught in certain contexts (Bonawitz et al., 2011; Yu et al., 2018). Bonawitz and her colleagues (2011) have described pedagogy as a "double-edged sword" when describing this tradeoff (p. 329). For example, 4-6-year-olds who were taught about a function on a novel toy in a pedagogical manner (i.e., direct instruction) showed less exploration while playing with the novel toy after the instruction occurred and discovered less of the other functions of the toy that they were not explicitly taught compared to the non-pedagogical and baseline conditions. The authors interpret these findings by discussing how the explicit instruction might have indicated to children that the toy only had the one function since only demonstrating one function may signal there are not any other functions to learn (Bonawitz et al., 2011). Therefore, this study highlights a potential benefit of discovery learning. Guided play has been proposed as a developmentally appropriate, effective guided discovery learning approach for preschool children (Weisberg et al., 2013, 2016; Toub et al., 2016). Therefore, the current study compared a direct instruction condition to a guided discovery approach using play and questions to guide preschoolers' learning about sinking and floating.

Sinking and Floating as a Target Learning Domain

For Study 2, we chose to focus on teaching children about sinking and floating. This science domain was selected to explore the relation between individual differences in EF and learning in a new science area that is not biological such as learning about animals in Study 1 and vitalism in work by other researchers (Bascandziev et al., 2018; Tardiff et al., 2020; Zaitchik et al., 2013). We also chose a different science topic instead of one from another academic domain (e.g., math, literacy) since fewer studies have examined the relation between EF and science achievement. Researchers have theorized that EF skills such as inhibition and cognitive flexibility are important for certain early science concepts, such as revising hypotheses in the face of contradicting evidence (Gropen, Clark-Chiarelli, Hoisington, & Ehrlich, 2011). One study with preschoolers found that children's EF measured at the beginning of the school year predicted improvements in their science achievement at the end of the preschool year. Moreover, the relation between EF and science achievement was stronger than the associations between EF and other academic subjects such as math and literacy (Nayfeld, Fuccillo, & Greenfield, 2013). There have also been studies examining how EF may predict and explain individual differences in young children's scientific skills. For example, Van der Graaf and colleagues (2018) found that kindergarteners with higher EF were better at experimentation and evaluating scientific evidence. These findings point to the importance of understanding how to most effectively teach preschoolers about science while considering the role their EF skills may play.

Children's Understanding of Sinking and Floating

Children as young as 3 years of age can predict whether an object will sink or float, but are not as accurate as older children and adults (Kohn, 1993). For example, when asked to predict whether blocks that varied systematically based on weight, volume, and density sink or float, 3-year-olds were correct 53% of the time (at chance) whereas 4-5-year-olds were correct 72% of the time and adults were correct 86% of the time (Kohn, 1993). Studies that have varied and paired stimuli that vary on weight, volume, and density can examine the systematic errors that participants make and what types of biases or misconceptions may interfere when they are deciding whether an object will sink or float. Studies with young children have found that they tend to make errors based on volume and weight (Kohn, 1993; Plovt & Cyr, 2017). For example, Plovt and Cyr (2017) found that the majority of preschoolers tended to demonstrate a volume bias, and their reaction times were slower when they needed to overcome this bias to decide which of two objects would sink more. It has also been found that the constraints of the objects themselves and whether differences in certain aspects such as density, mass, or volume are more apparent influence how individuals perform on sink and float prediction tasks (Kloos, Fisher, & van Order, 2010).

Studies on young children's understanding of the concepts of weight, volume, and density have shown that as children develop, they can understand these concepts more and eventually differentiate between weight and density. For example, Smith and colleagues (1985) found that most 3-year-olds focused only on weight, 5-7-year-olds were able to focus on weight and volume, and 8-9-year-olds were able to distinguish between weight and density. While density is a complicated concept, even for adults, there is evidence that young children understand basic concepts related to density and can predict whether objects will sink or float. A more in-depth examination of children's explanations for why objects sink or float has shown that most preschool and early school-aged children tend to focus on one aspect of the object such as weight, size, volume, or material kind in their reasoning (Butts, Hoffman, & Anderson, 1993; Hsin & Wu, 2011; Smith, Carey, & Wiser, 1985; Tenenbaum, Rappolt-Schlichtmann, & Zanger, 2004). This type of thinking during the preschool period aligns with Piaget's classic studies on weight, density, and buoyancy and the four stages of conceptual development he created from those experiments. Preschool children fall in Stage 1 of his theory, where their explanations of why objects sink or float tend to involve moral reasons or focus on a single aspect of the objects. However, as children get older, their explanations for why objects sink or float get more complex. They can incorporate more than one relevant aspect in their explanations and finally can correctly include all the information pertinent to make judgments based on density (e.g., Meindertsma, van Dijk, Steenbeek, & van Geert, 2014; Tenenbaum et al., 2004).

Sinking and Floating Interventions

Several studies have evaluated the effectiveness of interventions to improve preschoolers' and early school-aged children's understanding of sinking and floating. These interventions have occurred in various contexts such as museums (Tenenbaum et al., 2004), classrooms (Kallery, 2015; Leuchter, Saalbach, & Hardy, 2014; Hardy, Jonen, Moller, & Stern, 2006), in small groups outside of the classroom (Hong & Diamond, 2012; Butts et al., 1993; Hsin & Wu, 2011), or one-on-one with an experimenter (Baker, Haubmann, Kloos, & Fisher, 2011; Kloos & van Orden, 2005; Kloos & Somerville, 2001). These studies also have demonstrated it is possible to change children's understanding of sinking and floating (Butts et al., 1993; Hardy et al., 2006; Hsin & Wu, 2011; Hong & Diamond, 2012; Kloos & Somerville, 2001; Kloos & van Order, 2005; Leuchter et al., 2014).

Whereas most studies have compared different instructional techniques or approaches to teaching children about sinking and floating, some studies have compared an intervention to a control group who does not receive the intervention or provides the same demonstrations to all children about sinking and floating (e.g., Tenenbaum et al., 2004; Kloos, & Fisher, 2011; Kloos & van Orden, 2005; Kloos & Somerville, 2001). For example, a study by Tenenbaum and colleagues (2004) compared kindergarteners who visited a museum exhibit about buoyancy to a control group of children who visited exhibits not related to buoyancy. Prior to visiting the museum, they received a lesson related to the topic of their exhibit in their classroom led by their classroom teacher. They found that children in the experimental group did not improve significantly more in their accuracy of predicting whether objects will sink or float compared to the control group, but they did find the experimental group improved in the complexity of their explanations for why objects sink or float after the intervention (Tenenbaum et al., 2004). Other studies have created demonstrations that allow children to witness the misconceptions they have about sinking and floating such as mass or volume biases (Kloos, & Fisher, 2011; Kloos & van Orden, 2005; Kloos & Somerville, 2001). For instance, Kloos and van

Orden (2005) focused on 4-6-year-olds' understanding about how fast an object will sink. After measuring children's initial beliefs, they showed them demonstrations using a submarine toy that could be filled with small weights. They explained to children that the amount of empty space in the submarine was important for how fast it sank. The more empty space a submarine had, the slower it would be to sink. These demonstrations were created to conflict with their misconceptions about mass and volume to see if they would be able to revise their previous belief. They found that a small number of children were able to replace their mistaken belief (5 out of 24 children).

Most of the interventions on sinking and floating vary the amount of instructional guidance that children receive. This instructional guidance can be in the form of different scaffolding techniques such as guiding children's interactions with objects, asking open-ended questions, and directing their attention to the relevant aspects of the task. A couple of studies compared a pure discovery condition to a direct instruction condition. They have found that just being able to play with or manipulate objects in water is not sufficient for learning about sinking and floating to occur (e.g., Butts et al., 1993; Hsin & Wu, 2011; Baker et al., 2011). For example, Butts and colleagues (1993) found that 5-6-year-olds who were able to manipulate the objects and received direct instruction on why objects sink or float improved more in their explanations for why objects sank or floated than children who were only allowed to manipulate the objects and received no explicit instruction.

Other studies used instructional approaches that were more guided (e.g., Hsin & Wu, 2011; Hong & Diamond, 2012). Hong and Diamond (2012) compared a responsive teaching instructional approach to a responsive teaching plus explicit instruction approach. One hundred and four 4-5-year-olds were randomly assigned to one of the instructional groups or a control group. The control group received book reading sessions where they were read books about biology. All children received four 15-minute intervention sessions. Children in the instructional groups were taught about science vocabulary and concepts, general scientific skills such as categorization and sorting, and skills needed to conduct experiments (e.g., prediction, testing, and evaluating) related to sinking and floating. In the responsive teaching condition, the instructor was responsive to the child and what they were doing with the materials but did not direct the child's play or activities and did not provide any explicit information about sinking and floating. In the responsive teaching with explicit instruction, the children received a 10-minute direct instruction lesson before getting to play and interact with the materials and water bin. The teacher used the same responsive teaching strategies that were used in the responsive teaching only condition. They found that the two instructional groups had significantly greater gains in their science vocabulary and science compared to the control group. There were no significant differences between groups for general scientific skills such as categorizing and sorting. The responsive teaching plus explicit instruction group improved significantly more than the control group in their skills needed to conduct experiments. However, the responsive teaching plus explicit instruction group did marginally better than the responsive teaching only group for these skills (Hong & Diamond, 2012). These studies show that it is possible to change children's understanding of sinking and floating, but some guidance is needed for this to be successful.

The target concepts related to sinking and floating tend to vary across studies such as water displacement (Cohen, 2017), the amount of empty space in an object (Baker et al., 2011; Kloos & Somerville, 2001; Kloos & van Orden, 2005), science skills related to understanding sinking and floating such as vocabulary, categorizing and sorting, prediction, testing, and explaining (Hong & Diamond, 2012), and material kind as a proxy for density (Hsin & Wu, 2011; Kallery, 2015; Leuchter et al., 2014; Smith et al., 1985). Smith and colleagues (1985) proposed material kind (i.e., the material that objects are made out of) as a possible proxy for density that could be used to teach young children about sinking and floating. Since the density of the material that the object is made out of determines if it will sink or float, material kind is a relevant feature to focus on and may allow children to begin to understand that weight and density are distinct but related concepts. Concentrating on material kind will also enable children to see that objects made out of the same material will behave the same in water (i.e., sink or float) even if they vary on other dimensions such as size, weight, shape, color, etc. A few studies with young children have taught them that material kind is essential for deciding whether an object will sink or float (Kallery, 2015; Leuchter et al., 2014). One study had six classroom teachers instruct 4-5-year-olds about sinking and floating using material kind (Kallery, 2015). The intervention included lessons to teach children how material kind could influence whether both solid and hollow objects will sink or float. In the hollow objects lessons, they were informed that an object made out of a material that usually sinks could float if it has enough air inside of it. They used concept cartoons to test children's understanding of sinking and floating after the intervention. The concept cartoons presented children with two characters who were disagreeing about a common misconception, such as the mass bias (i.e., heavy objects sink and light objects float). After the assessor read the cartoon to children, they were asked what they thought and why. They found that children performed well on each of the three cartoons and discussed how material kind determined whether a solid object sinks or floats regardless of the object's weight (90% correct) or size (88.5% correct) and that they understood that the amount of air in a hollow object made out of a sinking material could float (78% correct; Kallery, 2015). Another study randomly assigned classrooms with 4-9-year-olds to either receive a sinking and floating curriculum or to not receive it (Leuchter et al., 2014). The curriculum lasted for four weeks and included lessons that allowed students to 1) recognize their misconceptions, 2) learn that material kind should be considered and challenged their misconceptions, and 3) incorporating material kind when talking about whether an object sinks or floats. They found that children who received the sinking and floating curriculum significantly improved in their ability to accurately predict whether an object sinks or floats and in their reasoning concerning material kind compared to the control group immediately after the intervention and when they were tested again three months later. The instructional group also had higher gains on a task used to measure their understanding of how material kind relates to hollow bodies compared to the control group when measured immediately after the intervention and three months later (Leuchter et al., 2014). The current study used material kind as the focus of the sinking and floating lessons based on these prior research studies and the suggestion that focusing on material kind would allow children to begin to overcome their misconceptions about weight, volume, and size.

Instructional Approaches and EF

One of the main arguments made against discovery learning approaches is that discovery learning environments and techniques tax the learners' working memory capacity, which leads to poorer learning (e.g., Kirshner, et al., 2006; Tuovinen & Sweller, 1999). This idea is based on the Cognitive Load Theory, which was proposed by Sweller and colleagues (Sweller, 1988; Sweller, van Merrienboer, & Paas, 1998) and has been used to explain why instructional materials need to be designed based on the limitations of the learners' cognitive architecture. According to this theory, humans have limited working memory capacities. Therefore, instructional materials and approaches should be designed to lessen the working memory demands for the learner to enhance learning. Cognitive load is defined as "the load that performing a particular task imposes on the cognitive system" (Sweller et al., 1998). Sweller makes the distinction between two aspects that make up the cognitive load for a task. Intrinsic cognitive load refers to "the inherent aspects of the mental task that must be understood for the learner to be able to carry out the task" (Tuovinen & Sweller, 1999, p. 335). Extraneous cognitive load refers to "…extraneous matters associated with the way the instructional material is taught that may add to the inherent nucleolus of the intrinsic cognitive load" (Tuovinen & Sweller, 1999, p. 335). Proponents of the Cognitive Load Theory argue that discovery learning involves the learner searching the problem space for relevant information, which takes away from the working memory resources that they need to be able to learn the content and store it in long-term memory (Kirschner et al., 2006). Therefore, instructional approaches that reduce the working memory demands for the learner should lead to better learning. Tuovinen and Sweller (1999) tested whether worked examples led to better learning than discovery learning. Their rationale was that worked examples would reduce the need for the learner to use their working memory resources to search the problem-space. They found that college students who were taught how to use a database program learned more from worked examples than from exploring the database program if they had no prior knowledge about databases. However, if the students had previous knowledge about databases, they performed similarly in the worked examples and exploration condition. The authors interpreted these findings by saying that students with prior knowledge were able to use their relevant knowledge and schemas about databases to guide their learning, which lessened the cognitive load of the task for them (Tuovinen & Sweller, 1999). While Cognitive Load Theory has been used to discuss how instructional materials should be created by keeping the cognitive abilities of the students in mind, it also points to the idea that individual differences in working memory and EF skills more generally may influence how effective a discovery learning approach may be. It also suggests that the EF demands of a direct and discovery learning approach are different and would provide a useful context through which to test whether individual differences in EF moderate the effectiveness of these different instructional approaches.

While there are very few studies looking at the relation between individual differences in EF and direct versus discovery learning approaches, Homer and Plass (2014) conducted a study with high school students and examined how individual differences in EF moderated the effectiveness of discovery and direct computer simulations to teach students about chemistry concepts. In the exploratory condition, students engaged in interactive lessons where they were guided through how to test different hypotheses. In the worked examples condition, students were presented with the sequence of steps needed to test the hypothesis and solve the problem. They conducted two studies: the first focused on Kinetic Molecular Theory and the second focused on the Ideal Gas Law. To measure EF, they used interference scores from the color-word Stroop task in both studies. To measure learning, they used assessments that measured both their comprehension and transfer of the material. In study 1, they found that students in the exploratory condition performed better on the transfer test of learning than students in the worked examples condition. However, they did not find a significant interaction between students' EF levels and instructional condition. In study 2, they chose a more complex chemistry topic to increase the cognitive demands. They only found a significant interaction between students' EF levels and instructional conditions for students' performance on the transfer learning assessment: Children with lower EF performed better on the transfer task in the worked examples condition than the exploration condition, whereas children with higher EF performed better on the transfer learning task in the exploration condition than those in the worked examples conditions. However, for the worked examples condition, there was not a statistically significant difference in performance between the low and high EF groups for their transfer scores (Homer & Plass, 2014). This study supports the idea that instructional approaches that have different cognitive demands can interact with students' EF skills to predict the effectiveness of those approaches for students' learning. We tested whether this idea was true for preschool children. Given EF is rapidly developing during the preschool years (Carlson et al., 2013) and research shows EF is related to learning, we examined whether individual differences in 4-5-year-olds' EF would interact with different instructional approaches for teaching sinking and floating. If so, it would suggest that EF should be considered when designing instructional materials and interventions to enhance young children's learning.

Current Study

In Study 2, we examined whether individual differences in EF predict 4-and 5-year-olds' learning of sinking and floating concepts and whether the effectiveness of different instructional approaches (direct instruction vs. discovery instruction) depends on the child's EF level. The current study addressed the following five aims:

1. *Assess whether individual differences in EF predict the expression of sinking and floating knowledge*

   Based on previous research showing individual differences are related to young children's expression of their knowledge in other science domains (Bascandziev et al., 2016; Tardiff et al., 2020; Zaitchik et al., 2013), we predict EF will be related to their expression of their prior knowledge about sinking and floating (pre-test).

2. *Compare the relations between individual differences in EF and the expression of knowledge in science, math, and literacy*

   Based on previous research, we predict children's EF will be a significant predictor of their ability to express their math, literacy, and science (i.e., sinking and floating) knowledge. We also predict that EF will be a stronger predictor of children's math achievement scores than their literacy achievement scores (Allan et al., 2014; Jacob & Parkinson, 2015). While there is less research on how the relation between EF and science compares to other academic domains, we predict that the strength of the relation between EF and science will be more similar to math than of literacy based on a prior study showing that EF most strongly associated with science than math and literacy in preschoolers (Nayfeld et al., 2013).

3. *Assess whether there are differences in learning in the different instructional conditions*

   Based on previous research (e.g., Hong & Diamond; Mayer, 2004), we predict children in the direct instruction group will learn more about sinking and floating than children in the discovery instruction group.

4. *Examine whether differences in learning in the different instructional conditions depend on the child's level of EF*

   We predict that children with higher EF will learn more from the discovery instruction condition than will children with lower EF (Homer & Plass, 2014). The discovery condition may carry a higher cognitive load since efficiently exploring the problem space requires one to keep in mind what has already been discovered in the problem space, to inhibit information that is not relevant to the learning task, and to be flexible enough to explore different aspects of the problem space instead of getting stuck on one area. We also predict that children with lower EF will learn more from the direct instruction condition than the discovery learning condition. Direct instruction will help children with lower EF know where to focus their attention and can scaffold their learning in ways that are more compatible with lower EF skills (e.g., reminding students of what they have previously been talking about, helping them form connections between different parts of the learning task, etc.).

5. *Compare EF as a predictor of expression versus construction of knowledge*

EF will be related to both the expression of prior knowledge (pre-test knowledge) and the construction of new knowledge (increases in learning post-intervention). However, based on the findings from Study 1and other studies looking at how individual EF components in EF predict expression and construction of knowledge, we will also examine the EF tasks separately as predictors. For Statue, we do not have a directional hypothesis. On the one hand, Statue may not be related to expression or construction of sinking and floating knowledge because we found that Statue was not associated with conceptual learning in Study 1 and the majority of the sinking and floating learning outcomes for Study 2 are conceptual. On the other hand, motor hyperactivity may influence attention during the instructional sessions and, therefore, influence their learning during the instructional sessions. For the cool EF tasks, we hypothesize it will be related to both expression and construction of knowledge given cool EF was a significant predictor of prior knowledge (expression of knowledge) and learning (construction of knowledge) in Study 1 and previous research findings showing that cool EF components predicted expression versus construction of knowledge (e.g., Bascandziev et al., 2018; Miller et al., 2016).

