EFFECTS OF FANTASY-ORIENTED PLAY

ON THE DEVELOPMENT

OF EXECUTIVE FUNCTIONS

by

RACHEL BETH THIBODEAU

ANSLEY T. GILPIN, COMMITTEE CHAIR

FRANCES A. CONNERS
ANGELA B. BARBER

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts
in the Department of Psychology
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2015

ABSTRACT

Although recent correlational studies have found a relationship between fantasy orientation (FO; i.e., a child's propensity to play in a fantastical realm) and higher-order cognitive skills called executive functions (EFs), no work has addressed the causality and directionality of this relationship. The present study experimentally determined the directionality of the observed relationship between FO and EFs through the training of FO in preschool-aged children. One hundred thirteen children between the ages of 2 and 5 were randomly assigned to one of three conditions: fantasy-play, realistic-play, or business-as-usual control. Results revealed that children who participated in a five-week fantasy-play intervention showed consistent improvements in inhibitory control and working memory beyond that of children who participated in realistic-play or business-as-usual control conditions. The data suggest that the relationship between FO play and EF development may be equifinal such that engaging in FO play is one of many ways to directly enhance EF development.

DEDICATION

This thesis is dedicated to my niece, Caroline Grace Dooley, whose ongoing development drives not only my research, but also my desire to learn more about children's cognitive development.

LIST OF ABBREVIATIONS AND SYMBOLS

a	Cronbach's index of internal consistency
β	Beta: standard regression coefficient
F	Fisher's *F* ratio: A ration of two variances
M	Mean: the sum of a set of measurements divided by the number of measurements in the set
N	Sample Size
η^2	Eta squared: measure of effect size
p	Probability associated with the occurrence under the null hypothesis of a value as extreme as or more extreme than the observed value
r	Pearson product-moment correlation
SD	Standard deviation: value of variation from the mean
<	Less than
>	Greater than
=	Equal to

ACKNOWLEDGMENTS

I am pleased to have this opportunity to thank the individuals who have helped me with this research project. First and foremost, I have to thank my advisor Ansley Gilpin for her constant encouragement, support, and guidance in completing this thesis. This thesis also would not have been possible without the help of members of the KID Lab including Melissa Brown, Allie Nancarrow, Brooke Meyer, and an army of undergraduate research assistants. I thank you all for the countless hours you have spent at preschools and in the lab to help me with data collection. Additionally, I would like to thank the members of my committee, Fran Conners and Angie Barber, for their valuable input and suggestions throughout. Finally, I would not be where I am today without the love and support from my family and friends. Your encouragement from miles away has been unwavering! Thank you for always pushing me to do my best and reminding me to be thankful in all circumstances.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Due to recent declines in academic performance throughout the United States, there has been a surge of research in the last 10 years aimed at understanding the development of cognitive control. This accumulation of research demonstrates the importance of cognitive control not only to everyday functioning, but also to academic achievement, job performance, and overall well-being (Dunn, 2010; Eakin et al., 2004; Moffitt et al., 2011; St. Clair-Thompson & Gathercole, 2006). One of the main concepts that has emerged from the body of research on cognitive control is a category of cognitive abilities referred to as executive functions (EFs).

Executive Functions

Executive functions are defined as higher-order thinking processes that allow individuals to override more automatic thoughts and behaviors for more adaptive and goal-directed responses (Carlson, 2005). Some of the most important cognitive processes included under the umbrella term of EFs are working memory, inhibitory control, and cognitive flexibility (Miyake et al., 2000; St. Clair-Thompson & Gathercole, 2006). Interestingly, research indicates that although these cognitive processes comprise distinct components of EF, they also share common underlying features and are highly related (Miyake & Freidman, 2012; Miyake et al., 2000). That is, the components that make up EFs display both unity and diversity (Miyake & Freidman, 2012; Miyake et al., 2000). Although this pattern is seen in adolescence and adulthood, in early childhood these domains of EF are highly integrated and not easily teased apart (Carlson, Moses, & Breton, 2002; Diamond & Taylor, 1996; Hughes, Ensor, Wilson, & Graham, 2010; Wiebe, Epsy, & Charak, 2008; Wiebe et al., 2011). Specifically, Wiebe and colleagues (2008) utilized confirmatory factor analysis to compare various models of executive processing in 2-3 year olds in order to determine if the components that comprise executive control (i.e., working memory and inhibition) are more domain-specific or domain-general in young individuals. These authors found that a single-factor model adequately explained the data. Although tasks of working memory and inhibitory control have distinct characteristics on a surface level, Wiebe et al. (2008) conclude that tasks of working memory and inhibitory control seem to measure a single cognitive ability in preschool aged children.

Expanding upon their previous research, Wiebe and colleagues (2011) performed another study using confirmatory factor analysis in a sample comprised of only young preschoolers (i.e., 3-year-olds). It is crucial to explore the structure of EFs in this age range, as 3 years of age is pivotal for the development of EFs (Diamond & Taylor, 1996). Prior to their study, analyzing the domains of executive control in young children proved difficult due to the lack of EF tasks suitable for this age group. However, with the development of more age appropriate measures of EF, Wiebe and colleagues (2011) were able to examine the relationships among several measures of working memory, inhibitory control, and cognitive flexibility. Similar to the results above, Wiebe et al. (2011) found that a single latent EF factor best modeled the data in this young population, with other research supporting this conclusion (Hughes et al., 2010).

Although EFs are measured in research with tasks that both measure it globally as well as tasks that measure a single aspect (e.g., working memory), it is very rare for a child to be presented with a situation that only requires the use of one EF in non-research settings. For example, children are often asked by parents and teachers to do various tasks at one time such as pick up your toys, put them in your room, and wash your hands before supper. Doing so requires that a child actively engages not only their working memory system (to remember the three instructions they were given), but also their attentional and inhibitory systems (to stay focused on the tasks at hand rather than sitting down to play with another toy). These domains are highly integrated in real life and not easily teased apart.

These tasks are often consolidated in real life, just like the skill sets are highly integrated in early childhood. In fact, the ability to perform tasks that require multiple EFs (e.g., the ability to hold two rules in mind while also inhibiting a prepotent response) improves between 3-6 years of age (Diamond & Taylor, 1996). The "Day-Night" task is a great example of this phenomenon. In this task, children are instructed to say the word "day" when shown a picture of the moon and to say the word "night" when shown a picture of the sun (Gerstadt, Hong, & Diamond, 1994). Not only does this task require the inhibition of a prepotent response to name what they see, but it also requires the children hold the instructions and rules in their working memory. Interestingly, 3 ½ to 4 ½ year olds experience difficulty in completing this task whereas 6 to 7 year olds are able to complete this task without much difficulty. However, when the task is manipulated so that it only taxes one component of EF, performance in younger children increases (Diamond, Kirkham, & Amso, 2002). Specifically, Diamond and colleagues (2002) instructed children to say "dog" when shown the picture of the moon and "pig" when shown a picture of the sun. Thus, children only had to use their working memory to remember the rules and no longer had to inhibit a prepotent response as the words they were instructed to say were unrelated to the pictures presented. In doing so, these children performed significantly better on this task than when it was presented in the standard format.

Despite these results, testing just one domain of EF may not lend externally valid results given that the domains of EF are highly integrated and that children are often presented with situations that require multiple EFs in daily life. Although each EF measure may not be individually predictive of a child's EF abilities, a large battery that tests multiple domains of EFs can provide a more accurate and externally valid depiction of a child's cognitive abilities. Below is a review of the components of EFs.

Working Memory

Working memory is defined as the temporary storage of information, which allows individuals to manipulate information as they cognitively process it (Baddeley, 1983; Baddeley, 1992). Two measures that are consistently used to assess working memory include the Forward Digit Span and Backward Digit Span (Halford, Maybery, & Bain, 1988). As their names imply, the Forward Digit Span requires that participants repeat strings of digits back to the experimenter just as they heard them, whereas the Backward Digit Span requires that participants repeat the strings of digits backwards. Preschoolers experience more difficulty with the Backward Digit Span because it requires that the participants first recall the digits spoken, and then manipulate this information in order to repeat these digits backwards (Davis & Pratt, 1996).