We also predict that individual differences in EF will be related to both academic achievement in math and literacy and children's learning about sinking and floating. However, we predict that EF will be a stronger predictor of children's academic achievement since this represents their ability to express knowledge that they have acquired over time and consolidated.

Method

Participants

Children were recruited from 15 childcare centers in and around the Twin Cities. The centers were part of two large childcare networks. The director at each center volunteered to participate in the study. Participants included 93 typically developing 4- and 5-year-olds (44 girls, *Mage* = 58.5 months, *SD* = 5.6, Range = 48 to 71 months). The sample was mostly Caucasian (50.1%), but some participants identified as Asian (3.2%), African-American (11.8%), Native American/Alaska Native (1.1%), White-Hispanic (2.2%), and bi-racial (11.8%). Parents of 10 children (10.8%) did not provide their child's ethnicity. About half of the sample was upper-middle-class (median family income = $175,000-$199,999) with 50.6% earning more than $150,000 annually. For the remaining families, 17.2% reported earning $50,000 or less, 6.5% reported earning $50,000-$100,000, and 12.9% reported earning $100,000-$150,000. Twelve families (12.9%) did not report their family income. The sample was also well educated, with 46.2% of primary caregivers having a graduate or professional degree, 4.3% with some graduate school, 23.7% with a Bachelor's degree, 11.8% with some college, and 3.2% with a high school diploma. Ten primary caregivers (10.8%) did not report their highest level of education.

Recruitment

A letter explaining the study was sent home by the participating centers with children of eligible families. After the letters were sent home, families were recruited during designated pick-up times at the participating locations, and informed consent was obtained. Children were eligible to participate in the study if they were the correct age

(48-71 months), were English speaking, did not have a premature birth, and had no known developmental delays or disorders or physical impairments that affect vision or hearing. After signing up for the study, parents were asked to complete a demographic questionnaire online via Qualtrics. For participating in the study, children received a sticker after each session and a small toy after finishing the last session.

An additional 24 children were recruited but did not participate in the study because they stopped attending the childcare center before data collection began (*n* = 19), or the child declined to participate in the pre-test session (*n* = 5). Thirty-eight additional children participated but were excluded due to prematurity (*n* = 8), speech delays (*n* = 1), motor delays (*n* = 1), ADHD (*n* = 2), previous brain injury (*n* = 1), not meeting the age eligibility requirement (*n* = 1), non-compliance/refusal to complete all of the instructional sessions (*n* = 2), inability to complete all the tasks in a timely manner due to attendance (*n* = 1), stopped attending the center before all sessions were completed (*n* = 6), inability to complete sinking and floating pre-test assessment (*n* = 2), not enough days between pre- and post-test sessions (*n* = 1), or the exclusion criteria could not be obtained from parents (*n* = 12).

Procedure

All children participated in a pre-test and post-test direct assessment session. Children in the instruction groups also participated in three one-on-one instructional sessions where they learned about sinking and floating. All sessions were conducted by a trained graduate student or undergraduate research assistant at the participating childcare centers during the school day. The pre-test session lasted about 30 minutes. Next, children were randomly assigned to one of three conditions: a direct instruction condition (*n* = 31), a discovery instruction condition (*n* = 32), or a no instruction control group (*n* = 30). Condition assignments were monitored by a graduate student researcher not involved in the instructional sessions to make sure that the number of children with high and low EF in each condition was balanced. Percentile scores from one of the EF measures used during the pre-test session, the Minnesota Executive Function Scale (MEFS), was used to classify children as having relatively high or low EF based on the U.S. norms. Children in the instruction groups were pulled individually to receive three 25-minute instructional sessions on sinking and floating. Children in the control group did not receive any instructional sessions. Finally, all children participated in a post-test session that lasted between 30 and 40 minutes depending on children's performance on the standardized tests. The goal was to have children complete all sessions within two weeks. However, due to the childcare centers' schedules, child absences, and children's attendance patterns (e.g., half days or only attending certain days of the week), this was not always possible. Therefore, the number of days between the pre- and post-test sessions was variable (*M* = 14.72 days, *SD* = 8.0, range = 4-45 days). See Figure 2 for a visual representation of the study design.

Pre-Test Session

During the pre-test session, children's EF and prior knowledge about sinking and floating were assessed. Children first completed the sinking and floating pre-test assessment and then the Minnesota Executive Function Scale (MEFS), Forward Word Span (FWS), Backward Word Span (BWS), and the Statue task.

Instructional Sessions

Children in the instruction groups received three 25-minute instructional sessions where they received lessons on sinking and floating. The sinking and floating lessons focused on material kind as a proxy for density, which has been used in previous studies with young children (e.g., Kallery, 2015; Leuchter et al., 2014). The first two lessons focused on demonstrating to children that size and weight alone do not determine whether an object will sink or float, but rather the material of the object is what is important (e.g., clay and metal sink and wood and plastic float). In the third lesson, they learned how to make an object that floats sink and an object that sinks float. The materials for both instructional conditions were the same for every instructional session (water tub, sinking and floating chart, objects, scale, etc.). However, how children were taught about material kind and sinking and floating differed depending on their instructional condition.

Children in the direct instruction condition were explicitly told that material kind matters throughout the lessons and demonstrations, and the experimenter explicitly labeled the materials and pointed out how the objects were the same or different. Children in the discovery instruction condition were not explicitly told that material kind matters but instead were asked open-ended questions to assist them in discovering this on their own. For example, when looking at wooden sticks that are all different sizes to explore how size influences whether something sinks or floats, the experimenter would ask the child in the discovery condition, “What’s the same about these?” and “What’s different about these?” In contrast, in the direct instruction condition, children would be explicitly told, “Look, these are all sticks, and they are all made out of wood, but they are different sizes. See, some are big, and some are small." See Appendix A for a summary of the concepts, materials, and activities for the three instructional sessions.

Post-Test Session

In the post-test session, children's knowledge of sinking and floating was assessed again using the sinking and floating assessment. Next, they completed the verbal and non-verbal subtests of the Stanford Binet IQ test and the Letter/Word (literacy) and Applied Problems (math) subtests of the Woodcock-Johnson Tests of Achievement.

Measures

Sinking and Floating Assessment

Tasks used in previous studies on young children's understanding of sinking and floating were used to measure children's sinking and floating knowledge. The assessment was administered at the beginning of the pre- and post-test sessions and had five parts. All assessments were double-coded and the intraclass correlation coefficients (*ICCs*) for each score are presented below along with descriptions of each part of the assessment.

Definition of Sink and Float. Children were asked to define "float" and then "sink." Children's responses were scored on a 0-2 scale based on the scoring in Hong and Diamond (2012). They received a 0 if their definition was incorrect or if no answer was given, a 1 if their answer was accurate and acceptable but not sophisticated enough, and a 2 if their answer was accurate and sophisticated, demonstrating a clear understanding of what float and sink mean. The scoring of the definitions was reliable (*ICC* =.94 for float and *ICC* = .95 for sink).

Sink and Float Predictions. Similar to assessments of sinking and floating used in several other studies, children were presented with 14 objects that differed in size, shape, and material kind and were asked whether they thought the object would sink or float. Different objects were used for this section for the pre- and post-test (see Appendix B for the set of objects used in the pre-test assessment and Appendix C for the post-test assessment objects). After indicating their prediction, children were asked why they thought the object would sink or float. Children's responses to both questions were recorded and coded. Children received one point for each correct prediction, and these were summed for a total prediction score (max of 14). Children's reasoning was also scored on a 0-5 scale based on what they focused their explanation on and how sophisticated their reasoning was (max of 70, see Appendix D for the scoring scheme). Reliability for the prediction score (*ICC* = .96) and explanation score (*ICC* = .98) were excellent.

Sinking and Floating Misconceptions. Concept cartoons from Kallery (2015) were used to assess children's understanding of common misconceptions of sinking and floating concepts. Children were presented with three concept cartoons, each measuring a different sinking and floating misconception: weight, size, and material kind. In each cartoon, there were two characters (one male and one female) who had different opinions. Children were told what each character said and then asked which character they thought was right and why. Children were then asked a simpler version of the question that was more direct to measure their understanding of the misconception (e.g., Luna thinks that if she makes her clay doll smaller, it will float. Do you think that will work? Why?). Pilot testing indicated that the concept cartoons in their original form might have been difficult to understand for some children, so the second question was asked to address this issue. See Appendix E for an example of one of the concept cartoons used along with its script.

Children's answers and reasoning to both questions for each cartoon were scored. Children received an accuracy score (max of 6), a misconception score (0-2 for each cartoon), a consistency score (max of 3), and a self-correct score (max of 3). For the accuracy score, children were given one point if they provided the correct answer and a zero if they provided an incorrect answer (two questions for each cartoon). For the misconception score, children received a 0 if they did not provide any evidence of whether or not they had the misconception presented in the cartoon, a 1 if they demonstrated evidence that they had the misconception, and a 2 if they demonstrated evidence that they did not believe the misconception. Since children were asked about the misconception for each cartoon twice, children's consistency in their reasoning and whether they self-corrected their answer were scored. For the consistency score, children received a 1 for consistent reasoning and a 0 for inconsistent reasoning. For the self-correct score, children received a 1 if they initially gave a wrong answer and then gave the correct answer when asked the second question. The reliability for each of these scores was acceptable (*ICC*s ranged from .61 to .98).

Children also were asked how boats float and how boats can carry heavy stuff without sinking. Children's responses were recorded and coded based on the accuracy of their responses. Children received a 0 if they did not provide an answer, a 1 if they provided an incorrect response, and a 2 if they provided a correct response. The reliability for each of these questions was excellent (*ICC* = .91 for how boats float and *ICC* = .90 for how boats can carry heavy stuff).

Changing Whether Something Sinks or Floats. In this adaptation of the assessment from Hong and Diamond (2012), children were presented with one object that sank and one object that floated and asked to provide ideas on how to change the outcome. Children were first presented with a plastic jar filled halfway with sand and told that the jar would sink if put in water. Then children were asked how they could make the jar float. Each idea that the child presented was recorded. Children were next presented with an empty water bottle and told that it would float when put in water. Children were then asked to provide ideas on how to make the bottle sink. Children's responses were scored for how many unique ideas they provided for each object, and then each idea was scored on a 0-4 scale based on whether the idea would work to make the object either sink or float and how specific their idea was. Children' ideas received a 0 if they did not provide any ideas, a 1 if they provided ideas that would not work to make the object sink or float (depending on the goal), just said to make it sink or float without any specific ways to accomplish this, or if they gave impossible ideas (e.g., using magic), a 2 if they provided an idea that would temporarily make the object sink or float (e.g., push the object down with your hand), a 3 if they provided ideas that would work to make it sink or float but were not specific (e.g., make it heavier), and a 4 if they provided ideas that would work and were specific (e.g., put rocks in the plastic bottle to make it heavier). The coding of each score reached acceptable levels of reliability (*ICC*s ranged from .86 to .95).

Transfer Question. Finally, we assessed whether children could transfer their ideas to new scenarios based on their current level of understanding of sinking and floating. Children were told a story about a caterpillar who wanted to cross a river and had four different objects she could choose from to use to cross the river (Leuchter et al., 2014). Children were presented with four dry clay objects (a sphere, a cylinder, a plate, and a boat) and were asked which one the caterpillar should pick and why. Children were asked if there were any other objects she could use until the child said there were no more objects she should use, or they chose all four objects. Children should have only selected the boat, and the other three objects (sphere, cylinder, and plate) were considered distractors. Children received a score between 0 and 2, depending on which objects they chose. Children received a 0 if they chose more than 2 of the objects, none of the objects, or only the distractor objects, a 1 if they chose the boat plus one of the distractors, and a 2 if they only chose the boat. The scoring reliability of this task was excellent (*ICC* = .97).

EF Measures

The same EF battery that was used in Study 1 was used in Study 2. The battery consisted of the MEFS, FWS/BWS, and the Statue Task.

IQ: Stanford-Binet Intelligence Scales for Early Childhood (5th edition; 10 min; Roid, 2005)

The Verbal Knowledge subtest measures children's verbal ability and consists of asking children to identify words by choosing pictures and defining words that get more challenging as they go. The Nonverbal Fluid Reasoning subtest measures children's ability to solve new problems such as completing sequences or identifying patterns. Standard scores with a mean of 10 and a standard deviation of 3 for each subtest were used in the analyses.

Academic Knowledge: Woodcock-Johnson III-NU (16 min; Blackwell, 2001)

The Letter/Word Identification subtest is a measure of emerging literacy. This subtest consists of identifying lower and uppercase letters and reading simple, commonly used words. For the Applied Problems subtest, children are read math problems out loud, which they are asked to solve. This subtest assesses quantitative reasoning, math achievement, and math knowledge. Children received one point for each correct response, and their correct answers on each subtest were summed to create a raw score. The raw scores were then standardized based on children's age. Standard scores were used in the analyses.

Parent Questionnaire

After families signed up for the study, they received an email invitation to take the Family Information Questionnaire (FIQ) online via Qualtrics. This was used to gather basic demographic information about family structure, the age, education, ethnicity of the parent and child, current occupation of the parent(s) living in the home, and family income.

Results

Preliminary Analyses

Descriptive statistics for each task are presented in Table 7. Although there was sufficient range on most measures of interest, there were floor effects for some of the sinking and floating measures. For example, the Caterpillar Transfer Task, Sink Definition, and Float Definition had a floor effect. For the Caterpillar Transfer Task, it appears that children did not understand the objective of the task and were distracted by the objects themselves and thinking about the caterpillar in the story. For the definitions of sink and float, there appeared to be little variability at each time point, and on average, children received very low scores on these measures. Therefore, these three scores were not considered further. Examination of the distribution of the scores for the remaining sinking and floating measures indicated that they were not normally distributed.

Therefore, each score was log transformed and the transformed scores were used in the main analyses. Correlations between all of the variables using the raw data are shown in Table 8. Given Caregiver Education Level and Family Income were highly correlated ($r(81) = .58, p <.001$), an SES Composite Score was created by standardizing each score and taking the average. This composite was used as a covariate in the main analyses.

We also conducted one-way Analyses of Variance (ANOVAs) to confirm that randomization to conditions was successful and that there were no systematic differences between the groups before the instructional sessions. These analyses revealed no statistically significant differences between conditions for any of the outcomes of interest, age, or gender at pre-test except for the Days Between Pre- and Post-Test ($F(2, 90) = 4.49, p = .014$) , Pre-Test scores for the Sand Jar Highest Score Received ($F(2,90) = 3.88, p = .02$) and Non-Verbal IQ scores ($F(2,88) = 3.80, p = .03$). Post-hoc analyses indicated that the Control group ($M = 11.37$ days, $SD = 1.3$) had significantly less days between pre- and post-test sessions than the two instructional groups ($M_{\text{Direct}} = 15.81$ days, $SD = 1.28$; $M_{\text{Discovery}} = 16.34$ days, $SD = 1.26$). However, there was no significant difference between the number of days between the pre- and post-test sessions for the Direct and Discovery conditions. For the Pre-Test Sand Jar Highest Score Received Scores, children in the Control group ($M_{Control} = 1.63$, $SD = .27$) had significantly lower scores than the Direct condition ($M_{Direct} = 2.65$, $SD = .26$) and marginally lower scores than the Discovery condition ($M_{Discovery} = 2.38$, $SD = .26$). For Non-Verbal IQ standard scores, the Control group ($M_{Control} = 8.93$, $SD = .65$) had significantly lower Non-Verbal IQ Standard Scores than the Direct ($M_{Direct} = 11.43$, $SD = .67$) and Discovery ($M_{Discovery} = 10.94$, $SD = .65$) conditions. However, there was no significant difference for Non-Verbal

IQ Standard Scores between the two instructional conditions. To account for these preexisting differences in our analyses, we controlled for the Number of Days Between Pre- and Post-Test Sessions for the analyses that included all three conditions. We also controlled for pre-test scores on all the sinking and floating measures when sinking and floating learning was a DV and controlled for Non-Verbal IQ scores in all analyses.

Data Reduction

Due to significant correlations among the BWS, MEFS, and Statue tasks (*rs* = .26-.44), we conducted a Principle Components Analysis (PCA) to confirm whether the three measures should be composited into one score. The PCA indicated that there was one factor with an eigenvalue greater than one. Standardized loadings were .67 for Statue, .77 for MEFS and .80 for BWS and together explained 56.12% of the variance. Therefore, an EF composite was created by averaging the standardized scores on each measure.

Given the sinking and floating assessment was created by combining tasks from previous research with similar aged children (Hong & Diamond, 2012; Kallery, 2015; Leuchter et al., 2014) and has never been used as a holistic assessment, we conducted a PCA to inform whether any of the scores obtained from the assessment should be composited. We included six sinking and floating scores in the PCA: Prediction Score, Explanation Score, Material Kind Frequency, Sand Jar Highest Score Received, Plastic Bottle Highest Score Received, and Total Concept Cartoon Score (accuracy plus reasoning). We conducted a PCA on the pre-test and post-test scores separately.

The analysis with the pre-test scores indicated that there were two factors that had an eigenvalue greater than 1. For the first unrotated factor, standardized loadings ranged from .20 (Prediction Score) to .80 (Explanation Score) and together explained 40.03% of the variance. For the second unrotated factor, standardized loadings ranged from - .39 (Concept Cartoon Total Score) to .85 (Prediction Score) and together explained 18.24% of the variance. The analysis with the post-test scores indicated that there was one factor with an eigenvalue greater than one. For this factor, standardized loadings ranged from .56 (Concept Cartoon Total Score) to .80 (Material Kind Frequency) and together explained 49.91% of the variance. We decided to take the first unrotated principle component factor because the loadings were all above .3 (ranged from .53 to .75) for the post-test scores and all but one score loaded onto the first unrotated principle component at pre-test (ranged from .22 to .74). Therefore, we created Pre-Test and Post-Test Sinking and Floating Overall Learning Composite scores by standardizing and averaging the scores on the six sinking and floating scores (Prediction Score, Explanation Score, Material Kind Frequency, Sand Jar Highest Score Received, Plastic Bottle Highest Score Received, and Total Concept Cartoon Score (accuracy plus reasoning)).

Missing Data

Variables of interest that had missing data ranged from 1.1% (FWS) to 13.98% (Pre-Test Total Concept Cartoon Score- accuracy plus reasoning). Given the Total Concept Cartoon Score was a sum of children’s accuracy and reasoning scores across three different cartoons, children were missing this score if they failed to answer any of the questions needed to create this summed score. According to Little’s MCAR test (Little, 1988), the missing data appeared to be missing completely at random ($\chi 2$ [731] = 693.56; $p = .836$). We used multiple imputation to impute missing data values using the Statistical Package for the Social Sciences Version 25, which resulted in 10 imputations that were used in the main analyses. We report the pooled parameter estimates provided by SPSS when available because the main analyses used the multiply imputed data. Given the R^2 values are not pooled automatically in SPSS, we averaged the R^2 values across the 10 imputations as recommended by van Ginkel (2019).

Main Analyses

Aim 1: EF Predicting Expression of Sinking and Floating Knowledge

First, we conducted hierarchical linear regression analyses for our Pre-Test Sinking and Floating Overall Learning Composite Score. We included Age in Months at Pre-Test, Verbal and Non-Verbal IQ, and SES Composite Scores as control variables in block 1. The EF Composite Score was included in block 2. These analyses indicated that EF Composite Scores was not a significant predictor of children's Pre-Test Sinking and Floating Overall Learning Composite Scores after accounting for all covariates ($B = .15$, $p = .16$; See Table 9).