Inhibitory Control

Inhibitory control refers to one's ability to suppress an automatic, prepotent response, and to initiate more appropriate responses (Stroop, 1935; Wright, Waterman, Prescott, & Murdoch-Eaton, 2003). Prepotent responses are those that have been learned by an individual through reinforcement and thus are automatically retrieved. For example, after years of practice, adults develop a prepotent response to read words on a page and ignore visual information surrounding the words. This phenomenon can be observed in a classic Stroop task whereby individuals are given color words (e.g., blue, red, green) and asked to name the color of the ink in which the word is printed. In other words, individuals must suppress their prepotent response to read the word in order to attend to the surrounding information. Research shows that adults have difficulty in naming the ink color and often read the word itself, thus indicating that reading a word is a prepotent response.

An influential study that sheds light on the developmental trajectory of inhibitory control and the components that make up this construct continues to impact and direct the literature in this area of research. Specifically, Diamond and Taylor (1996) studied inhibitory control abilities of 160 children ranging from 3.5 years of age to 7 years of age using the “Day-Night” task, described earlier. In this task, the child must inhibit the prepotent response to name what they see (e.g., “day” when shown a picture of the sun), and subsequently respond with a word that describes the opposite of the picture shown. In their study, Diamond and Taylor (1996) found that performance on measures of inhibitory control increase in children between the ages of 3 ½ and 6.

Taking into consideration the development of inhibitory control in young children, the Stroop task mentioned above (Stroop, 1935) is commonly used as a measure of cognitive inhibitory control in adults. Given that literate adults have a prepotent response to read the words printed on a page, the Stroop task requires that individuals inhibit the automatic response to read and alternatively name the color of the ink in which the word is printed. Although this task is a reliable and valid measure of cognitive inhibitory control in adults, it is not developmentally appropriate for young children who cannot read. That is, because children have not yet developed a prepotent response to reading the words over naming the color of the ink, this task is not a valid measure of cognitive inhibition for this age group. Interestingly, children often perform better than adults on this classic Stroop task because they have not yet formed a prepotent response for reading a word over attending to surrounding visual information.

Therefore, researchers have since developed various developmentally appropriate tasks to mirror the Stroop task (e.g., "Day-Night" task). However, many of these measures, including the "Day-Night" task, tax working memory in addition to inhibitory control making it difficult to determine if errors in performance are due to deficits in working memory or cognitive inhibitory control. Despite this issue, one task developed by Wright et al. (2003) requires that participants inhibit a prepotent response but does not tax the working memory system as much or demand reading skills. In this "Animal Stroop" task, children are shown drawings of animals, some of which have mismatched heads and bodies (e.g., a cow head on a pig body). Because young children develop a prepotent response for identifying entities by their faces early in life (Johnson, 1993), this task asks children to name the animal by its body and thus requires that children inhibit their automatic response to name an entity by its face. In their study of 155 children between the ages of 3 and 6, Wright et al. (2003) found that, similar to the results of a classic Stroop task, participants made more errors and had slower response times for Stroop trials (i.e., mismatched images) than control trials (i.e., matched images).

Attentional Shifting

Finally, cognitive flexibility, also referred to as attentional shift, refers to an individual's ability to shift their attention back and forth between two different domains (Monsell, 1996). This domain of EF is commonly measured using the Standard Dimensional Change Card Sort task (Frye, Zelazo, & Palfai, 1995; Zelazo, Muller, & Frye, 2003). The purpose of this task is to examine how well participants are able to perform after they are forced to switch to a new game with a new set of rules. Specifically, participants are shown two boxes, each with a different target image (i.e., a blue square and a red star). During the course of the task, participants will play two games: the shape game and the color game. The experiment begins by explaining the rules of the first game (e.g., the shape game). After the participant has correctly sorted five cards based on the rules of the current game, the experimenter stops the participant, tells them they are going to play a new game (e.g., the color game), and proceeds to explain the rules of this new game. The participant then proceeds according to the rules of the new game until they correctly sort five cards in a row.

Executive Function Development

A growing body of evidence indicates the importance of EF development throughout childhood and adolescence. Executive functions, which begin to develop shortly after birth, lay the foundation for cognitive and social development (Moffitt et al., 2011). Interestingly, previous research shows that EFs correlate with math, language, and literacy development and are more important for school readiness than IQ (Blair & Razza, 2007). Even as individuals progress into adulthood, EFs play a major role in physical and mental health, social relationships, and employment (Moffitt et al., 2011). For example, children who display low self-control and high impulsivity tend to have poorer health, earn less money in their respective careers, and commit more crimes as adults than children with high levels of self-control (Moffitt et al., 2011). Unfortunately, research indicates that EF abilities do not automatically develop and mature over the lifespan, but rather require rich environmental experiences (Center on the Developing Child at Harvard University, 2011). In fact, children raised in adverse environments (e.g., environments characterized by abuse and neglect) exhibit serious deficits in cognitive, attentional, and behavioral control suggesting that EFs are at risk for disruption at an early age (Center on the Developing Child at Harvard University, 2011). Therefore, it is imperative to determine methods that facilitate the development of EFs early in life in order to ensure that every child and adolescent has an equal opportunity to become a successful and contributing member of society as an adult.

Several programs attempting to improve children's EFs have since been developed. These programs utilize activities ranging from computer-based training, games, and school curriculum to physical activities such as yoga, martial arts, and aerobics (Diamond & Lee, 2011). Most of the studies investigating the effects of these programs on EFs have been conducted with school-aged children (i.e., age 5-13; Diamond & Lee, 2011). However, several of the computer-based training programs utilizing CogMed ©, a software that aims to increase working-memory demands, have been empirically studied in preschool aged children (Bergman Nutley et al., 2011; Lillard et al., 2013; Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2009). Although 4-year-olds who underwent five weeks of working memory training using the CogMed program displayed improvements in working memory beyond preschoolers in control groups, these effects did not necessarily generalize to other EF skills, such as problem solving (Bergman Nutley et al., 2011; Thorell et al., 2009). Computer programs aimed at specifically improving inhibition produced similar results. Specifically, 4- to 5-year-old children who participated in five weeks (15 min/day) of inhibition games showed improvements on two of the three inhibition games practiced when compared to controls (Thorell, et al., 2009). However, these improvements did not transfer to non-practiced inhibition tasks (Thorell et al., 2009).

Due to the fact that results of computer-based EF training programs have poor transfer effects between EF skills and that these programs are not easily accessible to all children (Diamond & Lee, 2011), researchers have recently turned their attention to finding natural environmental experiences that are involved in the development of EFs. For example, there is evidence that bilingualism, which requires cognitive flexibility to switch between languages, is related to the development of EF abilities in children (Bialystok, 1999; Bialystok, 2011; Bialystok & Martin, 2004). Specifically, when compared to monolingual children, bilingual children between the ages of 4 and 5 years are more successful at shifting between dimensions on a dimensional-change card-sort task (Bialystok, 1999; Bialystok & Martin, 2004). In addition, bilingual individuals also out-perform their monolingual counterparts on measures of cognitive inhibition (i.e., Stroop task; Bialystok, Craik, & Luk, 2008). Although bilingualism is a great example of a natural environmental experience that seems to facilitate the development of EFs, it is not practical to manipulate a construct like bilingualism for the purpose of experimentation. However, another example of a natural environmental experience that is easily manipulated and is ubiquitous in childhood is pretend play.

Play as a Mechanism of Executive Function Development

As Vygotsky theorized, complex pretend play may also provide a natural environmental experience in which cognitive skills can be developed (Vygotsky, 1978). Specifically, Vygotsky reasoned that imaginative play is instrumental to the development of children's ability to think about objects and events that are not immediately present (i.e., internal systems of representation; Vygotsky, 1967). Interestingly, play changes throughout development and follows a consistent timeline across children (Carlson & Zelazo, 2008). For example, early in life (i.e., 2 years old) children are able to give living characteristics to nonliving objects like stuffed animals (Carlson, White, Davis-Unger, 2014). At this age, children are also able to transform objects and use them for purposes other than their intended use (e.g., pretending a banana is a phone; Carlson et al., 2014). Eventually, as children enter the preschool years, their play becomes more abstract and less dependent on actual objects or props. For example, children at this age will pretend to bounce an imaginary ball or pretend to cook and eat an imaginary meal. Because they are able to utilize abstract concepts in their play, children at this age often enrich their play with fantastical themes.

Fantasy Orientation

The term fantasy orientation (FO) refers to a child's propensity to play in a fantastical realm and is often operationalized in children as the extent to which they engage in imaginary play and whether or not they have imaginary companions (Taylor, 1999). Research demonstrates the by three years of age, children are able to understand the concept of fantasy, differentiating fantasy from reality (Woolley & Wellman, 1990). Therefore, studies of FO can be conducted in children as young as 3 years old.

Interestingly, there are observable individual differences in children's play with respect to a child's level of FO. For example, children who are low in FO tend to supplement their play with real items such as using blocks to build towers. By contrast, children who are high in FO often engage in play that involves imaginary companions and fantastical story lines. In examining a child's FO, researchers often use questions from Singer and Singer's (1981) Imaginative Play Predisposition interview and Taylor and Carlson's (1997) Imaginary Companion and Impersonation interview. These measures ask children about their pretending behavior (e.g., do you ever pretend to be an animal or a machine), their imaginations, and their imaginary companion(s). In a study of 152 preschoolers, Taylor and Carlson (1997) found that one-third of their sample could be classified as high FO.

Relationship between Fantasy Orientation and Executive Function

Although there are clear individual differences in children's level of FO, to date few studies have investigated the impact of a child's level of FO on their cognitive control. Similar to the cognitive flexibility needed for bilingualism, the act of engaging in imaginary play and having imaginary companions requires that children switch back and forth between fantasy and reality (Estes, Wellman, & Woolley, 1989; Golumb & Kuersten, 1996) and thus use working memory to remember pretense rules and scripts, inhibit using pretense scripts in real life, and shift attention back and forth between reality and pretense. Therefore, it seems logical that FO would be related in some fashion to EF development. Sound empirical evidence of a relationship between FO and EF has only recently emerged. Specifically, Pierucci and colleagues (2013) interviewed preschoolers between the ages of 4 and 6 using measures of EF and FO. Pierucci et al. (2013) found that children exhibiting high FO (i.e., engaging in FO cognitions, belief in FO entities) displayed better cognitive inhibition and attentional shift than children exhibiting low and moderate levels of FO. More recently, Carlson and colleagues (2014) found positive correlations between EFs and the understanding of pretend versus reality ($r = .53$, $p < .01$) as well as the ability to perform pretend actions ($r = .21$, $p < .05$) among preschool children. Furthermore, when subjected to a regression analysis, Carlson and colleagues (2014) found that the understanding of pretend versus reality significantly predicted EF scores ($\beta = .22$, $p < .01$) and the ability to perform pretend actions marginally predicted EF scores ($\beta = .14$, $p < .10$).

Purpose of the Current Study

Despite the observed relationship between FO and EF in preschoolers, the correlational nature of these findings leave open questions regarding the directionality of these relationships (Carlson et al., 2014; Pierucci et al., 2013). One possibility is that having more advanced EFs provides individuals with the capacities needed to engage in FO behaviors. However, this explanation is less likely given that FO appears to be more stable than EF, with research indicating that EF develops as individuals progress through childhood and adolescence (Pierucci et al., 2013). Another possibility is that the act of engaging in FO behaviors facilitates the development of EFs. As mentioned above, FO activities require that children remember pretense rules and scripts, inhibit using pretense scripts in real life, and shift attention back and forth between reality and pretense. Therefore, the nature of engaging in FO activities may provide FO children more opportunities to practice EFs. To date, little has been done to empirically investigate the causality of this relationship between a child's FO and EF abilities (Carlson et al., 2014).

In fact, a recent meta-analysis in the field has called for experimental investigations of the effects of pretend play on cognitive control (Lillard et al., 2013), as many previous studies that have explored EFs in the context of pretend play have several methodological issues and are therefore inconclusive. For example, a Vygotskian classroom intervention known as Tools of the Mind has been implemented as an intervention to facilitate EF development (Bodrova & Leong, 2007; Lillard et al., 2013). This program has three main components that are performed on a daily basis: self-regulation, reading, and pretend play. The results across several empirical investigations of the Tools intervention are mixed. In those that do find improvements in EFs, researchers were unable to separate the effects of pretend play from the other aspects of the program (i.e., self-regulation and reading) to know if pretend play caused the observed increases in EF performance.

In addition, all previous experimental studies on play and EFs have specifically investigated pretense (Lillard et al., 2013). Pretense is a type of imaginative play in which children simply pretend to be an animal or another person (Woolley & Tullos, 2008). It does not necessarily involve the same fantastical elements as FO activities and cognitions. For example, pretense could involve pretending to drive a car whereas FO play would involve flying a car in outer space. Interestingly, Pierucci and colleagues (2013) found that unlike FO activities and cognitions, pretense did not correlate with any measures of EF. Thus it appears that fantasy may be crucial element to the observed relationship between play and EFs.

Therefore, the aim of the present study was to experimentally determine the causality and directionality of the observed relationship between FO and EFs through the training of FO in preschool aged children. In particular, the current study investigated whether or not the repeated act of engaging in FO behaviors actually facilitates the development of EFs beyond what would be observed due to simple maturation or engaging in non-fantasy play. One hundred thirteen preschoolers between the ages of 2-5 years were pre-tested on several measures of FO and EF (i.e., working memory, inhibitory control, and attention shift). Participants were then randomly assigned to one of three conditions: fantasy-play, realistic-play, or control. In the fantasy-play condition, groups of children worked with a research assistant to come up with a fantastical script and were encouraged to act it out. A realistic-play condition was included in which groups of children played games that did not require pretending. This condition helped to determine whether a pretending component of play is necessary for changes in EFs rather than just the act of engaging in play itself. Finally, a business-as-usual control condition did not receive any intervention. After five weeks of intervention, participants were tested using the matched pre-test measures of EF and FO. Post-test EF and FO scores were compared across groups.

Hypotheses

Because previous research demonstrates a relationship between being FO and performance on various EF tasks (Pierucci et al., 2013), it was hypothesized that engaging in fantasy-play would facilitate the development of EFs. Specifically, it was expected that there would be a significant difference between intervention conditions on post-test measures of EFs such that children in the fantasy-play condition would show an increase in EF abilities beyond that of children in the realistic-play or no intervention conditions. In addition, if there were differences between the realistic-play and no intervention conditions, it was anticipated that the realistic play condition would show greater increases in EF abilities than the no intervention condition because the act of engaging in play may inadvertently exercise cognition to some degree. It should be noted, FO was not hypothesized to change over the course of the intervention period. In fact, it is highly unlikely that FO would show any changes over time, as FO appears to be a stable trait in childhood (Pierucci et al., 2013). Rather, it was expected that engaging in fantastical play would create an enriched environment through which EFs could be scaffolded.

METHODOLOGY

Design

The present study utilized an experimental design with play condition as the independent variable and EF performance as the dependent variable. The independent variable had three levels of manipulation: fantasy-play, realistic-play, and control. After five weeks of participating in a play intervention, participants' scores of various EF measures (i.e., working memory, inhibitory control, attention shift) were compared across the three groups in order to determine if fantasy-play increased EF abilities beyond that of realistic play or no intervention.

Participants

The current study included 113 preschoolers between two and five years of age, with 54 girls and 59 boys. Eighty-nine percent of children were Caucasian, 6% were African American, 1% were Asian, and 4% were other or not reported. All participants were recruited from local preschools and day-care centers in Tuscaloosa, Alabama. Two children were excluded from the study because they scored below the 20^{th} percentile on the Peabody Picture Vocabulary Test, a measure of receptive vocabulary. The majority of children were from middle-class families in the southeastern United States (see Appendix E).