Given we found a marginally significant main effect of SES but not of EF Composite Scores, we ran an exploratory analysis to see if there was a significant SES x EF Composite interaction for children's Pre-Test Sinking and Floating Overall Learning Composite Scores. Although the SES x EF Composite interaction was not significant ($B = -.14$, $p =.11$), there was a trend for the effect of EF on children's Pre-Test Sinking and Floating Overall Learning Composite Scores to decrease as children's SES increased. Children from lower-income families with lower EF Composite Scores expressed less prior knowledge about sinking and floating than children from lower-income families with higher EF Composite Scores.

Aim 2: EF Predicting Expression of Knowledge in Science, Math, and Literacy

Next, we ran additional hierarchical linear regression models predicting children's scores on the WJ Letter/Word (Literacy) and WJ Applied Problems (Math) subtests separately**.** We included Age in Months at Pre-Test, Verbal and Non-Verbal IQ, and SES Composite Scores as control variables in block 1 and the EF Composite Score was included in block 2. As shown in Table 10, EF was not a significant predictor of WJ Letter/Word scores ($B = 4.62$, $p = .07$). For the model predicting Applied Problems scores, we found that EF was a significant predictor ($B = 8.00$, $p =.005$).

We conducted exploratory analyses to see if there was a significant SES x EF Composite Interaction for children's Letter/Word and Applied Problems Scores. There was not a significant interaction for Letter/Word ($B = -2.81$*,* $p = .26$), but there was a significant SES x EF Composite interaction for Applied Problems ($B = -3.88$, $p = .04$). The effect of EF on Applied Problems performance decreased as children's SES increased. Children with lower EF Composite Scores from lower-income families had lower Applied Problems scores than children with higher EF Composite Scores from lower-income families.

Given EF was only a significant predictor of children's math achievement scores after controlling for our covariates, comparing the strength of EF as a predictor of children's expression of knowledge in different academic domains (science, math, and literacy) was not necessary.

Aim 3: Differences in Learning Between Instructional Groups

Next, we looked to see if children in the instructional groups learned significantly more than children in the Control group. In this set of analyses, we conducted a series of hierarchical linear regression analyses with Age in Months at Pre-Test, Days Between

Pre- and Post-Test Sessions, Verbal and Non-Verbal IQ, and SES Composite Scores as covariates in block 1, children's Pre-Test Sinking and Floating Overall Learning Composite Scores in block 2, and Condition in block 3. Given the focus of the intervention was using material kind as a proxy for density, we first examined whether there were significant differences in the frequency of how often children mentioned material kind in their reasoning during the sink/float prediction task. There was a significant main effect of Condition for Material Kind Frequency with children in the Direct condition showing a greater increase in the use of material kind when reasoning about whether an object would sink or float (Figure 3). This analysis served as a manipulation check to see if children were picking up on the target information that was delivered during the intervention.

Next, we examined children's Post-Test Sinking and Floating Overall Learning Composite Scores. As can be seen in Table 12, Condition was a significant predictor of children's Post-Test Sinking and Floating Overall Learning Scores. Children in the Direct condition learned significantly more than children in the other conditions (Figure 4). After confirming that children in the Control group did not make significant gains in their sinking and floating knowledge from pre- to post-test, as expected, the remaining analyses only included children in the two instructional groups.

Aim 4: Interaction between EF and Instructional Condition

First, we looked to see if children's EF Composite Scores predicted their Post-Test Sinking and Floating Overall Learning Composite Scores controlling for their Age in Months at Pre-Test, Verbal and Non-Verbal IQ, SES Composite Scores, Pre-Test Sinking and Floating Knowledge, and Condition. Hierarchical linear regression analyses indicated that the EF Composite was not a significant predictor of children's Post-Test Sinking and Floating Overall Learning Composite Scores (*B* =.04, *p* =.74; See Table 13).

Next, we examined if there was an interaction between EF Composite Scores and Condition to predict children's Post-Test Sinking and Floating Overall Learning Composite Scores, controlling for Age in Months at Pre-Test, Verbal and Non-Verbal IQ Scores, SES Composite, and Pre-Test Sinking and Floating Overall Learning Composite Scores. There was not a significant interaction between Condition and EF Composite Scores to predict children's Post-Test Sinking and Floating Overall Learning Composite Scores (*p* = .61).

Given we found no significant main effect of EF or a significant EF x Condition interaction, we conducted exploratory analyses to better understand which children learned the most from the intervention and to see if EF was a significant predictor of their learning. First, we examined the effect of children's prior knowledge about sinking and floating on their learning from the instructional sessions. We created a change score for children's sinking and floating learning by subtracting their Post-Test Sinking and Floating Overall Learning Composite Score from their Pre-Test Sinking and Floating Overall Learning Composite Score. We then ran a series of hierarchical linear regression analyses to see if children's prior knowledge about sinking and floating was a significant predictor of the amount of change they made in their sinking and floating knowledge after the instructional sessions, controlling for Age in Months at Pre-Test, Non-Verbal and Verbal IQ, SES Composite, and Condition. As can be seen in Table 14, children's prior knowledge about sinking and floating was a significant predictor of the amount of change in their sinking and floating learning after the intervention (*B* = -.72, *p* =<.001).

This indicated that children with lower levels of prior knowledge about sinking and floating made greater gains as a result of the intervention compared to children with higher levels of prior knowledge.

Next, we examined whether children's prior knowledge interacted with their EF to predict their knowledge of sinking and floating after the intervention. The EF Composite x Prior Knowledge interaction was not significant (*B* = -.17, *p* = .16), however, we noted that among children with lower Prior Knowledge, those with higher EF learned more than children with lower EF (Figure 5). Given Cognitive Load Theory (e.g. Sweller, 1988) tends to focus specifically on working memory, we examined whether the interaction between each of the EF tasks and Prior Knowledge were significant predictors of children's sinking and floating learning scores after the intervention. Although not statistically significant, there was a trend for an interaction between BWS and children's Prior Knowledge controlling for Age in Months at Pre-Test, Non-Verbal and Verbal IQ, SES Composite, and Condition (*B* = -.23, *p* =.10). However, the MEFS x Prior Knowledge and Statue x Prior Knowledge interactions were not significant. Finally, we examined whether there was a significant three-way interaction between Condition, EF composite, and Prior Knowledge and between Condition, Prior Knowledge, and each of the individual EF tasks. None of these interactions was significant.

Aim 5: EF as a predictor of expression versus construction of knowledge

In this aim, we planned to see if EF differentially predicts children's expression (i.e., achievement) and construction of knowledge (i.e., learning) by examining the individual EF tasks in addition to the EF composite score. First, we planned to compare the strength of EF as a predictor for children's expression of knowledge about sinking and floating that they had already acquired (pre-test) and for their construction or learning of sinking and floating knowledge after the intervention (post-test). As mentioned before, we found that the EF composite was not a significant predictor of children's prior sinking and floating knowledge (See Aim 1) or of their construction of new sinking and floating knowledge (See Aim 4). When we tested each of the individual EF tasks separately, BWS, MEFS, and Statue scores were not significant predictors of children's expression of their sinking and floating knowledge controlling for Age in Months at Pre-Test, Non-Verbal and Verbal IQ, and SES Composite Scores. To further explore the significant main effect of SES found earlier, we tested whether any of the individual EF tasks interacted with SES to predict children's expression of prior knowledge about sinking and floating controlling for our covariates. There was only a marginally significant MEFS x SES interaction (B = -.01, p = .06; See Figure 6). The effect of children's MEFS scores on their expression of their prior knowledge about sinking and floating decreased as children's SES increased. Among children from lower-income families, those with lower MEFS scores expressed less prior knowledge about sinking and floating than those with higher MEFS scores. For the construction of sinking and floating knowledge by children in the instructional groups, when tested individually, none of the three EF tasks (BWS, MEFS, and Statue Scores) were significant predictors when controlling for the covariates. Given EF was not a significant predictor of children's expression or construction of knowledge, it was not necessary to compare the strength of EF as a predictor for children's expression and construction of their sinking and floating knowledge.

Next, we examined whether each of the EF tasks was a significant predictor of children's expression of their math and literacy knowledge. These analyses revealed that BWS (*B* = 4.83, *p* = .04) and MEFS (*B* = .27, *p* = .001) were significant predictors of children's Applied Problems Scores after controlling for all the covariates. For children's Letter/Word scores, only BWS was a significant predictor after controlling for the covariates (*B* = 4.83, *p* =.02). We also further explored the significant SES x EF Composite interaction we found earlier for Applied Problems by examining whether each of the EF tasks interacted with SES to predict children's scores. These exploratory analyses indicated only a significant BWS x SES interaction (*B* = -5.43, *p* = .008) to predict children's Applied Problems scores after controlling for our covariates. The effect of BWS on Applied Problems performance decreased as children's SES increased. Children from lower-income families with lower BWS Scores performed worse on the Applied Problems subtest than children from lower-income families with higher BWS Scores.

We aimed to examine how learning sinking and floating knowledge compared to children's expression of their math and literacy academic achievement. Given we only found that EF was a significant predictor of children's math achievement (expression of knowledge), but not of the expression of their literacy knowledge or construction of sinking and floating knowledge, we were not able to compare the relations between EF and expression versus construction of knowledge across domains.

Discussion

Study 2 extended the findings from Study 1 by examining how EF predicted children's expression and construction of knowledge in a new science domain and whether individual differences in EF influenced the effectiveness of different instructional approaches for children's learning about sinking and floating. We did not find a main effect of EF for children's expression or construction of sinking and floating knowledge. The results of the study also did not support our hypotheses for EF interacting with instructional conditions to predict children's learning. However, exploratory analyses to better understand these null findings suggested the relation between EF and expression and construction of knowledge might be more complex than initially thought. These findings revealed that there are certain factors such as SES and prior knowledge that might interact with EF to predict the expression and construction of knowledge, respectively.

Expression of Knowledge

Expression of Sinking and Floating Knowledge

Study 2 extended the findings from previous studies on the relation between young children's EF and science learning in a new science domain that has not been previously studied in relation to EF in preschoolers. Contrary to our hypothesis, EF was not a significant predictor of children's expression of their prior knowledge about sinking and floating, controlling for Age in Months at Pre-Test, Non-Verbal and Verbal IQ, and SES. However, exploratory analyses examining the interaction between the individual EF tasks and SES indicated there was a marginally significant MEFS x SES interaction to predict children's expression of their prior knowledge about sinking and floating. Children from lower-income families with lower MEFS Scores expressed less prior knowledge about sinking and floating than children from lower-income families with higher MEFS Scores. Not finding a significant main effect of EF contradicts previous research showing that individual differences in EF predict children's expression of science knowledge in other domains (e.g., Nayfeld et al., 2013; Tardiff et al., 2020; Zaitchik et al., 2013). One possible reason that we found a trend for an interaction between EF and SES instead of a main effect of EF could be due to the differences in SES for the sample used in our study compared to previous studies on individual differences in EF and science achievement in preschoolers. Most of the previous studies were conducted with mostly middle-class, Caucasian families and did not control for SES or caregiver education levels (Tardiff et al., 2020; Zaitchik et al., 2013). Other studies included ethnicity as a covariate instead of family income or caregiver education levels (e.g., Nayfeld et al., 2013). In Study 2, we included an SES composite combining caregiver education level and family income as a covariate in our analyses. Our finding that EF may have mattered more for lower-income children's ability to express their science knowledge than higher-income children's aligns with research showing that EF may be especially important for children growing up in high-risk contexts (Razza, Martin, & Brooks-Gunn, 2010; Wright, Masten, & Narayan, 2013) and research showing that EF mediates the association between SES and academic achievement (e.g., Dilworth-Bart, 2012; Fitzpatrick, McKinnon, Blair, & Willoughby, 2014). More research on how SES may influence the relation between individual differences in EF and expression of science knowledge is needed to better understand how EF relates to young children's science learning.

Expression of Math, Literacy, and Science Knowledge

We found that EF was only a significant predictor of children's math knowledge after controlling for Age in Months at Pre-Test, Non-Verbal and Verbal IQ, and SES.

This was contrary to our hypothesis that individual differences in EF would predict children's expression of knowledge in all three domains above and beyond the covariates given previous research (e.g., Allan et al., 2014; Best et al., 2011; Blair & Razza, 2007; Jacob & Parkinson, 2015; Nayfeld et al., 2013; Zaitchik et al., 2013). While there was not a main effect of EF for children's expression of knowledge in all three domains, there was a significant main effect of SES with children from lower-income families expressing less knowledge than children from higher-income families. This result replicates findings on the achievement gap starting during the preschool years in school readiness skills (Bradley & Corwyn, 2002; Duncan, Yeung, Brooks-Gunn, & Smith, 1998; Entwisle, Alexander, & Olson, 2005). Additionally, exploratory analyses indicated that there was a significant EF Composite x SES interaction for Applied Problems, but not for Letter/Word ID. For Applied Problems, additional analyses revealed that there was a significant BWS x SES interaction such that children from lower-income families with lower BWS scores expressed less prior knowledge about math than children from lower-income families with higher BWS scores. For expression of math knowledge, it appears that higher EF is associated with more math knowledge for all children, but especially for children from lower-income families. The finding that EF was not a significant predictor of children's literacy achievement scores was unexpected. One reason could be that literacy skills might require less EF than other academic subjects such as math and science. For example, once a child has learned their letters and how to sound out and read words, this skill becomes less effortful and more automatic. Meta-analyses have shown that the association between EF and math is stronger than for EF and literacy (e.g. Allan et al., 2014; Jacob & Parkinson, 2015) and that EF is a stronger predictor of children's science and math knowledge than their literacy (Nayfeld et al., 2013). The bivariate correlations in the current study support the idea that EF was more strongly related to science (*r* =.67) and math (*r* = .56) than literacy (*r* = .29). The lack of an EF x SES interaction for letter/word controlling for covariates including verbal ability aligns with previous research examining EF as a mediator between preschoolers' SES and math and literacy achievement (Dilworth-Bart, 2012; Fitzpatrick et al., 2014). For example, Fitzpatrick et al. (2014) found that EF no longer mediated the association between letter/word and SES after accounting for children's verbal ability. However, EF still mediated the relation between Applied Problems and SES once verbal ability was accounted for. More research is needed with more varied measures of children's math, literacy, and science achievement and with academic measures that have equal EF demands to further understand how EF may explain the relation between SES and academic achievement.

Construction of Knowledge

Construction of Sinking and Floating Knowledge

Contrary to our hypothesis, we did not find that EF Composite Scores were a significant predictor of children's construction of sinking and floating knowledge across instructional conditions after receiving instruction and controlling for our covariates. This was contrary to the findings from Study 1, which showed that Cool EF predicted children's construction of conceptual knowledge about animals. The difference in results across the two studies could be because children were asked to learn different science topics in each study, and sinking and floating may be a harder topic for preschoolers to learn. Study 2 also included additional covariates than Study 1, such as non-verbal IQ and

SES providing a more stringent test of the unique contribution that EF skills for children’s learning. Lastly, children’s learning was correlated with SES in Study 2, which allowed us to examine how SES and EF may interact to influence children’s learning. This could have been due to recruiting families for Study 2 from childcare centers in both higher and lower-income areas, which may have resulted in more representative differences in SES than are found among the IPP families that volunteered to participate in Study 1.

Exploratory analyses looking at children’s prior knowledge about sinking and floating revealed there was a trend for an interaction between children’s prior knowledge and EF (more specifically their BWS scores) to predict their sinking and floating learning. Children with lower prior knowledge and lower EF learned less than children with lower prior knowledge and higher EF. This suggests that EF might be particularly important for learning when children have less prior knowledge about the material being learned. This finding aligns with Cognitive Load Theory (Sweller, 1988; Kirschner et al., 2006) which suggests that having prior knowledge is one way the working memory demands of a learning environment can be reduced. Within this view, having prior knowledge allows the learner to better process new information by being able to use schemas they have already based on their prior knowledge which requires less working memory resources. This reduction in working memory demands in turn makes it easier for the learner to learn new information. Therefore, if a learner has less prior knowledge about the learning material, differences in working memory capacities will be important for determining how much they learn. The finding from the current study also aligns with empirical studies that have shown that learners with greater prior knowledge about the learning material learn more than those with less prior knowledge across different types of instruction such as worked examples and discovery learning (e.g., Tuovinen & Sweller, 1999). This highlights the importance of considering the role of prior knowledge for learning in future studies examining the role of individual differences in EF and learning.

Differences in Construction of Knowledge Between Instructional Conditions

Another main finding from the study is that there were significant differences in learning about sinking and floating depending on experimental condition. As hypothesized, children learned significantly more about sinking and floating in the Direct Instruction condition than in the Discovery or Control conditions. However, children in the Discovery condition learned less than we expected. One reason why children in the Discovery condition learned less compared to studies with preschoolers on guided play is that the problem space for the sinking and floating instructional sessions might have been too large. Previous studies examining the differences between direct instruction and guided play with young children tend to use problem spaces that are smaller and more constrained in terms of the number of objects and activities that are included. For instance, in the Bonawitz et al. (2011) study, children only had to learn about one novel toy with four functions and in the Fisher et al. (2013) study, children learned about four different types of shapes using a limited number of supplies (e.g., 16 cards with shapes printed on them with Velcro on the back, a felt board to stick the cards on, and sticks used to make shapes). In the current study, there were several different aspects of the instructional sessions that created a large problem space for children to work with and to be distracted by such as the tub with water, a sink/float chart, several different objects, and two to three different activities per session. While constraining the problem space and having more control over it is useful, it decreases the ecological validity of the learning experience. Our study was higher on ecological validity but at the expense of being lower on constraining the problem space. Future research needs to figure out the best balance between these two factors to better understand how EF may be influencing children's learning from a discovery condition.

Another aspect of the instructional conditions that may have influenced less learning in the discovery condition is the amount and type of guidance that was used. In Study 2, we used open-ended questions such as why questions and prompting children to compare and contrast the different materials. This was chosen because pedagogical questions (i.e., questions instructors know the answers to) have been shown to be an effective teaching method that alerts the child to what they should pay attention to but also allows them to further explore in a guided way (Yu et al., 2018; Yu, Bonawitz, & Shafto, 2019). The use of questions also allowed children to practice explaining to themselves why they think objects sink and float. There is a large body of research showing that self-explanation can be an effective instructional strategy (e.g., Lee & Anderson, 2013; Rittle-Johnson, 2006). Questions are just one way that instructors can provide guidance in a discovery learning environment. Other ways that this guidance can be provided is through the way that the materials are presented (e.g., materials such as those designed for Montessori classrooms that allow the child to self-correct and be active in discovering new information) or through the use of feedback (Honomichl & Chen, 2012). Future research is needed to better understand how much guidance is needed in a discovery condition for young children's science learning and what types of guidance are most effective.

Lastly, a reason why we saw little learning overall and only in the Direct Instruction condition could be due to the fact that changing children's conceptual understanding is difficult and takes time. For instance, research shows that children's understanding of concepts such as theory of mind and biological concepts follows specific developmental patterns and takes place over years (e.g., Carey, 1985; Carey, 2000; Carey, 2009; Wellman & Gelman, 1992). Although it is possible to change children's beliefs and concepts during a single lab session, previous research shows that changing their beliefs about sinking and floating is challenging (e.g., Kloos & Sommerville, 2001; Kloos & van Orden, 2005). For example, Kloos and van Orden (2005) found that only 20% of children in their sample showed changes in their sinking and floating knowledge after one session. Therefore, it is possible that the duration of our instructional sessions (three 25-minute session for a total of 75 minutes of instruction) was not sufficient to change children's beliefs and understanding about sinking and floating and to see enough variability in children's sinking and floating scores to test whether EF was a significant predictor of their learning. Future research should explore the interaction between EF and the amount of instructional support in a context where children are able to get a higher dosage of the instruction.

We also did not find a significant interaction between EF and instructional condition to predict children's sinking and floating learning. This null finding could be due to a lack of power to detect an interaction effect. The final sample was smaller than initially planned due to several logistical issues with conducting research in the field, such as children leaving the center before data collection began or could be completed (e.g., a large number of children left centers for the summer right after data collection started), or we were unable to get exclusion criteria information from parents after they signed up. Another potential reason we did not find a significant EF by condition interaction is that there was very little learning that occurred in the Discovery Learning condition compared to the Direct Instruction condition, making it hard to see if EF moderated the amount of learning across the different instructional conditions.

Expression vs. Construction of Knowledge

The current study also compared how EF predicts the expression and construction of children's knowledge. We were able to compare how EF predicts the same group of children's expression (Pre-Test) and construction (Post-Test) of their sinking and floating knowledge, controlling for Age in Months at Pre-Test, Verbal and Non-Verbal IQ, and SES. Although we predicted that EF would predict both expression and construction of children's sinking and floating knowledge, we did not find a main effect of EF for children's expression of their prior sinking and floating knowledge or for their construction of new knowledge. This was unexpected based on previous research showing that EF predicts the same young children's expression and construction of knowledge in other science domains (Bascandziev et al., 2016) and math (Miller et al., 2016). However, we did find that preliminary evidence that EF may interact with other individual differences variables in our sample. For example, we found a marginally significant interaction between SES and MEFS Scores to predict children's expression of their sinking and floating knowledge and a trend for an interaction between Prior Knowledge and BWS to predict children's construction of sinking and floating knowledge, collapsed across the two instructional conditions. The trend for an interaction between our working memory measure (BWS) and prior knowledge for children's construction of learning aligns with previous research looking at how individual types of EF tasks (i.e., working memory, cognitive flexibility, and inhibitory control) predict expression and construction of knowledge in the same group of young children that have found that working memory predicts their construction of knowledge in science and math (Bacandziev et al., 2016; Miller et al., 2016). Additionally, the finding that cool EF measures were interacting with these individual differences variables to predict children's construction of conceptual knowledge about sinking and floating instead of the Statue Task (motor inhibition) is similar to the finding from Study 1 that cool EF predicted children's construction of conceptual information about animals.

We also examined whether EF differentially predicted children's construction of sinking and floating knowledge and expression of math and literacy knowledge. We found a main effect of the EF Composite Scores, as well as main effects of BWS and MEFS separately, and a significant BWS x SES interaction for expression of math knowledge on the Applied Problems subtest of the WJ. Similar to the relation between BWS and Prior Knowledge for predicting construction of sinking and floating knowledge, cool EF seemed to play a part in the expression of knowledge in math, a different academic domain. These results demonstrate that particular types of EF may be more important for expression versus construction of knowledge and that other individual differences variables such as SES and prior knowledge need to be considered to get the full picture of how EF is related to young children's academic achievement and learning. This points to the importance of future research to determine if academic achievement (i.e., expression of knowledge already acquired) and learning (i.e., construction of knowledge as a result of instruction) are different constructs and require different types of EF in preschoolers.

Strengths and Limitations

Study 2 had several strengths. First, a science domain that has not been previously studied in relation to preschoolers' individual differences in EF was investigated. This is important due to the lack of research on children's science learning and EF compared to other academic domains such as math and literacy. It also helps expand our knowledge about how EF may influence young children's learning across different topics in the same academic domain to better understand if EF plays a different role in learning different types of domain knowledge. For instance, meta-analyses looking at academic achievement and EF across different academic domains such as math and literacy suggests that EF is more strongly related to math than literacy (e.g., Allan et al., 2014; Jacob & Parkinson, 2015), which we replicated here, but more research is needed on whether this is the case for children's construction of new knowledge in these domains.

The use of an experimental design with a control group with children randomly assigned to conditions was also a strength of the current study. The inclusion of a control group allowed us to measure the amount of change that would occur due to maturation or attending a childcare center. We could then see if the amount of change that children in the instructional groups made in their knowledge of sinking and floating was significantly different from what would be expected without instruction. Study 2 also included a number of control measures that have been found to be related to children's achievement and EF such as IQ, age, and SES. This allowed for a robust test of the unique contribution that EF makes to children's expression and construction of knowledge.

Despite these strengths, Study 2 also had some limitations. First, there was not much variability in children's performance on the sinking and floating assessment used at pre- and post-test. Due to a lack of a validated, standardized assessment focused solely on sinking and floating for preschoolers, we created an assessment by borrowing tasks from previous sinking and floating studies with similar-aged children (Hong & Diamond, 2012; Kallery, 2015; Leuchter et al., 2014). However, the children in our sample did not perform at the same level as children in previous studies on some of the tasks. For example, in the Kallery (2015) study, the majority of the children performed well on the concept cartoons (e.g., ranged from 78% to 90% correct), whereas in our sample, children's accuracy ranged from 2.2 to 10.8% at pre-test and 3.2% to 16.1% at post-test. One possible reason for the difference in performance is that in Kallery's study, the intervention and assessments using the concept cartoons was done by classroom teachers with whom children were very familiar whereas the children in Study 2 were not familiar with the instructor who conducted the instructional sessions. For the Caterpillar Transfer Task, children in Study 2 also performed worse than children from the Leutcher et al. (2014). For example, 10.8% of children at pre-test and 20.4% of children at post-test in our sample received the highest score on the task, whereas 16.2% of children at pre-test and 40% at post-test received the highest score on the task in the Leutcher et al. (2014) study. While the pre-test scores on this task were comparable across studies, the Leutcher study had almost double the number of children who were able to get the highest score on the Caterpillar Transfer Task at post-test. One reason for this difference may be that the age range in the Leutcher et al. (2014) study was wider for the younger children (52-85 months) than the age range for our study (48 to 71 months). Another large difference between the two studies is that children in the Leutcher et al. (2014) received a much higher dosage of a sinking and floating intervention focused on material kind (4 weeks) compared to the children in Study 2 (75 minutes across 3 sessions). A more sensitive measure of children's sinking and floating knowledge across different samples and intervention dosages is needed to better understand how individual differences in EF relates to the expression and construction of children's sinking and floating knowledge.

Another limitation of the study was its small sample size, making us underpowered to detect interaction effects between EF and Condition. Recruiting and conducting study sessions in childcare centers, as well as the need to train and maintain a large cadre of research assistants to ensure double-blind assessment and instructional sessions, posed a number of logistical issues that made it harder to reach our goal sample size of 120. Future studies should include larger sample sizes and ensure they recruit children with a range of EF, prior knowledge, and SES to further tease apart how all these variables may interact to predict children's learning of sinking and floating.

Future Directions

The findings from Study 2 highlight several avenues for future research. One important direction is to conduct interventions on conceptual learning that are higher in dosage. In Study 2, children received three 25-minute sessions on separate days. However, other research with similar-aged children included sinking and floating interventions that varied in length from a single 30-minute session to several sessions across four weeks. Since conceptual change takes time, it is important to know how much instruction young children need to change their conceptual understanding of sinking and floating. After determining how much instruction is required, longitudinal studies that measure both children's EF and conceptual learning over time as children receive a more time-intensive intervention is necessary to understand how EF predicts the construction of sinking and floating knowledge.

Another area for future research is to determine the best discovery condition for young children's learning of sinking and floating. This would allow us to determine the amount and types of guidance that work best for preschoolers for this science domain. It would also ensure greater learning in the discovery condition than occurred in Study 2 to examine whether individual differences in EF moderate the effectiveness of a discovery learning versus a direct instruction approach.

Lastly, more research is needed that uses the same EF measures and similar types of interventions to strengthen our understanding of the relation between EF and the construction of knowledge. While it is helpful to learn if children's EF predicts their learning in a new academic area such as sinking and floating, it does not allow us to go more in-depth and understand the mechanisms for why individual differences in EF predict children's construction of knowledge. For instance, in Study 2, prior knowledge was an important factor that marginally interacted with children's EF to predict their construction of sinking and floating knowledge. This points to the importance of using similar measures and types of interventions to be able to dissect the learning environment, the characteristics of the learner, the material being taught, and how it is presented to get a more complete picture of how and why individual differences in EF predict children's construction of academic knowledge.

Conclusion

In Study 2, we found evidence that the relation between EF and expression and construction of sinking and floating knowledge is more complex than initially hypothesized. We found that EF interacted with SES and Prior Knowledge to predict children's expression and construction of sinking and floating knowledge, respectively. While we found that children learned more in the Direct Instruction condition than the other two conditions, children in the Discovery condition did not learn very much. We also did not find an EF x Instructional Condition interaction to predict children's construction of new sinking and floating knowledge after the intervention. However, future research is needed to test this question with a larger sample size, a discovery learning condition that is more conducive to learning for preschoolers, and a more sensitive measure of preschoolers' sinking and floating knowledge. Understanding how EF may interact with different instructional approaches is important for personalizing interventions and instructional approaches to enhance young children's academic success.

Chapter 4: General Conclusion

The goal of this dissertation was to better understand the relation between EF and learning. While there is a large body of research showing that EF is related to academic achievement and learning, less is known about why this relation exists. Understanding how EF interacts with different aspects of the learning environment is one way to better understand the relation between individual differences in EF and learning. Therefore, in Study 1, we looked at aspects of the material to be learned, and in Study 2, we examined how the content was presented to the learner.

In Study 1, we found that cool EF predicted 4-year-olds' conceptual learning about animals, and their motor inhibition predicted their factual learning about animals. Additionally, children's EF mattered more for their learning when the content conflicted with their prior knowledge. Therefore, Study 1 provided evidence that the type of knowledge (i.e., factual or conceptual) and the amount of conflict between the material to be learned and the learner's prior knowledge (i.e., no prior knowledge, no conflicting prior knowledge, and conflicting prior knowledge) are important to consider when examining the relation between EF and learning.

In Study 2, 4-5-year-olds were taught about sinking and floating using a direct instruction or a discovery instruction approach. Results indicated that EF marginally interacted with SES to predict children's expression of their prior knowledge about sinking and floating. While children in the Direct Instruction condition learned more than children in the Discovery Learning condition, we did not find a significant EF x Instructional Condition interaction. However, when we collapsed across instructional conditions, we found a marginally significant interaction between children's prior knowledge of sinking and floating and EF to predict their learning from the intervention. Therefore, Study 2 provided promising results for EF predicting the expression and construction of sinking and floating knowledge, but only for certain groups of children (i.e., lower-income and lower prior knowledge). Although we did not find that EF moderated the effectiveness of a direct versus a discovery learning approach for constructing knowledge about sinking and floating, several limitations made it difficult for us to find this that should be addressed in future research to provide a more rigorous test of this research question.

Overall, Study 1 and Study 2 added to our understanding of the relation between EF and learning in two different science topics: animals and sinking and floating. Both studies took a more in-depth look into the relation between individual differences in EF and learning by examining factors that have not been extensively studied, such as the type of information being learned and the instructional approaches used to present the information to the learner. Examining different aspects of the learning environment in relation to EF is important to help us better understand why EF is associated with learning and under which circumstances and for whom individual differences in EF matter most for learning. The findings from Study 1 and Study 2 suggest that young children's EF level should be taken into consideration when tailoring instructional approaches and interventions to optimize children's learning.

Bibliography

Ackerman, D. J., & Friedman-Krauss, A. H. (2017). Preschoolers' executive function: Importance, contributors, research needs and assessment options (Research Report No. RR-17-22). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12148

Alfieri, L., Brooks, P.J., Aldrich, N.J., & Tenenbaum, H.R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*(1), 1-18. doi: 10.1037/a0021017.

Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., & Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: A meta-analysis. *Developmental Psychology*, *50*(10), 2368–2379. http://doi.org/10.1037/a0037493

Baker, S. T., Gjersoe, N. L., Sibielska-Woch, K., Leslie, A. M., & Hood, B. M. (2011). Inhibitory control interacts with core knowledge in toddlers' manual search for an occluded object. *Developmental Science*, *14*(2), 270–279. http://doi.org/10.1111/j.1467-7687.2010.00972.x

Baker, H., Haubmann, A., Kloos, H., & Fisher, A. (2011). Preschoolers' learning about buoyancy: Does it help to give away the answer? *Proceedings of the First Joint International Conference on Learning and Development and Epigenetic Robotics*. Frankfurt, Germany: IEEE.

Baptista, J., Osório, A., Martins, E. C., Verissimo, M., & Martins, C. (2016). Does social-behavioral adjustment mediate the relation between executive function and academic readiness? *Journal of Applied Developmental Psychology*, *46*, 22–30. http://doi.org/10.1016/j.appdev.2016.05.004

Barrouillet, P., & Lépine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology*, *91*(3), 183–204. http://doi.org/10.1016/j.jecp.2005.03.002

Bascandziev, I., Powell, L. J., Harris, P. L., & Carey, S. (2016). A role for executive functions in explanatory understanding of the physical world. *Cognitive Development*, *39*, 71–85. http://doi.org/10.1016/j.cogdev.2016.04.001

Bascandziev, I., Tardiff, N., Zaitchik, D., & Carey, S. (2018). The role of domain-general cognitive resources in children's construction of a vitalist theory of biology. *Cognitive Psychology, 104*, 1-28. https://doi.org/10.1016/j.cogpsych.2018.03.002

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

Blackwell, T. L. (2001). Test Review: Woodcock, RW, McGrew, KS, & Mather, N. (2001). Woodcock-Johnson® III Test. Riverside Publishing Company. Itasca, IL. *Rehabilitation Counseling Bulletin*, *44*(4), 232-235.

Blair, C., & Razza, R.P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten, *Child Development*, *78*(2), 647–663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N.D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition, 120* (3), 322-330. https://doi.org/10.1016/j.cognition.2010.10.001

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*, 371–99. https://doi.org/10.1146/annurev.psych.53.100901.135233

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of "hot" and "cool" executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly*, *24*(3), 337–349. http://doi.org/10.1016/j.ecresq.2009.06.001

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31,* 21–32.

Butts, D.P., Hoffman, H., & Anderson, M. (1993). Is hands-on experience enough? A study of young children's views of sinking and floating objects. *Journal of Elementary Science Education, 5*, 50-64. https://doi.org/10.1007/BF03170644

Carey, S. (1985). Conceptual change in childhood. Cambridge, MA: MIT Press.

Carey, S. (2000). Science Education as Conceptual Change. *Journal of Applied Developmental Psychology*, *21*(1), 13–19. http://doi.org/10.1016/S0193-3973(99)00046-5

Carey, S. (2009). *The origin of concepts.* New York, NY: Oxford University Press.

Carlson, S.M. & Moses, C.J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032-1053. https://doi.org/10.1111/1467-8624.00333

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*(2), 73-92. doi: 10.1002/icd.298

Carlson, S.M. (2005). Developmentally Sensitive Measures of Executive Function in Preschool Children. *Developmental Neuropsychology, 28* (2), 595-616, doi: 10.1207/s15326942dn2802_3

Carlson, S. M., Zelazo, P. D., & Faja, S. (2013). Executive Function. In P.D. Zelazo (Ed.), *Oxford handbook of developmental psychology, Vol. 1: Body and mind* (pp. 706-743). New York: Oxford University Press.

Carlson, S. M., & Zelazo, P. D. (2014). Minnesota Executive Function Scale. Saint Paul, MN: Reflection Sciences, LLC

Carlson, S.M., Claxton, L.J., Moses, L.J. (2015). The relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development, 16*(1), 186-197. doi: 10.1080/15248372.2013.824883

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, *54*(1), 241–257. http://doi.org/10.1016/S0301-0511(00)00058-2

Casey, B.J., Somerville, L.H., Gotlib, I.H., Ayduk, O., Franklin, N.T., Askren, M.K., …Yuichi Shoda (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the National Academy of Sciences, 108*(36), 14998-15003. doi: 10.1073/pnas.1108561108

Caviola, S., Mammarella, I. C., Cornoldi, C., & Lucangeli, D. (2012). The involvement of working memory in children's exact and approximate mental addition. *Journal of Experimental Child Psychology*, *112*(2), 141–160. http://doi.org/10.1016/j.jecp.2012.02.005

Chi, M.T.H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.)*, International Handbook of Research on Conceptual Change* (pp. 61-82). Hillsdale, NJ: Erlbaum.

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. http://doi.org/10.1016/j.ecresq.2015.12.009

Cohen, M. T. (2008). The effect of direct instruction versus discovery learning on the understanding of science lessons by second grade students. *NERA Conference Proceedings*. Rocky Hill: CT.

Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience*, *18*, 34–48. http://doi.org/10.1016/j.dcn.2015.11.005

Dilworth-Bart, J.E. (2012). Does executive function mediate SES and home quality associations with academic readiness? *Early Childhood Research Quarterly, 27*(3). https://doi.org/10.1016/j.ecresq.2012.02.002

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How Much Does Childhood Poverty Affect the Life Chances of Children? *American Sociological Review*, *63*(3), 406–423. doi: 10.2307/2657556

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First Grade and Educational Attainment by Age 22: A New Story. *American Journal of Sociology*, *110*(5), 1458–1502. http://doi.org/10.1086/428444

Fantuzzo, J., Perry, M. A., & McDermott, P. (2004). Preschool approaches to learning and their relationship to other relevant classroom competencies for low-income children. *School Psychology Quarterly*, *19*(3), 212–230. http://doi.org/10.1521/scpq.19.3.212.40276

Fisher, K.R., Hirsh-Pasek, K., Newcombe, N., & Golinkoff, R.M. (2013). Taking shape: Supporting preschoolers' acquisition of geometric knowledge through guided play. *Child Development, 84*(6), 1872-1878. doi: 10.1111/cdev.12091

Fitzpatrick, C., McKinnon, R.D., Blair, C.B., & Willoughby, M.T. (2014). Do preschool executive function skills explain the school readiness gap between advantaged and disadvantaged children? *Learning and Instruction, 30*, 25-31. https://doi.org/10.1016/j.learninstruc.2013.11.003

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*(3), 493–513. http://doi.org/10.1037/0022-0663.97.3.493

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., … Changas, P. (2013a). Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology*, *105*(1), 58–77. http://doi.org/10.1037/a0030127

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., Jordan, N. C., Siegler, R., Gersten, R., & Changas, P. (2013b). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105*(3), 683–700. https://doi.org/10.1037/a0032446

Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A., … Changas, P. (2014). Does working memory moderate the effects of fraction intervention? An aptitude–treatment interaction. *Journal of Educational Psychology*, *106*(2), 499–514. http://doi.org/10.1037/a0034341

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

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

Han, M., Moore, N., Vukelich, C., & Buell, M. (2010). Does play make a difference? How play intervention affects the vocabulary learning of at-risk preschoolers. *American Journal of Play*, *3*(1), 82-105.