Pre-Test Measures

Parent/Teacher Questionnaires. Parents completed a demographics questionnaire in which they were asked questions regarding the ethnicity, gender, and age of their child. In addition, teachers completed a Fantasy Orientation Questionnaire (Gilpin, 2009) in order to assess each child's fantasy orientation level. Teachers were asked a total of four questions. The first question asked whether or not the child believes in a list of seven fantastical figures. Answers to this question received a score of 1 if the parent/teacher indicated that the child believed the entity was real, .5 if they thought the child was unsure of the entity's existence, and 0 if the child believed the entity was pretend (see Appendix A). The total score for this first item was calculated as a percentage of the total points possible (determined by the number of questions answered) in order to accommodate responses that were left blank by teacher respondents. The second question asked the teacher to list the child's favorite books, games, TV shows, and videogames. Responses to this item were coded from a standardized list of responses used to score similar questions in previous research (Boerger, Tullos, & Woolley, 2009; Gilpin, 2009; Woolley, Boerger, & Markman, 2004). Answers that were more reality based received a score of 0 (e.g., hide-and-seek or Madeline), those that were representational or anthropomorphic but not fantastical received a score of 1 (e.g., cars or Curious George), and those that were high fantasy based received a score of 2 (e.g., superheros or Mary Poppins). See Appendix A for more details. Next, teachers were asked to rate a child's level of FO using a 5-point scale, with 1 indicating a strong interest in reality (e.g., plays sports), 2 indicating some interest in fantasy but mostly interested in reality, 3 indicating equal interests in fantasy and reality play/media, 4 indicating mostly interested in fantasy but some interest in reality, and 5 indicating a strong interest in fantasy (e.g., often engages in pretense, enjoys fantastical books; see Appendix A). The final question asked teachers to indicate whether or not a child has an imaginary friend.

In addition to asking teachers about students' FO levels, teachers were also asked to report on their students' EF behaviors at school. The Behavior Rating Inventory of Executive Function – Preschool (BRIEF-P; Gioia, Isquith, Retzlaff, & Espy, 2002), which contains 63 questions to gauge various clinical EF scales (i.e., Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, Monitor) was used. These scales form two broader indexes (i.e., Behavioral Regulation and Metacognition) and an overall score, the Global Executive Composite. Teachers were asked to report how often (i.e., never, sometimes, or often) certain behaviors have been a problem during the last six months (e.g., overreacts to small problems, becomes upset with new situations). The BRIEF-P was coded using standard procedures outlined in the BRIEF-P manual (Gioia et al., 2002). Higher scores on this measure indicate poorer executive control. Scores were eliminated from analyses if they were flagged as "inconsistent" as outlined by BRIEF-P standardized scoring or if they were greater than 2.5 standard deviations from the mean of all scores. This measure of EF is highly reliable (α = .80-.98) and is consistent across different administrations (retest = .82-.88).

Child Interview Pretense and Fantasy Orientation Measures. In addition to teacher reports of a child's FO, two child questionnaires were used to assess participants' level of FO. The first of the two measures was Singer and Singer's (1981) Imaginative Play Predisposition Interview. Children were asked questions about their favorite game, favorite toy, favorite story, and favorite television show. In addition, participants were asked what they like to do when they are alone and with other kids as well as what they think about before going to sleep. These questions were independently coded by two trained raters. Responses indicating interest in fantastical toys, stories, and/or games (e.g., fairies or superheroes) were given a score of 2. Responses involving anthropomorphized/animated toys, stories, and/or games (e.g., stuffed animals, toy cars) were given a score of 1 and responses involving realistic toys, stories, and/or games (e.g., checkers, bingo) were given a score of 0. Responses that were unintelligible were not scored. The total score for this measure was calculated as a percentage out of total points possible (depending on how many items were assigned scores), with a higher percentage indicating higher FO (see Appendix A).

The second measure that was used to assess participants' FO is Taylor and Carlson's (1997) imaginary companion and impersonation interview. This measure is divided into two separate scales. The first scale includes a maximum of 10 questions, beginning with whether or not a child has an imaginary companion. If the child responded "yes," they were asked to provide further details including their imaginary companion's name, if their imaginary companion is a toy or completely pretend, the gender and age of their companion, what they like and dislike about their companion, where their companion lives, and where their companion sleeps. The follow-up questions were used to confirm that the child actually had an imaginary companion and that they did not answer the first question based on a real friend. On this scale, a score of 1 indicated a child reporting an imaginary companion(s) and a score of 0 indicated a child saying they did not have one. The second scale of the imaginary companion and impersonation interview assessed participants' likelihood of engaging in pretend play and impersonating behaviors. Specifically, each child was asked if they pretend to be an animal, a different person, or anything unrelated to their self (e.g., an airplane). Answers to each of these questions were followed-up with questions probing for more details, such as what the child specifically pretends to be. Participants received a score of 1 for each question to which they answered "yes" (see Appendix A). Scores ranged from 0-3 points per participant, with higher scores indicating higher levels of FO.

Both measures of FO were chosen based on their use in past research (Pierucci et al., 2013; Sharon & Woolley, 2004; Taylor & Carlson, 1997; Taylor, Cartwright, & Carlson, 1993; Woolley et al., 2004). Inter-rater reliabilities for all FO measures range from 73% to 93%

(Sharon & Wooley, 2004; Taylor & Carlson, 1997). In addition, scores on each measure appear to remain consistent across different testing sessions and correlate highly with one another (Sharon & Wooley, 2004). See Appendix C for a comprehensive table of all FO measures.

Additionally, pretense was assessed using the Toy Phone Task (Tahiroglu, Mannering, & Taylor, 2011; Taylor, Sachet, Maring, & Mannering, 2013). Each participant was presented with a toy phone and asked if they would like to pretend to call a friend with whom they like to play. The child was then allowed to play with the phone as they wished. During this time, the experimenter recorded the child's behaviors on five dimensions: dialing a number, holding the phone to their ear, talking on the phone, listening to the other person, and generating a conversation that went beyond stereotyped greetings such as "hi" and "how are you". For each action that the child performed, they received a score of 1. Thus, the total score for this measure ranged from 0-5. In addition, the experimenter rated the length of the conversation. A score of 0 was given if the child did not engage in any pretend conversation, a score of 1 was given to a short conversation (i.e., less than 10 sec), a score of 2 was given to a medium length conversation (i.e., between 10 and 30 sec), and a score of 3 was given to a long conversation (i.e., greater than 30 sec). Higher scores indicate higher levels of FO.

Child Interview Executive Function Measures. Four previously established measures were used to assess various aspects of EF (see Appendix D). In order to assess working memory, the forward digit span task was used (Davis & Pratt, 1996). During this task, the experimenter showed the child a puppet and told the child that the puppet likes to repeat everything they say. They then demonstrated by saying "3, 6" and then having the puppet repeat back "3, 6." After one additional example, the experimenter allowed the child to take on the role of the puppet and repeat the digits after the experimenter had stated them. A child's total score on this task was equal to the highest number of digits that he or she successfully repeated. The forward digit span is an appropriate measure of working memory for young children and has been shown to be reliable across different testing sessions ($r = .78$; Henry, 2001). In addition, this task is correlated with other measures of working memory (i.e., backward digit span, word span, pattern span, spatial span, listening span), thus providing evidence for the construct validity of this measure (Henry, 2001).

Although a classic Stroop task is typically used to assess inhibitory control, this measure is not developmentally appropriate for the children of the present study. Therefore, an age appropriate Animal Stroop task (Wright et al., 2003) was used as one measure to assess inhibition. This task, which was administered using a computer, began with a warm-up during which time participants were shown cartoon pictures of four animals that would appear throughout the task (i.e., duck, sheep, pig, cow). Participants then completed two trials containing 24 images each. One of the trials was a matching trial in which the animal's body matched the animal's face. Participants were told to name the animal as quickly as possible. The other trial was a Stroop trial in which the animal's face did not match the animal's body (e.g., pig head on a cow body). Participants were told to name the animal's body as quickly as possible and thus had to inhibit their automatic response of recognizing the animal by its face. Included in this Stroop trial was a control condition in which the animal bodies had cartoon human faces. Again, participants were instructed to name the animal's body as quickly as possible. Response times and errors were recorded. Errors included incorrect responses, stutters, and corrections. Because children are highly distractible, response times were easily influenced by distractions. For example, there may have been a 30 second distraction when the mean response time was 2 seconds. Therefore, in order to clean the data from extreme distractions, if any one item was more than three standard deviations from the mean response time in that trial, this score was excluded from calculating the average response time. For errors, if any child scored more than three standard deviations from the mean, their data were excluded from analyses. Each child's score was calculated as a percentage (((stroop-matching)/matching)x100), with higher percentages indicating poorer cognitive inhibition. High internal consistency has been demonstrated for both reaction time and error measures among all four trials ($\alpha = .56$-.93; Wright et al., 2003). In addition, this task is sensitive to developmental changes and individual differences in inhibitory control, thus indicating its validity as a measure of cognitive inhibition (Wright et al., 2003).

The Day/Night Task was used as an additional measure of inhibitory control (Gerstadt et al., 1994). This Stroop-like task presents each participant with two pictures: a picture of moon and stars (i.e., "nighttime") and a picture of a blue sky and grassy field (i.e., "daytime"). Participants were instructed to say "day" when the experimenter pointed to the picture of the moon and stars, and to say "night" when the experimenter pointed to the picture of the blue sky and grassy field. Before completing 16 test trials, each child underwent two practice trials during which time each picture was pointed to once. If a child answered incorrectly during either of these practice trials, the experimenter repeated both rules and the practice trials as necessary. If the child responded correctly on 75% of the practice trials, it was assumed that they did not understand the task and their data were not included in the final analyses. Number of correct, incorrect, and self-corrected (i.e., initially incorrect but self-corrected to appropriate answer) responses were recorded. Errors included incorrect responses, stutters, and corrections. Correct answers were given a score of 1, self-correct answers were given a score of .5, and incorrect answers were given a score of 0 resulting in a maximum possible score of 16. Higher scores on this measure indicate better cognitive inhibition. This measure has an inter-rater reliability of .90 and alpha values ranging from .56-.93 (Carlson & Moses, 2001; Gerstadt et al., 1994; Wright et al., 2003). In addition, it is correlated with other established measures of inhibitory control (Carlson & Moses, 2001).

Finally, participants' ability to shift their attention was assessed using the Standard Dimensional Change Card Sort task (Zelazo, 2006; Zelazo et al., 2003). Children were presented with two small boxes labeled with two different target images (e.g., a blue square and a red star). The experimenter told the child that they were going to play two games, the "shape game" and the "color game." They began by explaining the shape game. The child was told to place all the star cards (both blue and red) in the container labeled with the red star, and all square cards (both blue and red) in the container labeled with the blue square. After the child correctly sorted five cards consecutively, the experimenter told the child to stop playing the shape game and start playing the color game. The experimenter explained to the child that they should now place all blue cards (stars and squares) in the container labeled with a blue square, and all of the red cards (stars and squares) in the container labeled with a red star. The task ended when the child correctly sorted five cards consecutively. The order of the color/shape game was randomized. This task was scored as the number of inaccurate card sorts after the rule switch. This particular task was chosen because it is developmentally appropriate for children and is reliable and valid (Zelazo, 2006; Zelazo et al., 2003).

Control Measures. Because previous research demonstrates that vocabulary is strongly related to EF performance in preschool children (Carlson & Moses, 2001), vocabulary level was assessed during pre-testing in order to match the experimental groups on vocabulary level and thus control for individual differences. The Peabody Picture Vocabulary Test, Fourth Edition

Form B (PPVT-4; Dunn & Dunn, 2007) was administered to each participant. Children were shown a colored book that contained several pages with four pictures on each page. The experimenter said a word that related to one of the pictures on the page and asked the child to point to the picture that best matched the word. After the child passed the training items, the experimenter progressed through the pages of the book, with each item getting progressively more challenging. The PPVT-4 consists of 19 sets with 12 items in each set for a total of 228 items. Participants were tested until they incorrectly identify eight vocabulary words within one set of 12 items, at which time the test ended.

The PPVT-4 was coded using standard procedures outlined in the Form B manual. The PPVT-4 exhibits excellent internal consistency and test-retest reliability ($\alpha > .90$; Dunn & Dunn, 2007). In addition, this task correlates with several other language measures giving evidence for the construct validity of the PPVT-4 (Dunn & Dunn, 2007).

Post-Test Measures

During post-test assessment, teachers, who were blind to the hypotheses and conditions, completed the same BRIEF-P used during the pre-test period in order to examine any teacher observed changes in EFs during the five week intervention period. Additionally, each participant was tested on all of the FO measures utilized during the pre-test period as well as the same measure of attention shift, working memory, and the Animal Stroop task. The following new measures of EF were also included in order to rule out practice effects (see Appendix D).

The Grass/Snow task was used as an additional measure of inhibitory control during post-testing in place of the Day/Night task used during pre-test (Carlson & Moses, 2001). This Stroop-like task presented each participant with two pictures: a picture of a grassy field (i.e., "grass") and a picture of a snowy field (i.e., "snow"). Participants were instructed to point to the grassy field when the experimenter said "snow", and to point to the snowy field when the experimenter said "grass". Before completing 16 test trials, each child underwent two practice trials during which time each target word was said once. If a child pointed incorrectly during either of these practice trials, the experimenter repeated both rules and the practice trials as necessary. If the child responded correctly on 75% of the practice trials, it was assumed that they did not understand the task and their data were not included in the final analyses. Number of correct, incorrect, and self-corrected (i.e., initially incorrect but self-corrected to appropriate answer) responses were recorded. Correct answers were given a score of 1, self-correct answers were given a score of .5, and incorrect answers were given a score of 0 resulting in a maximum possible score of 16. Higher scores on this measure indicate better inhibitory control. In addition to being correlated with other established measures of inhibitory control (Carlson & Moses, 2001), this measure has alpha values ranging from .56-.93 (Carlson & Moses, 2001; Gerstadt et al., 1994; Wright et al., 2003).

It should also be noted that 22 participants completed a Forward Word Span task (Carlson et al., 2002) instead of the Forward Digit Span task of working memory at post-test. This measure tests children's working memory by asking them to remember and repeat lists of single-syllable, non-semantically related words in a forward order (e.g., experimenter says "book, cup" and child replies, "book, cup"). The list size increased with each successive trial starting with 2 words. Each trial was scored on a scale from 0-2. A child received a score of 0 if they omitted a word, substituted a word, or inserted a word. A child received a score of 1 if they repeated the correct words but in an incorrect order. Finally, a child received a score of 2 if they made no errors. There were 15 total trials. The task ended when a child received a score of 0 on three consecutive trials. In piloting this measure with the first 22 participants, it became apparent that this measure produced large floor effects. Thus it was replaced by the Forward Digit Span at post-test for the remaining participants. However, data on this measure is still used to assess working memory at post-test for the first 22 participants of the current study. These scores were transformed to be on the same scale as the Forward Digit Span task. Specifically, each child's score on the Forward Word Span was multiplied by 7 (i.e., the total score possible on the Forward Digit Span task) and then divided by 16 (i.e., the highest score achieved on the Forward Word Span task). In doing this transformation, the mean and standard deviation of transformed Forward Word Span task scores were consistent with the mean and standard deviation of the Forward Digit Span task scores at post-test. Therefore, all post-test scores of working memory are reported as Forward Digit Span scores.

Procedure

After obtaining consent for participation from parents and assent for participation from the children, all participants were individually interviewed by research assistants on the pre-test measures of FO, pre-test measures of EF, and PPVT-4 described above in order to assess baseline levels of FO and EF. The order of administration of the measures was randomized across participants. Each participant was interviewed for approximately 45 min and received stickers for participating. Parents and teachers also completed demographic and FO/EF questionnaires, respectively, at this baseline period.

Participants were then randomly assigned to one of three intervention conditions: fantasy-play, realistic-play, and control. All conditions were matched on pre-test measure scores in order to ensure that the groups did not differ on EF, FO, or vocabulary levels prior to intervention. In the fantasy-play condition, research assistants worked with groups of children to come up with a fantastical script and then encouraged them to act it out. For example, children may have pretended that their feet were wheels and that they could skate on them or that they were birds flying around a forest. In addition, children may have pretended to go on an adventure to the moon and interact with space creatures. The scripts were generated by the students, but the research assistants provided scaffolding (e.g., encouragement, questions that inspired imagination such as "What do you see on the moon?", and suggestions as needed). In the realistic-play condition, participants engaged in action oriented group activities that did not require any imagination and were not designed to specifically tap executive function skills. Research assistants were given a list of commonly played activities including, but not limited to duck duck goose, action songs (e.g., "Head Shoulders Knees and Toes"), coloring, and hot potato. Games in the fantasy-play condition and realistic-play conditions were well matched in terms of child interest, activity-level and opportunities for child-initiated choices. Finally, in the control condition, participants did not participate in any intervention games, proceeding with business as usual.

Prior to the beginning of the intervention period, research assistants were trained on how to implement each intervention. During this training, research assistants were given a list of example scenarios/activities to complete. After discussing these activities as well as the intervention protocol in detail, research assistants were given an opportunity to practice implementing each condition and ask questions. Throughout the intervention period, research assistants were asked to record what occurred during each training session using a daily log provided by the researchers (e.g., how engaged each child was, how solitary their play was, how fantastical their play was; see Appendix B). These logs were analyzed in order to ensure that interventions were implemented as planned and that both intervention groups were performing similar levels of activity on a daily basis.