Hardy, I., Jonen, A., Moller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking." *Journal of Educational Psychology, 98(*2), 307-326. doi: 10.1037/0022-0663.98.2.307

Homer, B.D. & Plass, J.L. (2014). Level of interactivity and executive functions as predictors of learning in computer-based chemistry simulations. *Computers in Human Behavior, 36*, 365-375. https://doi.org/10.1016/j.chb.2014.03.041

Hong, S.H. & Diamond, K.E. (2012). Two approaches to teaching young children science concepts, vocabulary, and scientific problem-solving skills. *Early Childhood Research Quarterly, 27*, 295-305. doi:10.1016/j.ecresq.2011.09.006

Honomichl, R.D., & Chen, Z. (2012). The role of guidance in children's discovery learning. *WIREs Cognitive Science, 3*, 615-622. doi: 10.1002/wcs.1199.

Hsin, C-T., & Wu, H-K. (2011). Using scaffolding strategies to promote young children's scientific understandings of floating and sinking. *Journal of Science Education and Technology*, *20*, 656-666. doi: 10.1007/s10956-011-9310-7.

Jacob, R., & Parkinson, J. (2015). The Potential for School-Based Interventions That Target Executive Function to Improve Academic Achievement: A Review. *Review of Educational Research*, *85*(4), 512-522. http://doi.org/10.3102/0034654314561338

Kallery, M. (2015). Science in early years education: introducing floating and sinking as a property of matter. Internatioal jOurnal of Early Years Education, 23(1), 31-53. doi: 10.1080/09669760.2014.999646.

Keil, F. C. (1992). *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: The MIT Press.

Khng, K. H., & Lee, K. (2009). Inhibiting interference from prior knowledge: Arithmetic intrusions in algebra word problem solving. *Learning and Individual Differences*, *19*(2), 262–268. http://doi.org/10.1016/j.lindif.2009.01.004

Kirschner, P.A., Sweller, J., & Clark, R.E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86. doi: 10.1207/s15326985ep4102_1

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15*(10), 661-667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

Kloos, H., Fisher, A. & van Orden, G.C. (2010). Situated naïve physics: Task constraints decide what children know about density. *Journal of Experimental Psychology: General, 130*(4), 625-637. doi: 10.1037/a0020977

Kloos, H., & Fisher, A. (2011). Can a preschooler's mistaken belief benefit learning? *Swiss Journal of Psychology, 64*(3), 195-205. doi: 10.1024/1421-0185.64.3.195

Kloos, H., & Somerville, S.C. (2001). Providing impetus for conceptual change: The effect of organizing the input. *Cognitive Development, 16*(2), 737-759. https://doi.org/10.1016/S0885-2014(01)00053-3

Kohn, A.S. (1993). Preschoolers' reasoning about density: Will it float? *Child Development, 64*(6), 1637-1650. https://doi.org/10.1111/j.1467-8624.1993.tb04204.x

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A Developmental Neuropsychological Assessment*. Psychological Corporation

Laski, E. V., & Dulaney, A. (2015). When prior knowledge interferes, inhibitory control matters for learning: The case of numerical magnitude representations. *Journal of Educational Psychology, 107*(4), 1035–1050. https://doi.org/10.1037/edu0000034

Lee, H.S., & Anderson, J.R. (2013). Student learning: What has instruction got to do with it? *Annual Review of Psychology, 64*, 445-469. https://doi.org/10.1146/annurev-psych-113011-143833

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development*, *26*(3), 282–294. http://doi.org/10.1016/j.cogdev.2011.01.002

Leuchter, M., Saalbach, H., & Hardy, I. (2014). Designing science learning in the first years of schooling. An intervention study with sequenced learning material on the topic of 'floating and sinking.' *International Journal of Science Education, 36*(10), 1751-1771.doi: 10.1080/09500693.2013.878482

Markman, E. M., & Callanan, M. A. (1983). An analysis of hierarchical classification. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence*, Vol. 2, Erlbaum, Hillsdale, NJ.

Mayer, R.E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59* (1), 14-19. doi:10.1037/0003-066X.59.1.14

McKenzie, B., Bull, R., & Gray, C. (2003). The effects of phonological and visual-spatial interference on children's arithmetical performance. *Educational and Child Psychology*, *20*(4), 19–36.

Meindertsma, H.B., van Dijk,M.W.G., Steenbeek, H.W., & van Geert, P.L. (2014). Stability and variability in young children's understanding of floating and sinking during one single-task session. *Mind, Brain, and Education, 8*(3), 149-158. doi: https://doi.org/10.1111/mbe.12049

Miller, M. R., Rittle-Johnson, B., Loehr, A. M., & Fyfe, E. R. (2016). The Influence of Relational Knowledge and Executive Function on Preschoolers' Repeating Pattern Knowledge. *Journal of Cognition and Development*, *17*(1), 85–104. http://doi.org/10.1080/15248372.2015.1023307

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T.

D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions. *Current Directions in Psychological Science*, *21*(1), 8–14. http://doi.org/10.1177/0963721411429458

Modrek, A.S., Kuhn, D., Conway, A., & Arvidsson, T.S. (2019). Cognitive regulation, not behavioral regulation, predicts learning. *Learning and Instruction*, *60*, 237-244. https://doi.org/10.1016/j.learninstruc.2017.12.001

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2693–2698. http://doi.org/10.1073/pnas.1010076108

Nayfeld, I., Fuccillo, J., & Greenfield, D. B. (2013). Executive functions in early learning: Extending the relationship between executive functions and school readiness to science. *Learning and Individual Differences*, *26*, 81–88. http://doi.org/10.1016/j.lindif.2013.04.011

Nesbitt, K. T., Farran, D. C., & Fuhs, M. W. (2015). Executive function skills and academic achievement gains in prekindergarten: Contributions of learning-related behaviors. *Developmental Psychology*, *51*(7), 865–878. http://doi.org/10.1037/dev0000021

Neuenschwander, R., Röthlisberger, M., Cimeli, P., & Roebers, C. M. (2012). How do different aspects of self-regulation predict successful adaptation to school? *Journal of Experimental Child Psychology, 113*(3), 353–371. http://doi.org/10.1016/j.jecp.2012.07.004

Piaget, J. (1970). Science of education and the psychology of the child. Trans. D. Coltman. Orion.

Piekny, J. & Maehler, C. (2013). Scientific reasoning in early and middle childhood: The development of domain-general evidence evaluation, experimentation, and hypothesis generation skills. *British Journal of Developmental Psychology, 31*, 153-179. doi:10.1111/j.2044-835X.2012.02082.x

Plovt & Cyr, (2017). Toward a durable prevalence of scientific conceptions: Tracking the effects of two interfering misconceptions about buoyancy from preschoolers to science teachers. Journal of Research in Science Teaching, 54(9), 1121-1142. doi:10.1002/tea.21396

Portilla, X. A., Ballard, P. J., Adler, N. E., Boyce, W. T., & Obradović, J. (2014). An integrative view of school functioning: Transactions between self-regulation, school engagement, and teacher-child relationship quality. *Child Development*, *85*(5), 1915–1931. http://doi.org/10.1111/cdev.12259

Prager, E. O., Sera, M. D., & Carlson, S. M. (2016). Executive function and magnitude skills in preschool children. *Journal of Experimental Child Psychology*, *147*, 126–139. http://doi.org/10.1016/j.jecp.2016.01.002

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110–122. http://doi.org/10.1016/j.lindif.2009.10.005

Razza, R. A., Martin, A., & Brooks-Gunn, J. (2010). Associations among family environment, sustained attention, and school readiness for low-income children. *Developmental Psychology, 46*(6), 1528-1542. https://doi.org/10.1037/a0020389

Razza, R. A., Martin, A., & Brooks-Gunn, J. (2015). Are Approaches to Learning in Kindergarten Associated with Academic and Social Competence Similarly? *Child and Youth Care Forum*, *44*(6), 757–776. http://doi.org/10.1007/s10566-015-9307-0

Reuter, J.M. (2017). *Classroom preschool science learning: The learner, instructional tools, and peer-learning assignments*. Doctoral Dissertation retrieved from https://doi.org/doi:10.7282/T3PR7ZWS

Rhodes, S. M., Booth, J. N., Campbell, L. E., Blythe, R. A., Wheate, N. J., & Delibegovic, M. (2014). Evidence for a Role of Executive Functions in Learning Biology. *Infant and Child Development, 23*(1), 67–83. https://doi.org/10.1002/icd.1823

Rhodes, S. M., Booth, J. N., Palmer, L. E., Blythe, R. A., Delibegovic, M., & Wheate, N. J. (2016). Executive functions predict conceptual learning of science. *British*

Journal of Developmental Psychology, *34*(2), 261–275. http://doi.org/10.1111/bjdp.12129

Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children’s self-regulation and classroom quality to children’s adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, *45*(4), 958–972. http://doi.org/10.1037/a0015861

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development, 77*(1), 1-15. https://doi.org/10.1111/j.1467-8624.2006.00852.x

Robinson, K. M., & Dubé, A. K. (2013). Children’s additive concepts: Promoting understanding and the role of inhibition. *Learning and Individual Differences*, *23*(1), 101–107. http://doi.org/10.1016/j.lindif.2012.07.016

Roid, G. H. (2005). Stanford-Binet Intelligence Scales for Early Childhood manual (5th ed.). Itasca, IL: Riverside Publishing.

Sasser, T. R., Bierman, K. L., & Heinrichs, B. (2015). Executive functioning and school adjustment: The mediational role of pre-kindergarten learning-related behaviors. *Early Childhood Research Quarterly*, *30*(Part A), 70–79. http://doi.org/10.1016/j.ecresq.2014.09.001

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, *124*(2), 209–215. http://doi.org/10.1016/j.cognition.2012.04.005

Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: A case study of the development of the concepts of size, weight, and density, *Cognition, 21*, 177-237. https://doi.org/10.1016/0010-0277(85)90025-3

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.G. & Paas, F.G.W.C. Cognitive Architecture and Instructional Design. *Educational Psychology Review, 10*, 251–296 (1998). https://doi.org/10.1023/A:1022193728205

Tardiff, N., Bascandziev, I., Carey, S., & Zaitchik, D. (2020). Specifying the domain-general resources that contribute to conceptual construction: Evidence from the child's acquisition of vitalist biology. *Cognition, 195*, 1-12. https://doi.org/10.1016/j.cognition.2019.104090

Tenenbaum, H.R., Rappolt-Schlichtmann, G., & Zanger V.V. (2004). Children's learning about water in a museum and in the classroom. *Early Childhood Research Quarterly, 19*, 40-58. doi: 10.1016/j.ecresq.2004.01.008

Toub, T.S, Rajan, V., Golinkoff, R.M., & Hirsh-Pasek, K. (2016). Guided play: A solution to the play versus learning dichotomy. In D.C. Geary & D.B Berch (Eds.), *Evolutional perspectives on child development* (pp.117-141). New York, NY: Springer.

Tuovinen, J.E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology, 91*(2), 334-341. https://doi.org/10.1037/0022-0663.91.2.334

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). "mice: Multivariate imputation by chained equations in R." *Journal of Statistical Software*, *45*(3), 1-67. https://www.jstatsoft.org/v45/i03/.

Van der Graaf, J., Segers, E., & Verhoeven, L. (2018). Individual differences in the development of scientific thinking in kindergarten. *Learning and Instruction, 56*, 1-9. https://doi.org/10.1016/j.learninstruc.2018.03.005

van Ginkel, J. R. (2019). Significance tests and estimates for R^2 for multiple regression in multiply imputed datasets: A cautionary note on earlier findings, and alternative solutions. *Multivariate behavioral research, 54,* 514-529. https://doi.org/10.1080/00273171.2018.1540967

Vitiello, V. E., & Greenfield, D. B. (2017). Executive functions and approaches to learning in predicting school readiness. *Journal of Applied Developmental Psychology*, *53*, 1–9. http://doi.org/10.1016/j.appdev.2017.08.004

Vosniadou, S. (2014). Examining cognitive development from a conceptual change point of view: The framework theory approach. *European Journal of Developmental Psychology, 11*(6), 645-661. http://dx.doi.org/10.1080/17405629.2014.921153

Vosniadou, S., Pnevmantikos, D., Makris, N., Ikospentaki, K., Lepenioti, D., Chountala, A., & Kyrianakis, G. (2015). Executive functions and conceptual change in science and mathematics learning. In R. Dale, C. Jennings, P. P. Maglio, T. Matlock, D. C. Noelle, A. Warlaumont, & J. Yoshimi (Eds.), *Proceedings of the 37th annual meeting of the Cognitive Science Society: Mind, Technology, and Society*. Pasadena, CA (pp. 2529–2534). Merced, CA: University of California.

Vygotsky, L. S. (1978). Mind in society: Development of higher psychological processes.

(M.Cole, V. John-Steiner, S. Scribner, & E. Soutberman, Eds.). Cambridge, MA: Harvard University Press.

Weisberg, Hirsh-Pasek, & Gonlinkoff, (2013). Guided play: Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education, 7*(2), 104-112. https://doi.org/10.1111/mbe.12015

Weisberg, D.S., Hirsh-Pasek,K., Golinkoff, R.M., Kittredge, A.K., & Klahr, D. (2016). Guided play: Principles and practices. *Current Directions in Psychological Science, 25*(3), 177-182. https://doi.org/10.1177/0963721416645512

Wellman, H. M., & Gelman, S. A. (1992). Cognitive Development: Foundational Theories of Core Domains. *Annual Review of Psychology*, *43*(1), 337–375. http://doi.org/10.1146/annurev.ps.43.020192.002005

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

Wright, M. O. D., Masten, A. S., & Narayan, A. J. (2013). Resilience processes in development: Four waves of research on positive adaptation in the context of adversity. In *Handbook of Resilience in Children: Second Edition* (pp.15-37). https://doi.org/10.1007/978-1-4614-3661-4_2

Xenidou-Dervou, I., van Lieshout, E. C. D. M., & van der Schoot, M. (2014). Working memory in nonsymbolic approximate arithmetic processing: A dual-task study with preschoolers. *Cognitive Science*, *38*(1), 101–127. http://doi.org/10.1111/cogs.12053

Yeniad, N., Malda, M., Mesman, J., Van Ijzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, *23*(1), 1–9. http://doi.org/10.1016/j.lindif.2012.10.004

Yu, Y., Bonawitz, E., & Shafto, P. (2019). Pedagogical questions in parent-child conversations. *Child Development, 90* (1), 147-161. https://doi.org/10.1111/cdev.12850

Yu, Y., Landrum, A.R., Bonawitz, E., & Shafto, P. (2018). Questioning supports effective transmission of knowledge and increased exploratory learning in pre-kindergarten children. *Developmental Science*, *21*(6), 1-8. https://doi.org/10.1111/desc.12696

Zaitchik, D., Iqbal, Y., & Carey, S. (2014). The Effect of Executive Function on Biological Reasoning in Young Children: An Individual Differences Study. *Child Development*, *85*(1), 160–175. http://doi.org/10.1111/cdev.12145

Zaitchik, D., Solomon, G. E. A., Tardiff, N., & Bascandziev, I. (2016). Conceptual change: Where domain-specific learning mechanisms meet domain-general cognitive resources. In D. Barner & A. S. Baron (Eds.), *Core knowledge and conceptual change* (pp. 73– 88). Oxford, UK: Oxford University Press.

Table 1

Descriptive statistics for all measures

Measure	**n**	**Range**	***M (SD)***
Age (in months)	61	49-59	53.72 (2.71)
Stanford Binet Verbal (raw score)	59	5-15	10.08 (2.59)
FWS (highest level)	61	2-5	3.62 (.71)
BWS (highest level)	61	1-4	1.82 (.85)
MEFS (max of 100)	61	17-78	46.22 (10.34)
Statue (max of 30)	55	0-30	21.11 (8.47)
Animal Familiarity Task (max of 17)	61	3-12	7.90 (1.53)
No Conflict, Factual- Memory Check Questions (max of 2)	61	0-2	1.34 (.70)
No Conflict, Factual- Comprehension Questions (max of 4)	61	0-4	3.15 (1.03)
No Conflict, Factual- Application Questions (max of 2)	61	0-2	.82(.72)
No Prior Knowledge, Conceptual- Practice Trials Score (max of 3)	61	0-3	1.43 (.78)
No Prior Knowledge, Conceptual- Test Trials Total Score (max of 9)	61	5-9	7.31 (1.21)
No Prior Knowledge, Conceptual- Test Trials Reasoning Score (max of 9)	61	0-9	4.52 (2.62)
No Conflicting Prior Knowledge, Conceptual- Pretest (max of 1)	61	0-1	.05(.22)
No Conflicting Prior Knowledge, Conceptual- Practice Trials Score (max of 3)	61	0-3	1.79 (.71)
No Conflicting Prior Knowledge, Conceptual - Test Trials Total Score (max of 8)	61	0-7	4.13 (1.49)
No Conflicting Prior Knowledge, Conceptual- Test Trials Reasoning Score (max of 8)	61	0-7	2.21 (2.25)
No Conflicting Prior Knowledge, Conceptual- # of Questions Asked Mr. Hippo (max is unlimited)	60	0-5	.47 (.98)
Conflicting, Prior Knowledge, Conceptual- Pretest Score (max of 1)	61	0-1	.21 (.41)
Conflicting, Prior Knowledge, Conceptual- Learning Trial Score (max of 1)	61	0-1	.26 (.44)
Conflicting, Prior Knowledge, Conceptual- Posttest/Test Total Score (max of 2)	61	0-2	1.05 (.87)
Conflicting, Prior Knowledge, Conceptual- Reasoning Score for Test Trials (max of 2)	61	0-2	.48 (.81)

Note. MEFS = Minnesota Executive Function Scale. FWS =Forward Word Span. BWS = Backward Word Span.

Table 2

Bivariate correlations for study variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Child Age	1														
2. Child Gender (Female)	.06	1													
3. SB Verbal (*N*=59)	-.10	.01	1												
4. Prior Knowledge	.25*	-.12	.16	1											
5. FWS	.02	-.10	.43**	.29*	1										
6. BWS	.13	-.07	.20	.12	.36**	1									
7. MEFS	.15	.16	.12	.27*	.31*	.42**	1								
8. Statue (*N* =55)	-.14	.33*	.17	-.02	.13	.00	.19	1							
9. Cool EF Composite	.15	.05	.28*	.31*	.64**	.71**	.88**	.17	1						
10. Factual Learning Total Score	.08	.09	.35**	.25	.34**	.12	.27*	.36**	.32*	1					
11. Conceptual, No Prior Knowledge Total Score	.06	.16	.55**	.33**	.46**	.26*	.30*	.35**	.42**	.43**	1				
12. Conceptual, No Conflicting Prior Knowledge Total Score	.07	.16	.19	.53**	.34**	.20	.36**	.06	.40**	.18	.45**	1			
13. Conceptual, Conflicting Prior Knowledge Total Score	.05	.11	.29*	.54**	.33**	.29*	.39**	.14	.45**	.18	.57**	.51**	1		
14. Overall Conceptual Learning Composite	.07	.17	.42**	.57**	.46**	.29*	.42**	.22	.52**	.32*	.82**	.80**	.84**	1	
15. Overall Learning Composite	.09	.17	.47**	.56**	.50**	.29*	.45**	.31*	.54**	.61**	.83*	.73**	.77**	.95**	1

Note. SB Verbal = Stanford Binet Intelligence Scale- Verbal Subtest. FWS =Forward Word Span. BWS = Backward Word Span. MEFS = Minnesota Executive Function Scale. EF = Executive Function. *N* =61, unless otherwise noted.
*$p < .05$, ** $p < .01$, *** $p < .001$.

Table 3

Regression analyses with EF predicting Overall Learning Composite Score

Variable	***B (SE)***	***t***	***p***	R^2	ΔR^2
Block 1				.23	-
Age in Months	.04 (0.031)	1.18	.24		
SB Verbal IQ	.14 (0.033)	4.10	.000		
Block 2				.46	.23***
Age in Months	-.0003(.03)	-.01	.99		
SB Verbal IQ	.11(.03)	3.81	.000		
Prior Knowledge Score	1.57 (.32)	4.86	.000		
Block 3				.59	.13**
Age in Months	-.004(.02)	-.15	.878		
SB Verbal IQ	.08 (0.03)	3.04	.004		
Prior Knowledge Score	1.35(.30)	4.53	.000		
Cool EF Composite	.30 (0.09)	3.30	.002		
Statue Total Score	.02 (0.008)	1.92	.062		

Note. SB = Stanford Binet Intelligence Scale. EF = Executive Function. *N* = 61
**p* < .05, ** *p* < .01., ****p* <.000

Table 4

Regression analyses with EF predicting Overall Conceptual Composite Score

Variable	***B (SE)***	***t***	***p***	***R²***	***ΔR²***
Block 1				.18	-
Age in Months	.035 (0.036)	.97	.33		
SB Verbal IQ	.13 (0.038)	3.54	.001		
Block 2				.43	.25***
Age in Months	-.008(.032)	-.26	.80		
SB Verbal IQ	.10(.033)	3.16	.003		
Prior Knowledge Score	1.84(.37)	4.96	.000		
Block 3				.53	.10**
Age in Months	-.01 (.03)	-.49	.63		
SB Verbal IQ	.08 (.03)	2.40	.02		
Prior Knowledge Score	1.59(.35)	4.47	.000		
Cool EF Composite	.32 (0.11)	3.03	.004		
Statue Total Score	0.01 (0.01)	1.10	.28		

Note. SB = Stanford Binet Intelligence Scale. EF = Executive Function. *N* = 61
**p* < .05, ** *p* < .01., ****p* <.000

Table 5

Regression analyses with EF predicting Factual Learning Total Score

Variable	***B (SE)***	***t***	***p***	***R^2***	***ΔR^2***
Block 1				.13	-
Age in Months	.06 (0.07)	.92	.36		
SB Verbal IQ	0.20 (0.07)	2.85	.006		
Block 2				.16	.03
Age in Months	.03 (.07)	.50	.62		
SB Verbal IQ	.18(.07)	2.55	.01		
Prior Knowledge Score	1.13(.81)	1.40	.17		
Block 3				.27	.11*
Age in Months	.04 (.07)	.62	.54		
SB Verbal IQ	.13 (.07)	1.89	.06		
Prior Knowledge Score	.92 (.80)	1.15	.25		
Cool EF Composite	.31 (.24)	1.28	.21		
Statue Total Score	0.05 (.02)	2.09	.04		

Note. SB = Stanford Binet Intelligence Scale. EF = Executive Function. *N*= 61
**p* < .05, ** *p* < .01., ****p* <.000

Table 6

Hierarchical linear model analyses to examine the interaction between Cool EF and Task Conflict predicting children's Conceptual Learning Scores

Variable	***B (SE)***	***t***	***p***
Cool EF Composite	.16 (.04)	4.09	.000
Statue Total Score	.002(.003)	.82	.42
No Conflicting Prior Knowledge	.015(.04)	.43	.67
No Prior Knowledge	.28(.04)	7.74	.000
Age in Months	-.005(.008)	-.67	.50
SB Verbal IQ	.017(.008)	2.06	.04
Prior Knowledge Score	.43(.09)	4.64	.000
Cool EF x No Conflicting Prior Knowledge	-.11(.05)	-2.27	.02
Cool EF x No Prior Knowledge	-.11(.05)	-2.43	.02

Note. SB = Stanford Binet Intelligence Scale. EF = Executive Function. Reference group = conflicting prior knowledge task. *N* =61

Table 7

Descriptive statistics for all measures

Measure	***n***	**Range**	***M (SD)***
MEFS	93	11-91	48.9 (16.3)
FWS	92	1-5	3.62 (.74)
BWS	90	1-3	2.01 (.74)
Statue	90	0-30	21.53 (8.59)
SB Non-Verbal Standard Score	91	1-19	10.46 (3.80)
SB Verbal Standard Score	89	3-15	9.89 (2.41)
SB Summed Standard Scores	87	7-31	20.51 (5.00)
Letter/Word Standard Score	90	69-141	103.19 (13.18)
Applied Problems Standard Score	88	70-137	110.19 (12.72)
Pre-Test Definition of Float (max of 2)	93	0-2	.76 (.81)
Post-Test Definition of Float (max of 2)	93	0-2	.91 (.87)
Pre-Test Definition of Sink (max of 2)	93	0-2	.84 (.90)
Post-Test Definition of Sink (max of 2)	93	0-2	.85 (.88)
Pre-Test Material Kind Frequency	93	0-12	1.34 (2.44)
Post-Test Material Kind Frequency	93	0-14	3.95 (4.31)
Pre-Test Prediction Accuracy Score (max of 14)	93	3-12	7.59 (1.82)
Post-Test Prediction Accuracy Score (max of 14)	93	3-14	8.94 (2.34)
Pre-Test Explanation Score (max of 70)	93	0-70	34.81 (14.93)
Post-Test Explanation Score (max of 70)	93	0-70	37.86 (17.95)
Pre-Test Total Accuracy Score for Concept Cartoons (0-6)	82	0-6	3.07 (1.44)
Post-Test Total Accuracy Score for Concept Cartoons (0-6)	89	0-6	3.37 (1.48)
Pre-Test Total Misconception Score (max of 6)	91	0-5	1.88 (1.40)
Post-Test Total Misconception Score (max of 6)	92	0-6	2.15 (1.50)
Pre-Test Total Concept Cartoon Score (accuracy + reasoning) (max of 12)	80	2-11	5.03 (1.96)
Post-Test Total Concept Cartoon Score (accuracy + reasoning) (max of 12)	89	0-11	5.54 (2.31)
Pre-Test Sand Jar Highest Score Received (max of 4)	93	0-4	2.23(1.51)
Post-Test Sand Jar Highest Score Received (max of 4)	93	0-4	2.26 (1.55)
Pre-Test Plastic Bottle Highest Score Received (max of 4)	93	0-4	2.22 (1.44)
Post-Test Plastic Bottle Highest Score Received (max of 4)	93	0-4	2.39 (1.54)
Pre-Test Caterpillar Score (max of 2)	93	0-2	.44 (.68)
Post-Test Caterpillar Score (max of 2)	93	0-2	.60 (.81)

Note. MEFS = Minnesota Executive Function Scale. FWS= Forward Word Span. BWS = Backward Word Span. SB = Stanford Binet Intelligence Scale

Table 8

Bivariate correlations among all variables

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Child Age	1												
2. Child Gender (1= female)	.03	1											
3. MEFS	.46**	.004	1										
4. FWS (*n* =92)	.27**	-.008	.33**	1									
5. BWS (*n* =90)	.38**	.14	.44**	.38**	1								
6. Statue (*n*=90)	.30**	.17	.26*	.26*	.32**	1							
7. EF Composite (*n* =89)	.52**	.13	.76**	.43**	.77**	.71**	1						
8. SB Non-Verbal IQ (*n* =91)	.04	.14	.28**	.09	.19	.14	.26*	1					
9. SB Verbal IQ (*n* =89)	.18	.18	.36**	.24*	.29**	.30**	.43**	.33**	1				
10. WJ Letter Word (*n* =90)	-.15	.04	.25*	.04	.36**	.07	.29**	.39**	.19	1			
11. WJ Applied Problems (*n* =88)	-.02	-.02	.50**	.29**	.48**	.28*	.56**	.42**	.51**	.52**	1		
12. Pre-Float Definition	.23*	-.02	.25*	.18	.25*	.05	.25*	.001	.27*	.09	.26*	1	
13. Post-Float Definition	.18	.05	.31**	.02	.21*	.25*	.36**	.35**	.47**	.12	.26*	.51**	1
14. Pre-Sink Definition	.34**	.05	.32**	.19	.25*	.24*	.35**	.20	.41**	.10	.31**	.62**	.48**
15. Post-Sink Definition	.38**	.06	.31**	.13	.24*	.22*	.34**	.33**	.41**	.10	.20	.33**	.76**
16. Pre-Material Kind Frequency	.19	.12	.22*	.12	.20	.28**	.29**	.14	.15	.19	.10	.07	-.11
17. Post-Material Kind Frequency	.21*	.15	.28**	.14	.24*	.13	.28**	.43**	.36**	.23*	.26*	.23*	.20
18. Pre-Total Prediction Score	.03	-.05	.09	.01	.05	.07	.10	-.03	.17	-.01	.02	.26*	.12
19. Post-Total Prediction Score	.37**	.12	.20	.11	.26*	.21*	.29*	.14	.30**	.09	.06	.16	.16
20. Pre-Explanation Score	.44**	.09	.51**	.33*	.41**	.36**	.59**	.11	.51**	.14	.43**	.39**	.28**
21. Post-Explanation Score	.38**	.27*	.42**	.18	.34**	.29**	.46**	.31**	.62**	.19	.29**	.29**	.34**
22. Pre-Concept Cartoon Total Score (Accuracy +Reasoning) (*n* =80)	.29**	.01	.29**	.21	.30**	.29*	.39**	.01	.35**	.04	.30**	.23*	.29**
23. Post-Concept Cartoon Total Score (Accuracy +Reasoning) (*n* =89)	.14	.08	.29**	.08	.25*	.11	.30**	.15	.42**	.19	.32**	.26*	.29**
24. Pre-Sand Jar Highest Score	.46**	-.04	.40**	.24*	.26*	.15	.36**	.28**	.39**	.12	.32**	.30**	.29**
25. Post-Sand Jar Highest Score	.37**	-.01	.36**	.21*	.22*	.11	.31**	.24*	.17	.05	.24*	.15	.11
26. Pre-Plastic Bottle Highest Score	.44**	.04	.20	.12	.12	.16	.22*	.002	.20	-.11	.00	.15	.23*
27. Post-Plastic Bottle Highest Score	.50**	.00	.32**	.30**	.35**	.28**	.43**	.20	.35**	.09	.18	.13	.31**
28. Pre-Caterpillar Total Score	.11	-.01	.21*	.10	.32**	.14	.31**	-.04	.11	.06	.09	.11	.03
29. Post-Caterpillar Total Score	-.03	.15	.01	.07	-.01	-.01	.003	.11	.19	.04	.13	.15	.21*
30. Pre-Sink/Float Learning Composite (*n* =80)	.57**	.03	.58**	.31**	.46**	.48**	.67**	.21	.47**	.18	.42**	.33**	.27*
31. Post-Sink/Float Learning Composite (*n* =89)	.52**	.08	.52**	.29**	.41**	.28**	.53**	.29**	.49**	.16	.36**	.31**	.30**
32. Family Income (*n* =81)	.12	.10	.40**	.15	.53**	.23*	.52**	.30**	.50**	.42**	.59**	.31**	.34**
33. Caregiver Education (*n* =83)	-.004	.18	.22	.04	.29**	.19	.31**	.22*	.42**	.44**	.43**	.26*	.29**

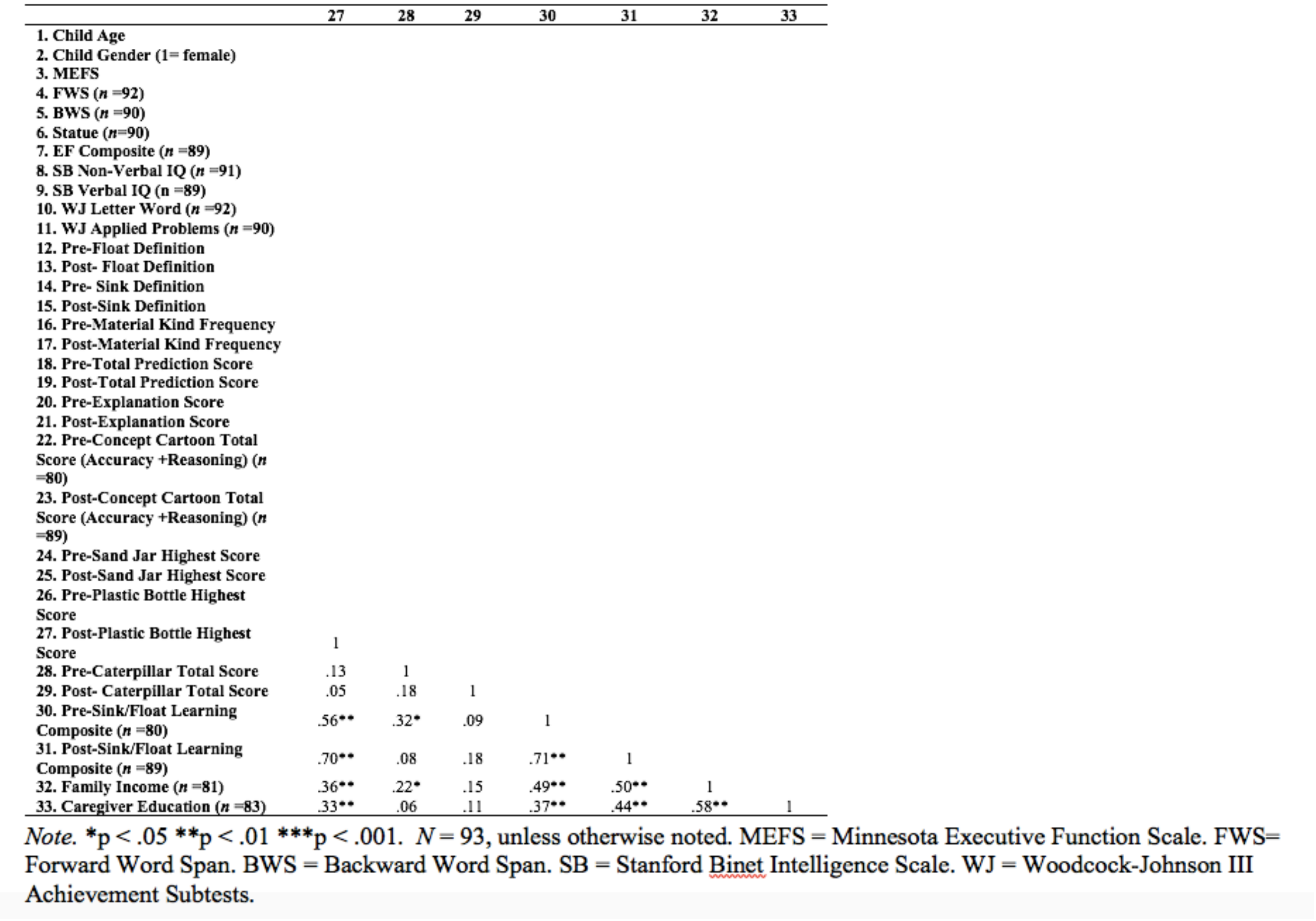
Table 8 Continued

Bivariate correlations among all variables

	14	15	16	17	18	19	20	21	22	23	24	25	26
1. Child Age													
2. Child Gender (1= female)													
3. MEFS													
4. FWS (*n* =92)													
5. BWS (*n* =90)													
6. Statue (*n*=90)													
7. EF Composite (*n* =89)													
8. SB Non-Verbal IQ (*n* =91)													
9. SB Verbal IQ (*n* =89)													
10. WJ Letter Word (*n* =92)													
11. WJ Applied Problems (*n* =90)													
12. Pre-Float Definition													
13. Post-Float Definition													
14. Pre-Sink Definition	1												
15. Post-Sink Definition	.45**	1											
16. Pre-Material Kind Frequency	.18	-.07	1										
17. Post-Material Kind Frequency	.31**	.16	.37**	1									
18. Pre-Total Prediction Score	.02	.05	.15	.15	1								
19. Post-Total Prediction Score	.22*	.25*	.33**	.65**	.13	1							
20. Pre-Explanation Score	.35**	.23*	.33**	.28**	.27**	.32**	1						
21. Post-Explanation Score	.39**	.33*	.27**	.74**	.16	.58**	.64**	1					
22. Pre-Concept Cartoon Total Score (Accuracy +Reasoning) (*n* =80)	.34**	.31**	.06	.11	-.01	.24*	.49**	.34**	1				
23. Post-Concept Cartoon Total Score (Accuracy +Reasoning) (*n* =89)	.24*	.21*	.19	.44**	-.03	.36**	.32**	.53**	.37**	1			
24. Pre-Sand Jar Highest Score	.33**	.39**	.31**	.35**	.05	.44**	.42**	.43**	.27*	.225*	1		
25. Post-Sand Jar Highest Score	.18	.21*	.19	.32**	-.09	.41**	.35**	.43**	-.07	.24*	.47**	1	
26. Pre-Plastic Bottle Highest Score	.25*	.31**	.19	.24*	.14	.40**	.28**	.34**	.23*	.19	.33**	.35**	1
27. Post-Plastic Bottle Highest Score	.23*	.42**	.18	.32**	.08	.36**	.42**	.47**	.20	.23*	.56**	.47**	.40**
28. Pre-Caterpillar Total Score	.10	-.09	.23*	-.05	.11	-.04	.28**	.004	.07	.00	.16	.13	.01
29. Post-Caterpillar Total Score	.09	.17	-.07	.26*	.19	.18	.14	.22*	-.04	.07	.06	.18	.09
30. Pre-Sink/Float Learning Composite (*n* =80)	.40**	.32**	.53**	.40**	.42**	.51**	.78**	.60**	.56**	.30**	.64**	.39**	.57**
31. Post-Sink/Float Learning Composite (*n* =89)	.38**	.33**	.35**	.69**	.12	.72**	.63**	.81**	.30**	.61**	.62**	.70**	.46**
32. Family Income (*n* =81)	.30**	.28*	.22	.27*	.18	.27*	.53**	.48**	.28*	.30**	.20	.35**	.05
33. Caregiver Education (*n* =83)	.27*	.18	.18	.36**	.21	.24*	.27*	.43**	.17	.33**	.14	.21	.13

Table 8 Continued

Bivariate correlations among all variables



	27	28	29	30	31	32	33
1. Child Age							
2. Child Gender (1= female)							
3. MEFS							
4. FWS (*n* =92)							
5. BWS (*n* =90)							
6. Statue (*n*=90)							
7. EF Composite (*n* =89)							
8. SB Non-Verbal IQ (*n* =91)							
9. SB Verbal IQ (n =89)							
10. WJ Letter Word (*n* =92)							
11. WJ Applied Problems (*n* =90)							
12. Pre-Float Definition							
13. Post- Float Definition							
14. Pre- Sink Definition							
15. Post-Sink Definition							
16. Pre-Material Kind Frequency							
17. Post-Material Kind Frequency							
18. Pre-Total Prediction Score							
19. Post-Total Prediction Score							
20. Pre-Explanation Score							
21. Post-Explanation Score							
22. Pre-Concept Cartoon Total Score (Accuracy +Reasoning) (*n* =80)							
23. Post-Concept Cartoon Total Score (Accuracy +Reasoning) (*n* =89)							
24. Pre-Sand Jar Highest Score							
25. Post-Sand Jar Highest Score							
26. Pre-Plastic Bottle Highest Score							
27. Post-Plastic Bottle Highest Score	1						
28. Pre-Caterpillar Total Score	.13	1					
29. Post- Caterpillar Total Score	.05	.18	1				
30. Pre-Sink/Float Learning Composite (*n* =80)	.56**	.32*	.09	1			
31. Post-Sink/Float Learning Composite (*n* =89)	.70**	.08	.18	.71**	1		
32. Family Income (*n* =81)	.36**	.22*	.15	.49**	.50**	1	
33. Caregiver Education (*n* =83)	.33**	.06	.11	.37**	.44**	.58**	1

Note. *p < .05 **p < .01 ***p < .001. *N* = 93, unless otherwise noted. MEFS = Minnesota Executive Function Scale. FWS= Forward Word Span. BWS = Backward Word Span. SB = Stanford Binet Intelligence Scale. WJ = Woodcock-Johnson III Achievement Subtests.