Based on previous EF training studies that have successfully increased EF (Bergman Nutley et al., 2011; Thorell et al., 2009), all groups underwent five weeks of intervention. Training occurred every school day for a total of 25 training sessions that lasted 15 minutes each. After five weeks of intervention, all participants underwent post-testing using the measures described above. Post-test interviewers and teachers were blind to the hypotheses of the study and the conditions of each child.

RESULTS

Preliminary Analyses

Participants were randomly assigned to one of the three play conditions (i.e., fantasy-play, realistic-play, or control). The breakdown of the play groups were as follows. There were a total of 39 children in the fantasy-play condition (19 male, 20 female), 33 children in the realistic-play condition (17 male, 16 female), and 41 children in the control condition (18 male, 23 female). The average age of children in the fantasy-play group was 3.8 years, 4 years in the realistic-play group, and 3.9 years in the control condition. The groups did not differ significantly on age, gender, or ethnicity and were evenly distributed across the 4 preschools included in the study (see Appendix E).

In order to ensure that the play conditions did not differ prior to the intervention, scores on all pre-intervention FO, EF, and vocabulary measures were individually submitted to a one-way ANOVA with play condition (i.e., fantasy-play, realistic-play, and control) as the independent variable. The results of these analyses indicated that there were no differences between groups on any measures prior to the intervention (see Table 1, 2, 3, 4) with the exception of the Global Executive Composite scale of the BRIEF-P, $F(2, 106) = 3.104$, $p < .05$, $\eta^2 = .06$. Post-hoc analyses indicated that the realistic-play group (M = 66.84) scored significantly higher than the control group (M = 57.53) on this scale. Differences between the realistic-play group and the fantasy-play group (M = 60.49) as well as differences between the fantasy-play group and the control group on this composite scale were not statistically different.

Because groups were not equal on this measure prior to the intervention, this specific scale of the BRIEF-P will not be analyzed during post-intervention analyses. With the exclusion of this specific scale, groups were equal on all measures of FO and EF prior to the start of the play intervention. It should be noted that some variables did not meet the homogeneity of variance assumption of ANOVA. When this occurred, a Welch correction was used. The remaining analyses were corrected when necessary.

Table 1

Pre-Intervention Group Differences on Measures of FO

Measure	**Condition**	**N**	**Range of Scores**	**Mean**	**SD**	**F**	**Sig.**	**η^2**
IC	Fantasy	39	0-1	.205	.409	.180	.836	.00
	Realistic	33		.242	.435			
	Control	38		.236	.446			
IMP	Fantasy	39	0-3	1.90	1.07	1.04	.357	.02
	Realistic	33		2.00	1.17			
	Control	39		1.64	1.06			
IPP	Fantasy	39	0-1	.463	.186	.275	.760	.00
	Realistic	33		.452	.172			
	Control	39		.450	.170			
TPT Total	Fantasy	39	0-5	2.90	1.45	1.32	.271	.02
	Realistic	33		3.21	1.41			
	Control	39		2.64	1.58			
TPT Length	Fantasy	38	0-3	1.08	.969	1.10	.336	.02
	Realistic	33		1.06	.899			
	Control	39		.795	.923			

Note. IC – Imaginary Companion Interview; IMP – Impersonation Interview; IPP – Imaginative Play Predisposition Interview; TPT Total – Toy Phone Task Total Score; TPT Length – Toy Phone Task Length of Conversation.

Table 2

Pre-Intervention Group Differences on EF Measures

Measure	Condition	N	Range of Scores	Mean	SD	F	Sig.	η^2
FDS	Fantasy	39	0-7	3.95	.887	1.19	.309	.02
	Realistic	32		4.03	1.03			
	Control	39		4.28	1.05			
AS Errors	Fantasy	37	-.83-10	2.32	2.09	.483	.618	.01
	Realistic	31		1.92	1.84			
	Control	36		1.83	2.70			
AS RT	Fantasy	38	-.28-2.55	.717	.505	.271	.763	.01
	Realistic	32		.675	.368			
	Control	38		.758	.506			
CS Errors	Fantasy	38	0-25	2.89	4.69	2.47^	.092	.03
	Realistic	31		1.61	3.07			
	Control	38		3.39	3.77			
D/N	Fantasy	27	0-16	12.33	4.31	2.46^	.097	.05
	Realistic	26		13.46	2.49			
	Control	27		11.19	4.98			

Note. FDS – Forward Digits Span (working memory); AS Errors – Animal Stroop Errors (inhibitory control); AS RT – Animal Stroop Reaction Time (inhibitory control); CS Errors – Card Sort Errors after Rule Switch (attention shift); D/N – Day Night Task total score (inhibitory control). ^Welch correction used.

Table 3

Pre-Intervention Group Differences on Teacher Reported FO

Measure	Condition	N	Range of Scores	Mean	SD	F	Sig.	η^2
TQ Belief	Fantasy	39	0-1	.698	.177	.994^	.375	.02
	Realistic	33		.757	.216			
	Control	28		.658	.274			
TQ Favorite	Fantasy	36	0-2	1.75	.500	.292	.748	.01
	Realistic	30		1.73	.521			
	Control	33		1.82	.392			
TQ Games	Fantasy	38	0-2	1.24	.714	.595	.554	.01
	Realistic	32		1.16	.723			
	Control	39		1.33	.621			
TQ FO Level	Fantasy	36	1-5	2.94	1.04	1.26^	.290	.03
	Realistic	33		3.27	1.07			
	Control	38		2.92	.850			
TQ IC	Fantasy	38	0-1	.105	.311	1.97^	.148	.04
	Realistic	33		.212	.415			
	Control	38		.053	.226			

Note. TQ Belief – Teacher Questionnaire Belief in Fantastical Entities; TQ Favorite – Teacher Questionnaire Favorite TV Shows, Video Games, Books; TQ Games – Teacher Questionnaire Favorite Games; TQ FO Level – Teacher Questionnaire FO Level; TQ IC – Teacher Questionnaire Imaginary Companion Report. ^Welch correction used.

Table 4

Pre-Intervention Group Differences on Teacher Reported EF

Measure	Condition	N	Mean	SD	F	Sig.	η^2
BRIEF Inhibit	Fantasy	38	54.32	8.34	1.98^	.148	.04
	Realistic	29	59.76	13.51			
	Control	37	54.98	13.09			
BRIEF Shift	Fantasy	39	44.03	16.03	.441	.644	.01
	Realistic	33	41.76	20.29			
	Control	37	45.70	16.46			
BRIEF Emotional Control	Fantasy	39	48.64	17.60	.027^	.973	.00
	Realistic	33	47.61	23.38			
	Control	38	48.71	18.00			
BRIEF Working Memory	Fantasy	39	49.69	14.52	.152^	.869	.00
	Realistic	32	47.47	19.52			
	Control	38	49.45	15.01			
BRIEF Plan/Organize	Fantasy	39	47.36	17.48	.795	.454	.01
	Realistic	33	42.97	19.25			
	Control	37	47.84	16.33			
BRIEF Self-Control Index	Fantasy	38	51.82	9.62	1.18^	.313	.02
	Realistic	32	56.63	15.55			
	Control	38	52.21	15.05			
BRIEF Flexibility Index	Fantasy	39	48.85	12.49	.006^	.994	.00
	Realistic	33	49.24	16.94			
	Control	38	48.92	14.11			
BRIEF Emergent Metacognition Index	Fantasy	39	51.38	11.95	.028	.972	.00
	Realistic	32	50.69	13.78			
	Control	37	51.16	11.77			
BRIEF Global Executive Composite	Fantasy	39	60.49	16.96	2.104	.049*	.02
	Realistic	32	66.84	16.67			
	Control	38	57.53	13.67			

Note. All values presented are standardized T-scores found in the BRIEF-P manual. Range of possible scores was not included because the range differed for each participant depending on age and gender. *$p < .05$. ^Welch correction used.