Table 9

Regression analyses with EF Composite predicting Pre-Test Sinking and Floating Overall Learning Composite Score

Variable	*B (SE)*	*T*	*p*	R^2	ΔR^2
Block 1				.41	-
Age in Months	.05 (.01)	5.16	< .001		
Non-Verbal IQ	.006 (.01)	.44	.66		
Verbal IQ	.05 (.02)	2.61	.009		
SES Composite	.18 (.07)	2.69	.007		
Block 2				.43	.02
Age in Months	.04 (.01)	3.47	.001		
Non-Verbal IQ	.002 (.01)	.11	.91		
Verbal IQ	.05 (.02)	2.23	.03		
SES Composite	.13 (.07)	1.80	.07		
EF Composite	.15 (.11)	1.42	.16		

Note. EF = Executive Function. SES= Socioeconomic Status. *N* = 93
*$p < .05$, ** $p < .01$.

Table 10

Regression analyses with EF Composite predicting Woodcock-Johnson Letter/Word Scores

Variable	*B (SE)*	*t*	*p*	R^2	ΔR^2
Block 1				.30	-
Age in Months	-.37 (.23)	-1.62	.11		
Non-Verbal IQ	.94 (.36)	2.61	.009		
Verbal IQ	-.23 (.52)	-.45	.66		
SES Composite	6.54 (1.77)	3.69	<.001		
Block 2				.33	.03
Age in Months	-.64 (.27)	-2.37	.02		
Non-Verbal IQ	.80 (.36)	2.22	.03		
Verbal IQ	-.43 (.51)	-.83	.41		
SES Composite	5.20 (1.89)	2.76	.007		
EF Composite	4.62 (2.49)	1.86	.07		

Note. EF = Executive Function. SES= Socioeconomic Status. *N* =93.
$^*p < .05$, $^{**} p < .01$.

Table 11

Regression analyses with EF Composite predicting Woodcock-Johnson Applied Problems Scores

Variable	*B (SE)*	*t*	*p*	R^2	ΔR^2
Block 1				.45	-
Age in Months	-.09 (.20)	-.43	.67		
Non-Verbal IQ	1.01 (.34)	2.95	.004		
Verbal IQ	1.28 (.46)	2.80	.005		
SES Composite	6.42 (1.46)	4.39	<.001		
Block 2				.56	.11**
Age in Months	-.56 (.25)	-2.28	.03		
Non-Verbal IQ	.77 (.30)	2.58	.01		
Verbal IQ	.94 (.42)	2.27	.02		
SES Composite	4.09 (1.55)	2.64	.009		
EF Composite	8.00 (2.63)	3.04	.005		

Note. EF = Executive Function. SES= Socioeconomic Status. $N = 93$.
$^*p < .05$, $^{**}\ p < .01$.

Table 12

Regression analyses with Condition predicting Post-Test Sinking and Floating Overall Learning Composite Scores

Variable	*B (SE)*	*t*	*p*	R^2	ΔR^2
Block 1				.50	-
Age in months	.05 (.01)	5.13	<.001		
Non-Verbal IQ	.03 (.02)	2.16	.03		
Verbal IQ	.05 (.02)	2.38	.02		
SES Composite	.25 (.07)	3.35	.001		
Days between Pre- and Post-Test	.01 (.01)	.96	.34		
Block 2				.58	.08**
Age in Months	.03 (.01)	2.82	.005		
Non-Verbal IQ	.03 (.01)	2.31	.02		
Verbal IQ	.03 (.02)	1.57	.12		
SES Composite	.18 (.07)	2.44	.02		
Days between Pre- and Post-Test	.004 (.01)	.47	.64		
Pre-Test Sinking and Floating Overall Learning Composite	.44 (.11)	3.98	<.001		
Block 3				.66	.08**
Age in Months	.03 (.01)	3.13	.002		
Non-Verbal IQ	.02 (.01)	1.69	.09		
Verbal IQ	.04 (.02)	2.26	.02		
SES Composite	.19 (.06)	2.99	.003		
Days between Pre- and Post-Test	<.00 (.01)	-.01	.99		
Pre-Test Sinking and Floating Overall Learning Composite	.34 (.10)	3.23	.001		
Condition: Control	-.15 (.12)	-1.27	.20		
Condition: Direct	.36 (.11)	3.27	.001		

Note. EF = Executive Function. SES= Socioeconomic Status. $N = 93$.
$^*p < .05$, $^{**} p < .01$.

Table 13

Regression analyses with EF Composite predicting Post-Test Sink and Floating Overall Learning Composite Scores

Variable	*B (SE)*	*t*	*p*	R^2	ΔR^2
Block 1				.59	-
Age in Months	.04 (.01)	4.30	<.001		
Non-Verbal IQ	.01 (.02)	.30	.76		
Verbal IQ	.08 (.03)	2.76	.006		
SES Composite	.28 (.08)	3.54	<.001		
Condition: Direct	.39 (.12)	3.43	.001		
Block 2				.63	.04*
Age in Months	.03 (.01)	2.72	.007		
Non-Verbal IQ	.009 (.02)	.50	.62		
Verbal IQ	.07 (.03)	2.36	.02		
SES Composite	.22 (.08)	2.70	.007		
Condition: Direct	.37 (.11)	3.38	.001		
Pre-Test Sinking and Floating Overall Learning Composite	.28 (.13)	2.24	.03		
Block 3				.63	.00
Age in Months	.03 (.01)	2.18	.03		
Non-Verbal IQ	.007 (.02)	.37	.71		
Verbal IQ	.06 (.03)	2.31	.02		
SES Composite	.21 (.09)	2.41	.02		
Condition: Direct	.37 (.11)	3.31	.001		
Pre-Test Sinking and Floating Overall Learning Composite	.27 (.13)	2.01	.045		
EF Composite	.04 (.12)	.33	.74		

Note. EF = Executive Function. SES= Socioeconomic Status. *N* = 63 (Instructional Groups Only).
$*p < .05$, $** p < .01$.

Table 14

Regression analyses with Sinking and Floating Prior Knowledge predicting Change in Sinking and Floating Overall Learning Composite Scores

Variable	*B (SE)*	*t*	*p*	R^2	ΔR^2
Block 1				.06	-
Age in Months	.001 (.012)	.10	.92		
SB Non-Verbal IQ	.02 (.02)	.97	.33		
SB Verbal IQ	.03 (.03)	.78	.44		
SES Composite	.06 (.10)	.58	.56		
Block 2				.14	.08*
Age in Months	-.003 (.01)	-.23	.82		
SB Non-Verbal IQ	.02 (.02)	.78	.43		
SB Verbal IQ	.03 (.03)	.94	.35		
SES Composite	.06 (.10)	.65	.51		
Condition: Direct	.32 (.14)	2.29	.02		
Block 3				.48	.34**
Age in Months	.03 (.01)	2.72	.007		
SB Non-Verbal IQ	.009 (.02)	.50	.62		
SB Verbal IQ	.07 (.03)	2.36	.02		
SES Composite	.22 (.08)	2.70	.007		
Condition: Direct	.37 (.11)	3.38	.001		
Prior Sinking and Floating Knowledge (Pre-Test)	-.72 (.13)	-5.74	<.001		

Note. SES= Socioeconomic Status. *N* =63 (Instructional Groups Only).
**p* < .05, ** *p* < .01.

Table 15

Regression analyses with BWS x Prior Knowledge Interaction predicting Change in Sinking and Floating Overall Learning Composite Scores

Variable	*B (SE)*	*t*	*p*	R^2	ΔR^2
Block 1				.14	-
Age in Months	-.003 (.01)	-.23	.82		
SB Non-Verbal IQ	.02 (.02)	.78	.43		
SB Verbal IQ	.03 (.03)	.94	.35		
SES Composite	.06 (.10)	.65	.51		
Condition: Direct	.32 (.14)	2.29	.02		
Block 2				.48	.34**
Age in Months	.03 (.01)	2.42	.02		
SB Non-Verbal IQ	.008 (.02)	.45	.65		
SB Verbal IQ	.07 (.03)	2.36	.02		
SES Composite	.22 (.09)	2.53	.01		
Condition: Direct	.38 (.11)	3.35	.001		
BWS	.01 (.09)	.10	.92		
Prior Sinking and Floating Knowledge (Pre-Test)	-.72 (.13)	-5.67	<.001		
Block 3				.51	.03
Age in Months	.03 (.01)	2.31	.02		
SB Non-Verbal IQ	.005 (.02)	.27	.79		
SB Verbal IQ	.06 (.03)	1.96	.05		
SES Composite	.20 (.09)	2.40	.02		
Condition: Direct	.35 (.11)	3.21	.001		
BWS	.05 (.09)	.52	.60		
Prior Sinking and Floating Knowledge (Pre-Test)	-.32 (.28)	-1.14	.25		
BWS x Prior Knowledge	-.23 (.14)	-1.65	.10		

Note. SES= Socioeconomic Status. BWS = Backword Word Span. *N* =63 (Instructional Groups Only).
**p* < .05, ** *p* < .01.

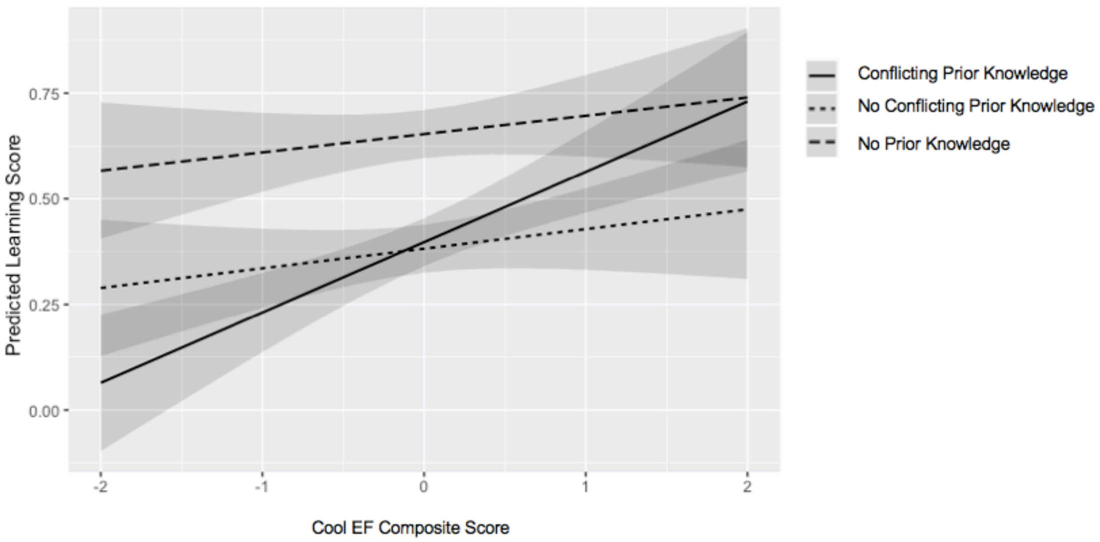


Figure 1. Interaction between Cool EF Composite scores and the amount of conflict with prior knowledge in the learning task, controlling for age in months, verbal IQ, prior knowledge, and Statue task scores. (*N* =54).

Note. EF = Executive function

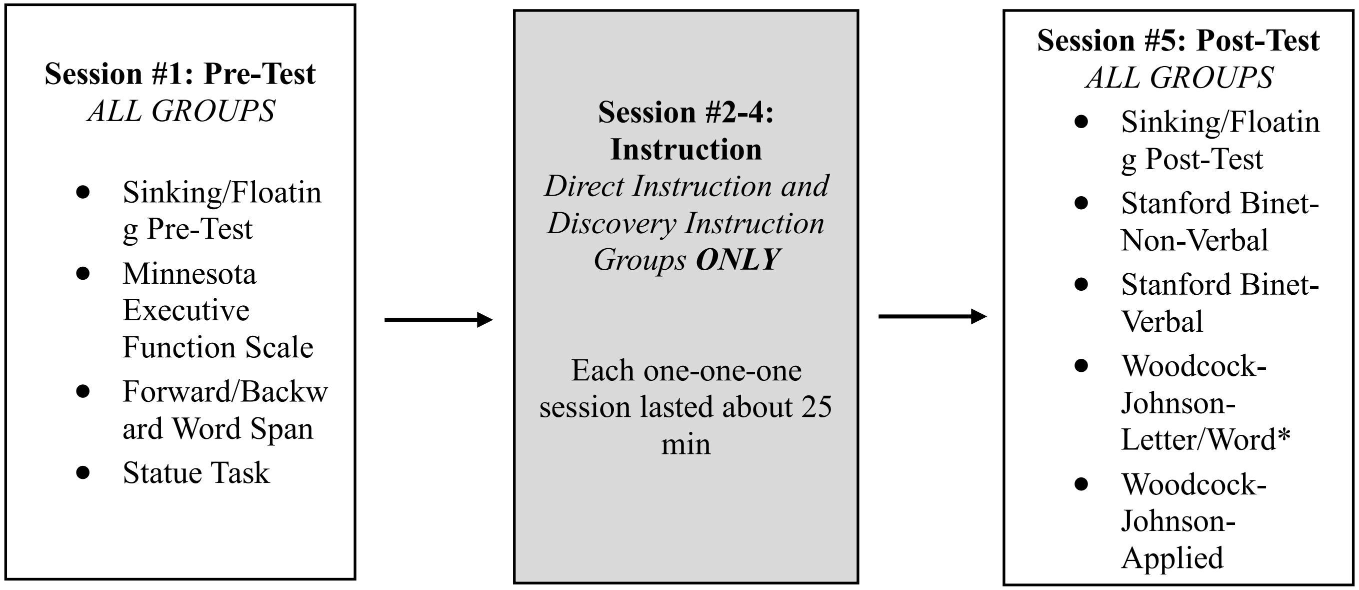


Figure 2. Study Design.

**Note*. Some children in the instructional sessions were administered the Letter/Word and Applied Problems subtests of the Woodcock-Johnson at the end of the second instructional session. Children in the control group were administered the Woodcock-Johnson subtests in between the pre- and post-test sessions, if possible. If the Woodcock-Johnson subtests were not completed before the post-test session, they were completed during the post-test session.

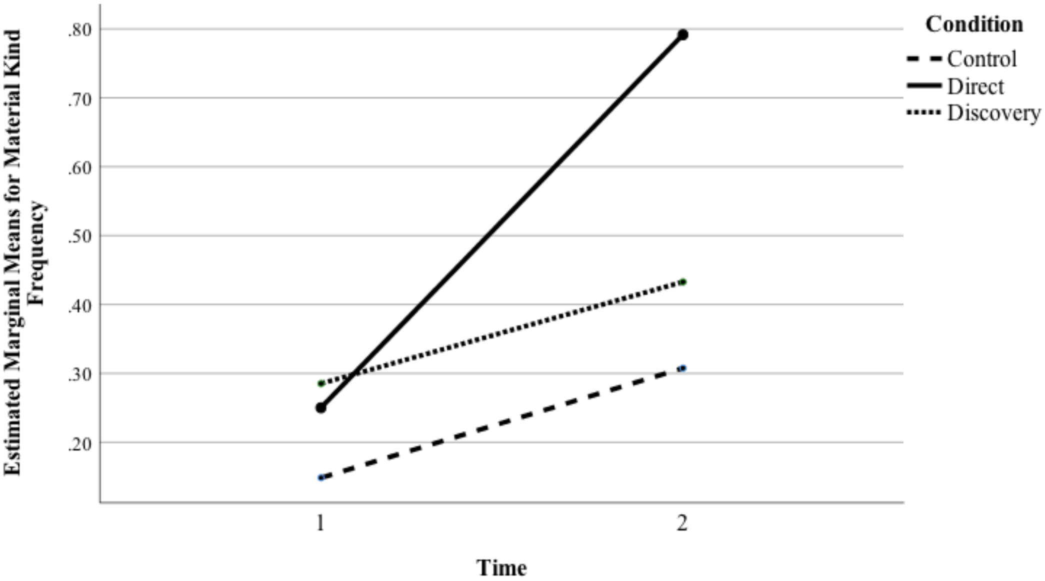


Figure 3. Randomly selected imputed dataset to display the effect of Condition on children’s frequency of mentioning material kind in their reasoning during the sink/float prediction task from Pre-to Post-Test, controlling for Non-verbal IQ, Verbal IQ, Age in Months at Pre-Test, Days Between Pre- and Post-Test Sessions, and SES Composite (*N* = 93).

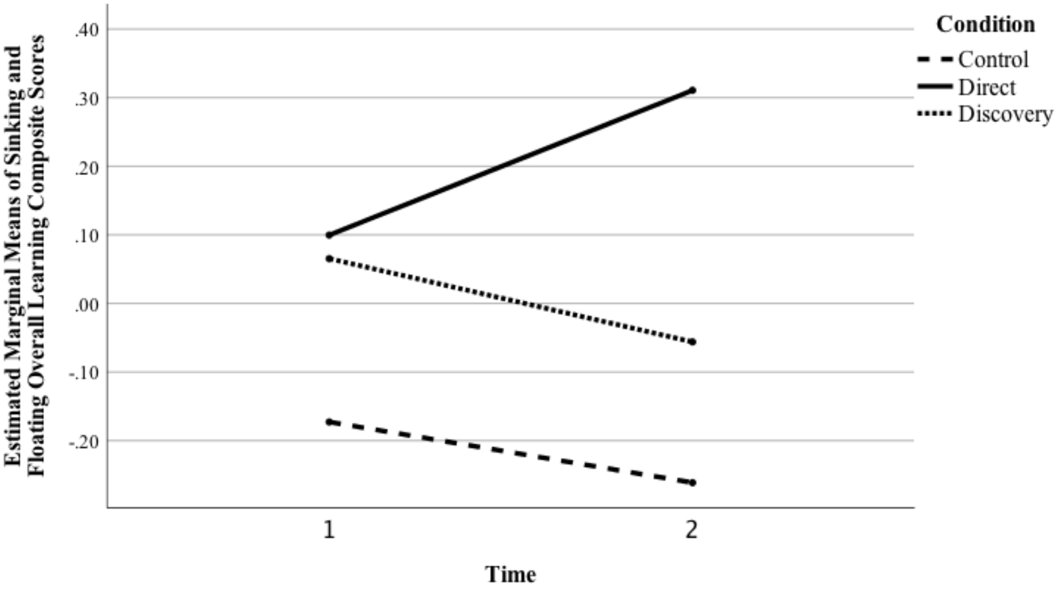


Figure 4. Randomly selected imputed dataset to display the effect of Condition on children's Sinking and Floating Overall Learning Composite Scores from Pre-to Post-Test, controlling for Non-Verbal IQ, Verbal IQ, Age in Months at Pre-Test, Days Between Pre- and Post-Test Sessions, and SES Composite (*N* =93).