Finally, it should be noted that not all of the measures used in the present study were normally distributed (i.e., Day/Night Task, Grass/Snow Task, Card Sort Task). Ceiling effects were observed on each of these measures indicating that they were not developmentally appropriate for our sample. Although prior research indicates that EFs are highly integrated in early childhood (Carlson et al., 2002; Diamond & Taylor, 1996; Hughes et al., 2010; Wiebe et al., 2008; Wiebe et al., 2011), it is not likely that we will observe consistent effects on all measures of EF because of these ceiling effects. Results from tasks that had sufficient variance and were normally distributed are presented below.

Differences between Conditions at Post-Test

Scores on all post-intervention FO and EF measures were individually submitted to a one-way ANOVA with play condition (i.e., fantasy-play, realistic-play, and control) as the independent variable. As hypothesized, the three groups differed in meaningful ways after the intervention (see Table 5, 6, 7). Specifically, there were significant differences between groups on the Toy Phone Task total score (FO), $F(2, 101) = 3.924$, $p = .02$, $\eta^2 = .07$ and the Toy Phone Task length of conversation (FO), $F(2, 97) = 4.829$, $p = .01$, $\eta^2 = .09$. Specifically, the fantasy-play group had higher total scores on this task (M = 3.59) than the control group (M = 2.56) as well as longer conversations (M = 1.34) than the control group (M = .70; see Figure 1). There were no statistically significant differences between the fantasy-play and realistic-play group or the realistic-play group and control group on either measure. In addition to 3-way group differences in the Toy Phone Task, there were also marginally significant differences between means on the inhibition subscale of the BRIEF-P (EF), $F(2, 97) = 2.570$, $p = .08$, $\eta^2 = .05$. Next 2-way group differences in FO and EF are presented.

Table 5

Post-Intervention Group Differences on Measures of FO

Measure	**Condition**	**N**	**Range of Scores**	**Mean**	**SD**	**F**	**Sig.**	η^2
IC	Fantasy	38	0-1	.368	.489	.467^	.629	.01
	Realistic	31		.290	.461			
	Control	34		.264	.448			
IMP	Fantasy	39	0-3	1.85	1.09	.945	.392	.02
	Realistic	29		1.48	1.12			
	Control	30		1.67	1.03			
IPP	Fantasy	39	0-1	.546	.193	2.02	.139	.04
	Realistic	30		.463	.177			
	Control	33		.475	.192			
TPT Total	Fantasy	39	0-5	2.90	1.45	3.92	.023*	.07
	Realistic	33		3.21	1.41			
	Control	39		2.64	1.58			
TPT Length	Fantasy	38	0-3	1.34	.938	4.83	.01**	.09
	Realistic	29		.897	.900			
	Control	33		.697	.847			

Note. IC – Imaginary Companion Interview; IMP – Impersonation Interview; IPP – Imaginative Play Predisposition Interview; TPT Total – Toy Phone Task Total Score; TPT Length – Toy Phone Task Length of Conversation. $*p < .05$. $**p < .01$. ^Welch correction used.

Table 6

Post-Intervention Group Differences on EF Measures

Measure	Condition	N	Range of Scores	Mean	SD	F	Sig.	η^2
FDS	Fantasy	39	0-7	4.69	1.76	2.25^	.113	.04
	Realistic	31		3.94	1.26			
	Control	34		4.36	1.54			
AS Errors	Fantasy	38	-.75-6	.920	1.53	.839^	.437	.02
	Realistic	29		1.42	1.87			
	Control	33		.913	1.32			
AS RT	Fantasy	39	-.33-1.73	.590	.383	.167	.846	.00
	Realistic	30		.601	.432			
	Control	34		.641	.365			
CS Errors	Fantasy	38	0-25	1.58	3.21	1.76^	.182	.03
	Realistic	28		3.14	4.78			
	Control	34		2.92	4.03			
G/S	Fantasy	34	0-16	11.54	4.72	.058	.943	.00
	Realistic	26		11.71	4.40			
	Control	30		11.97	5.61			

Note. FDS – Forward Digits Span (working memory); AS Errors – Animal Stroop Errors (inhibitory control); AS RT – Animal Stroop Reaction Time (inhibitory control); CS Errors – Card Sort Errors after Rule Switch (attention shift); D/N – Day Night Task total score (inhibitory control). ^Welch correction used.

Table 7

Post-Intervention Group Differences on Teacher Reported EF

Measure	Condition	N	Mean	SD	F	Sig.	η^2
BRIEF Inhibit Scale	Fantasy	35	51.71	9.18	2.83^	.067	.05
	Realistic	28	58.64	12.96			
	Control	37	54.00	13.91			
BRIEF Shift Scale	Fantasy	36	48.81	9.49	.092	.912	.00
	Realistic	28	49.82	11.92			
	Control	36	48.83	10.26			
BRIEF Emotional Control Scale	Fantasy	36	52.58	11.94	.239	.787	.00
	Realistic	30	53.60	12.74			
	Control	37	52.92	12.38			
BRIEF Working Memory Scale	Fantasy	36	53.75	11.10	.312	.733	.01
	Realistic	29	52.03	11.80			
	Control	36	54.19	11.38			
BRIEF Plan/Organize Scale	Fantasy	36	53.64	12.26	.693	.502	.01
	Realistic	28	50.25	11.94			
	Control	37	52.22	10.17			
BRIEF Self-Control Index	Fantasy	36	52.61	11.43	1.22	.300	.02
	Realistic	28	57.21	10.30			
	Control	37	53.28	13.78			
BRIEF Flexibility Index	Fantasy	36	50.69	10.86	.040	.961	.00
	Realistic	28	51.25	11.57			
	Control	37	51.46	13.04			
BRIEF Emergent Metacognition Index	Fantasy	36	53.86	11.64	.447	.641	.01
	Realistic	29	51.34	11.84			
	Control	37	53.01	10.63			

Note. All values presented are standardized T-scores found in the BRIEF-P manual. Range of possible scores was not included because the range differed for each participant depending on age and gender. **p* < .05. ^Welch correction used.

Fantasy vs. Realistic Only. Because it was hypothesized that the fantasy-play group would show improvements over the realistic-play group due to the nature of FO play, a series of one-way ANOVAs were calculated to examine differences on FO and EF measures between the fantasy-play group and the realistic-play group after the intervention. These two groups did not differ on any measures of FO/pretense except for the "length of conversation" score on the Toy Phone task, $F(1, 65) = 3.842$, $p = .05$, $\eta^2 = .06$. Specifically, the fantasy-play group (M = 1.34) carried on longer pretend conversations than the realistic-play group (M = .90; see Figure 1). Because there were no differences on this measure between the fantasy-play and realistic-play group prior to the intervention, Toy Phone "length of conversation" scores were submitted to a repeated measures ANOVA with pre-intervention scores and post-intervention scores as the within-subjects factor in order to examine the nature of the changes observed. There were not statistically significant changes in Toy Phone Task length of conversation scores from pre-intervention to post-intervention in the fantasy-play group, $F(1, 36) = 2.61$, $p = .115$, or the realistic-play group, $F(1, 28) = 1.061$, $p = .312$. However, the change in the fantasy-play group was marginally significant, and with more power it is expected that a significant increase in Toy Phone length of conversation scores will be observed in the fantasy-play group. See Figure 6.

Differences also existed between the fantasy-play group and the realistic-play group on measures of EF. Specifically, the fantasy-play group remembered significantly more digits on the Forward Digit Span task (i.e., better working memory; M = 4.69) than the realistic-play group (M = 3.94), $F(1, 67) = 4.380$, $p = .040$, $\eta^2 = .06$ (Welch correction used; see Figure 1). In order to examine the nature of the changes observed, Forward Digit Span scores were submitted to a repeated measures ANOVA with pre-intervention scores and post-intervention scores as the within-subjects factor. In the fantasy-play condition, results revealed that there was a significant increase in scores from pre-intervention (M = 3.95) to post-intervention (M = 4.69) in the fantasy-play group, $F(1, 38) = 10.405$, $p = .003$, $\eta^2 = .22$. In other words, participants in the fantasy-play condition remembered more digits after engaging in the fantasy-play intervention than prior to the intervention. However, in the realistic-play condition, there was no change in scores from pre-intervention (M = 4.03) to post-intervention (M = 3.97), $F(1, 29) = .065$, $p = .80$, $\eta^2 = .00$. See Figure 2.