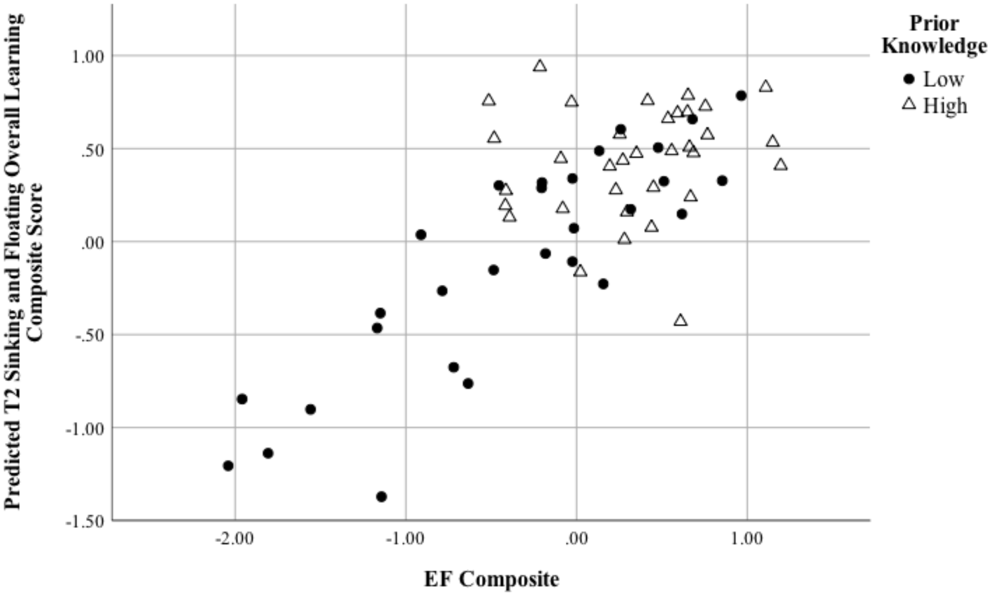


Figure 5. Randomly selected imputed dataset to display the interaction between EF Composite Scores and children's Prior Knowledge of sinking and floating (Pre-Test) to predict their Post-Test Sinking and Floating Overall Learning Composite Scores, controlling for Non-Verbal IQ, Verbal IQ, Age in Months at Pre-Test, SES Composite, and Condition (*N* =63).

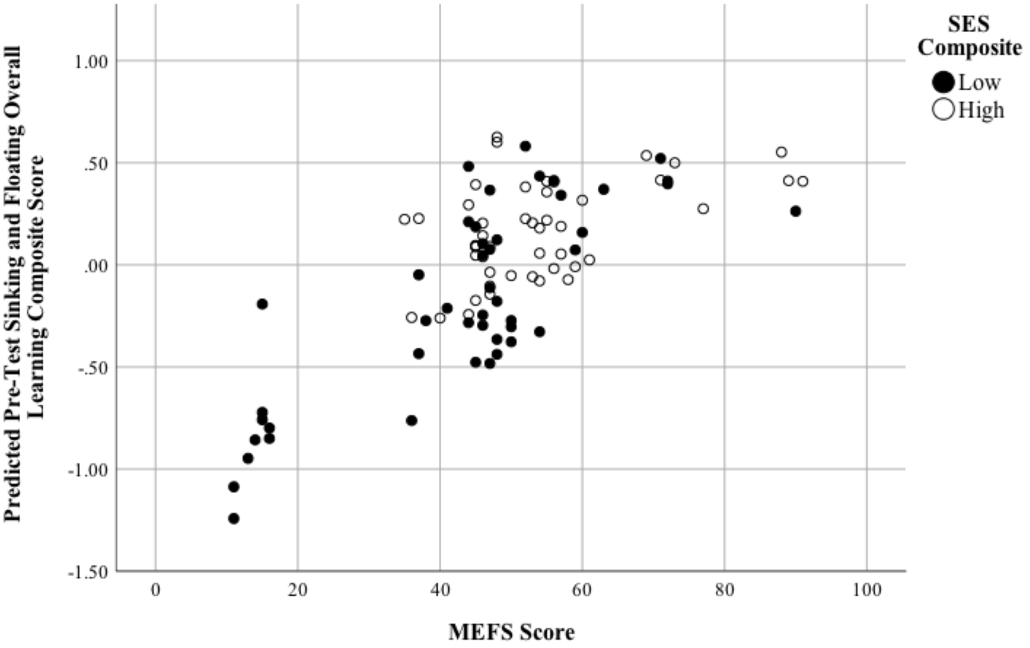


Figure 6. Randomly selected imputed dataset to display the interaction between MEFS and SES Composite Scores to predict children's Pre-Test Sinking and Floating Overall Learning Composite Scores, controlling for Non-Verbal IQ, Verbal IQ, and Pre-Test Age in Months (*N* =93).

Appendix A

Summary of concepts, materials, and example activities for each instructional session

Instructional Session	Topics/Concepts	Materials	Example of Activity
Session # 1	**Size-** objects with same shape and material, but different sizes	• Water tub • Sink/float chart • Five wooden dowel sticks that vary in size (2 in, 4 in, 6 in, 10 in, 12 in) • Five clay dowel sticks that vary in size (2 in, 4 in, 6 in, 10 in, 12 in)	Children are shown five wooden sticks. Children in the **direct instruction condition** are explicitly told that they are all sticks and made of wood, but they are different sizes. Children in the **discovery instruction condition** are asked what is the same and what is different about the sticks. Then they see if they float or sink and sort them on a sink/float chart. This is repeated with the set of clay sticks. Once all the sticks have been sorted, the instructor either explicitly says that the material of the sticks is important for whether they sink or float since big and large sticks in each material behaved the same (**direct instruction condition**) or asks the child why he or she thinks all of the sticks on the sink side sank and what is the same or different about them and why all of the sticks on the float side floated and what is the same or different about them (**discovery instruction condition**)
	Weight - objects with same shape and material but different weights	• Water tub • Sink/float chart • Balance scale with bear weights • 2 plastic balls (one heavier than the other)	Children are shown a balance scale and the instructor demonstrates how to weigh the metal balls one at a time using the bear counters. The instructor then puts the two metal balls in the water, says they sink, and then puts them on the sink side of the chart. The demonstration for children in the **direct instruction condition** involves more of a discussion about what weight means and the instructor narrates everything she is doing. She

		- 2 wooden balls (one heavier than the other) - 2 metal balls (one heavier than the others) - 2 clay balls (one heavier than the other)	also tells children that both of the metal balls sink, even though one was heavier than the other, because they are made of metal. In the **discovery instruction condition**, the experimenter does not discuss weight with the child and uses minimal verbalizations during the demonstration. Once the instructor has demonstrated that both balls sink, she asks the child why they think they both sank. Next, the child is given one pair of balls (wooden, clay, and plastic) at a time to weigh, place in the water, and then sort on the chart. Children in the **direct instruction condition** are told that the objects both sank or floated even though one was heavier than the other after they put the two balls in the water. Children in the **discovery instruction condition** are only asked why they think they both sank or floated. After all the objects have been sorted, children in the **direct instruction condition** are told that what the objects are made of is important to determine if an object will sink or float whereas children in the **discovery instruction condition** are asked what is the same and different about the different objects and why they think all of the objects on the sink side sank and all the objects on the float side floated.
	Material Kind- density cubes that are the same size and shape	- Water tub - Sink/float chart - 4 small density cubes (wood,	Children are introduced to four small density cubes and are either told how they are the same or different (**direct instruction**) or asked how they are the same or different (**discovery instruction**). The experimenter demonstrates how one cube floats and one cube that sinks. Then children are told

	but different materials	plastic, clay, and metal) - 4 big density cubes (wood, plastic, clay, and metal)	it is their turn to test out the rest of the objects and put them on the chart where they go. Then the instructor brings out four big density cubes and the child is either told how they are the same or different (**direct instruction**) or asked how they are the same or different (**discovery instruction**). Then they are told it is their turn to test out the objects and put them on the chart where they go. After all the objects have been sorted on the chart, children are explicitly told how the objects are similar or different and that the material that objects are made of is important for them to sink or float (**direct instruction**) or asked how they are the same and different and why they think all the objects that sink sank and all the objects that float floated (**discovery instruction)**
Session # 2	Review lessons learned in first session by looking at objects that vary in size and material kind but are the same shape	- Water tub - Sink/float chart - 4 spoons of different sizes in each of the four target materials (clay, wood, metal, and plastic) - 4 rings of different sizes in each of the four target materials (clay,	Children are introduced to the four spoons and are either told how they are the same or different (**direct instruction**) or asked how they are the same or different (**discovery instruction**). The instruction demonstrates how one spoon sinks and one spoon floats. Then they are told it is their turn to test out the rest of the objects and put them on the chart where they go. Then the instructor brings out four rings and the child is either told how they are the same or different (**direct instruction**) or asked how they are the same or different (**discovery instruction**). Then they are told it is their turn to test out the objects and put them on the chart where they go. After all the objects have been sorted on the chart, children are explicitly told how the objects are similar or different and that

		wood, metal, and plastic)	the material that objects are made of is important for them to sink or float (**direct instruction**) or asked how they are the same and different and why they think all the objects that sink sank and all the objects that float floated (**discovery instruction)**
	Material Kind Application Activity	• Water tub • A plastic white string • Various objects with holes in them made out of the target materials (wood, plastic, metal, and plastic) and some non-target materials (Styrofoam, sponges)	Children are introduced to the materials and then told they can make chains that sink and chains that float. The instructor than demonstrates making a chain that floats. In the **direct instruction condition**, children are told they need to put things that float on the chain to make it float and the instructor labels the materials of the objects she is putting on her chain. In the **discovery instruction condition,** the instructor says she will make a chain that floats, puts on an object that will float, and then puts the chain in the water and says it floats. The instructor then demonstrates how to make a chain that sinks in the same way as described above for each condition. The child is then told it is their turn to make chains. The instructor encourages them to make at least one chain that floats and one that sinks during the activity. After the child has made and tested a chain in the water, the instructor either explicitly labels the materials the child used and whether they sink or float (d**irect instruction condition**) or the instructor asks the child why they think their chain sank or floated **(discovery instruction condition)**.
Session # 3	How to make something that floats sink	• Water tub • Plastic cups • Various objects in	Children are shown that the plastic cup floats and then shown one way to make the plastic cup sink (e.g., putting objects that sink inside the cup).

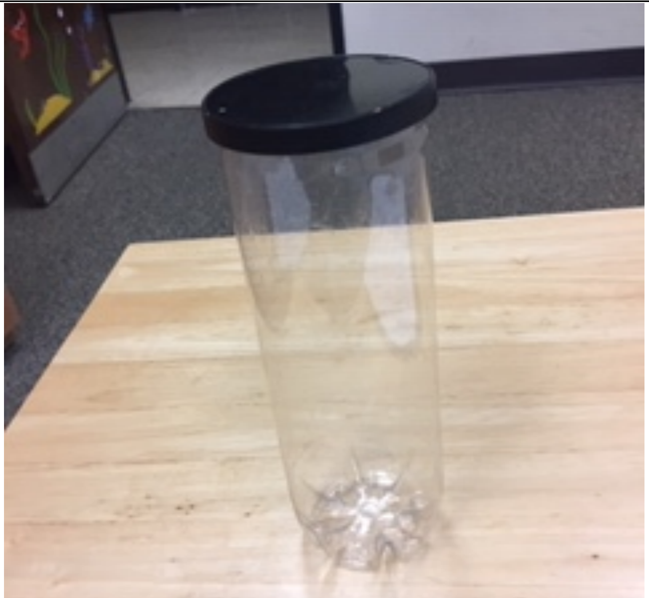
		different sizes and made of target and non-target materials that can fit inside the plastic cups	Then children are given a turn to make the cup sink. In the **direct instruction condition**, the instructor explicitly explains the different ways of making the cup sink that the child tries and why those worked. In the **discovery instruction condition**, children are asked why and how they made the cup sink.
	How to make something that sinks float	• Water tub • Ball of clay	Children are shown a ball of clay and the instructor demonstrates that it sinks. The instructor then changes the ball into a log shape and puts it in the water. It sinks again. The instructor then changes the log into a flat pancake shape. It sinks again. The instructor then changes the pancake into a boat shape. The clay floats. Children in the **direct instruction** condition are told one way to make something that sinks float is by changing the shape and are introduced to the idea of an air cavity being important for things to be able to float like boats. Children in the **discovery instruction condition**, are asked why they think the clay boat floats now when the other objects sink and to think about how the different objects (ball, log, pancake, and boat) are the same and different

			Children are then given a clay ball and given a chance to try to make it float.
	How to make something that sinks float	• Water tub • Metal toy boat • 6 x 6 squares of tinfoil • Pennies	Children are shown that a metal toy boat floats. Children in the **direct instruction condition** are then told about how boats float (e.g., water displacement and buoyancy) and how the amount of water that a boat displaces relates to how much they are able to carry without sinking. Children in the **discovery instruction condition** are asked how they think boats can float and how they can carry heavy loads without sinking. Then children are introduced to the foil sheets and pennies and asked to make a boat out of the foil and then they can see how many pennies it can carry before sinking. Children are encouraged to make at least two different boats. After children have made multiple boats, children in the **direct instruction condition** are told that the size, shape, and sturdiness of the boat are important for how many pennies it can carry and the experimenter explicitly tells them how their boats were different and why they could hold different amounts of pennies. Children in the **discovery instruction condition** are asked why they think their boats could hold different amounts of pennies, what was the same and different about their boats, and finally what they think boats need to have to be able to carry lots of pennies without sinking.

Appendix B

Pre-Test Prediction Objects

Object	**Picture**	**Material**	**Sink/Float**
Dowel		Wood	Float
Gem		Clay	Sink
Large Density Cube (1.5 in)		Plastic	Float

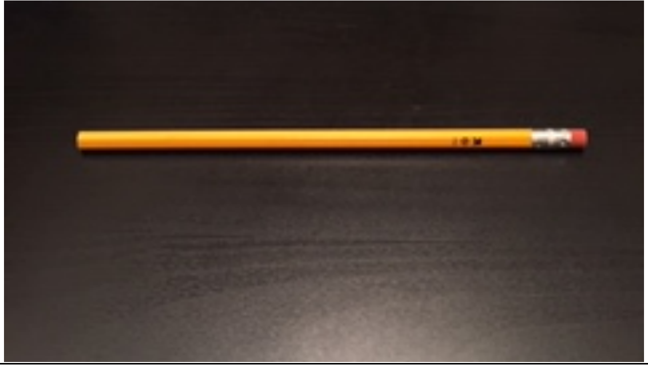
Large serving spoon		Metal	Sink
Pot		Clay	Sink
Washer		Metal	Sink
Empty tennis ball canister		Plastic	Float

Block		Wood	Float
Ball		Clay	Sink
Drawer knob		Metal	Sink
Ring		Plastic	Float
Spool		Wood	Float

Candle		Wax	Float
Eraser		Rubber	Sink

Appendix C

List of Post-Test Prediction Objects

Object	Picture	Material	Sink/Float
Puzzle Piece		Wood	Float
Large block		Wood	Float
Pencil		Wood	Float

Straw		Metal	Sink
Carabiner Clip		Metal	Sink
Paper clip		Metal	Sink
Plate		Clay	Sink

Spoon		Clay	Sink
Heart		Clay	Sink
ID Holder		Plastic	Float
Box		Plastic	Float

Ping-pong ball		Plastic	Float
Bottle Stopper		*Plastic and rubber*	Float
Rubber band		*Rubber*	Sink

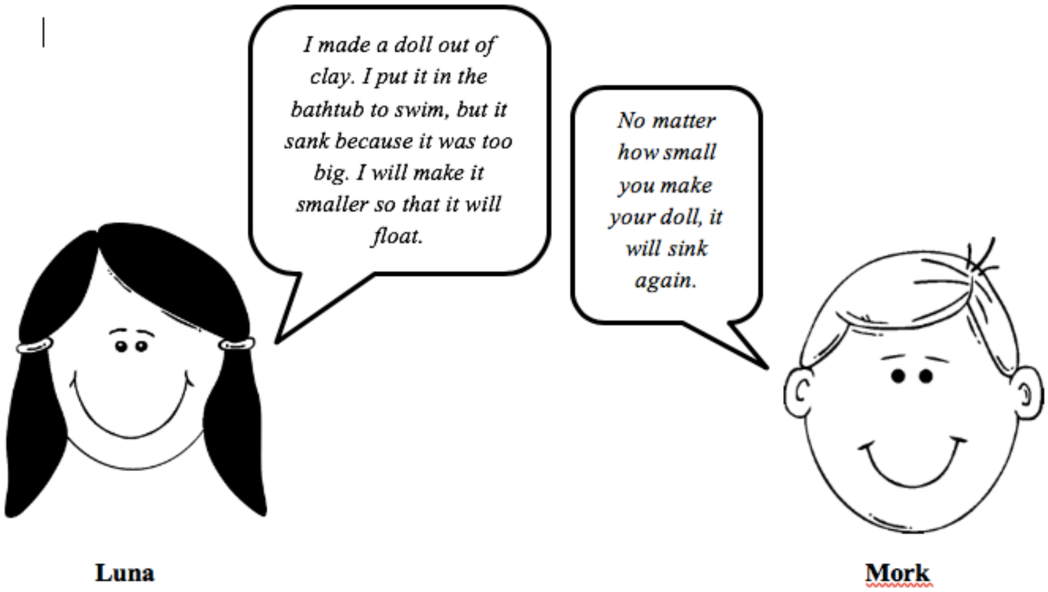
Appendix D

Coding Scheme for Explanation Score for the Sinking and Floating Prediction Task

Explanation	**Explanation Points**
Explanation is related to material kind correctly **OR** in relation to weight of water-density type explanation **OR** hollow bodies **OR** incorrectly identifies the material but has a correct explanation (e.g. saying the plastic container sinks because it's metal)	5
Explanation is related to material kind incorrectly (e.g., wood sinks)	4
Explanation is not material kind or focuses on one or more of the following factors (e.g., weight, size, consistency, shape) correctly (e.g., paper clip is light) **OR** incorrectly (e.g., a paper clip is heavy)	3
Explanation is prior experience alone or object itself	2
Explanation is irrelevant to sinking/floating (e.g., it moves like a domino)	1
No explanation given (e.g., I don't know, because, etc.)	0

Appendix E

Example of a Concept Cartoon used in the Sinking and Floating Assessment (From Kallery, 2015)



Experimenter: “This is Luna (*point to Luna*). Luna says, “I made a doll out of clay. I put it in the bathtub to swim, but it sank because it was too big. I will make it smaller so that it will float. Then Mork (*point to Mork*) says to Luna, “No matter how small you make your doll, it will sink again.”

Experimenter: “What do you think? Who do you think is right, Luna or Mork?”

Experimenter:” Why?”

Experimenter: “So Luna has a clay doll that sinks but she thinks she can make it float if she makes it smaller. Do you think that will work? Why or why not?”