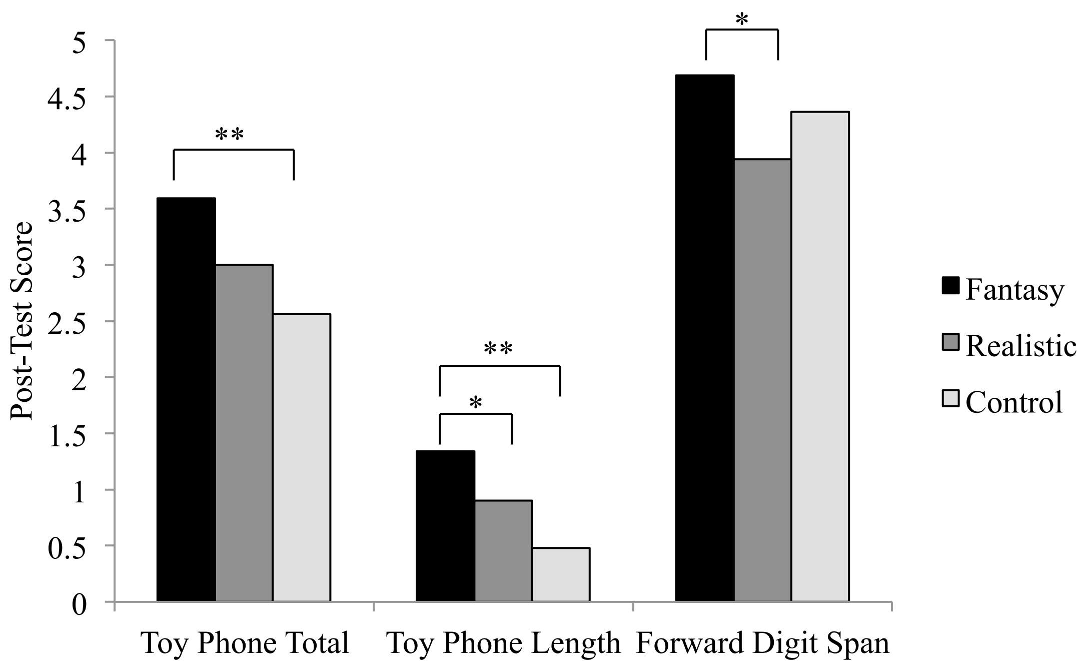

Figure 1 – FO and EF Differences between Play Conditions Post-Intervention. Differences between play conditions on measures of FO and EF after the play intervention. *$p < .05$. **$p < .01$.

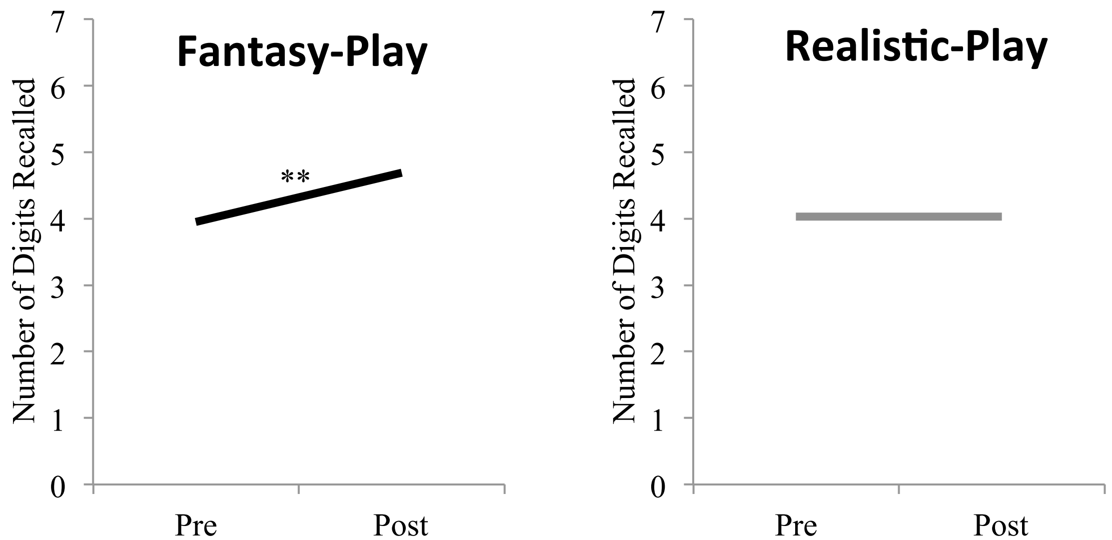

Figure 2 – Changes in Forward Digit Span Scores. Changes in Forward Digit Span scores from pre-intervention to post-intervention. $^{**}p < .01$. The fantasy-play group made significant improvements in working memory from pre- to post-intervention whereas the realistic-play group did not make any changes.

In addition to the working memory results above, the fantasy-play group scored significantly lower on the inhibition subscale of the BRIEF-P (M = 51.71) than the realistic-play group (M = 58.64), $F(1, 47) = 5.71$, $p = .02$, $\eta^2 = .09$ (Welch correction used), indicating better teacher reported inhibitory control in the fantasy-play group after the intervention (see Figure 3). As a reminder, teachers were blind to the hypotheses as well as the condition of each child.

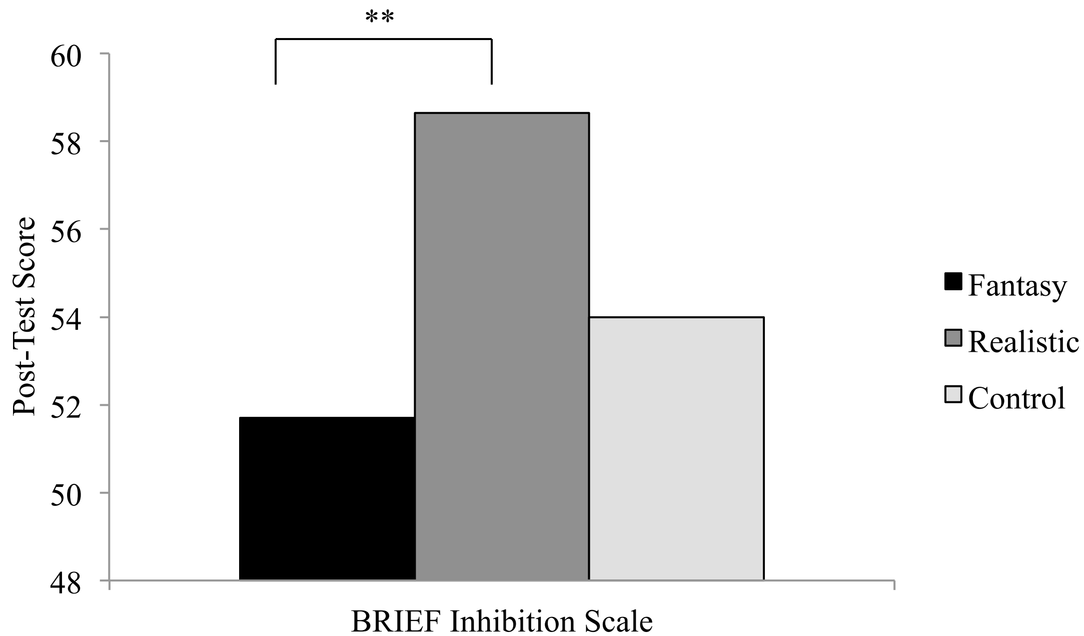

Figure 3 – BRIEF-P Differences between Play Conditions Post-Intervention. Difference between conditions on BRIEF Inhibition Scale post-intervention. **$p < .01$. Note that higher scores indicate poorer inhibitory control.

Again, in order to examine the nature of the changes observed, scores on the inhibition subscale of the BRIEF-P were submitted to a repeated measures ANOVA with pre-intervention scores and post-intervention scores as the within-subjects factor. In the fantasy-play condition, results revealed that there was a significant decrease in scores from pre-intervention (M = 54.14) to post-intervention (M = 51.71), $F(1, 34) = 5.925$, $p = .02$, $\eta^2 = .15$, indicating that participants in the fantasy-play condition were rated by their teachers has having better inhibitory control after the intervention than prior to the intervention. In the realistic-play condition, there was no change in scores from pre-intervention (M = 59.04) to post-intervention (M = 58.64), $F(1, 24) = .001$, $p = .98$, $\eta^2 = .00$. See Figure 4.

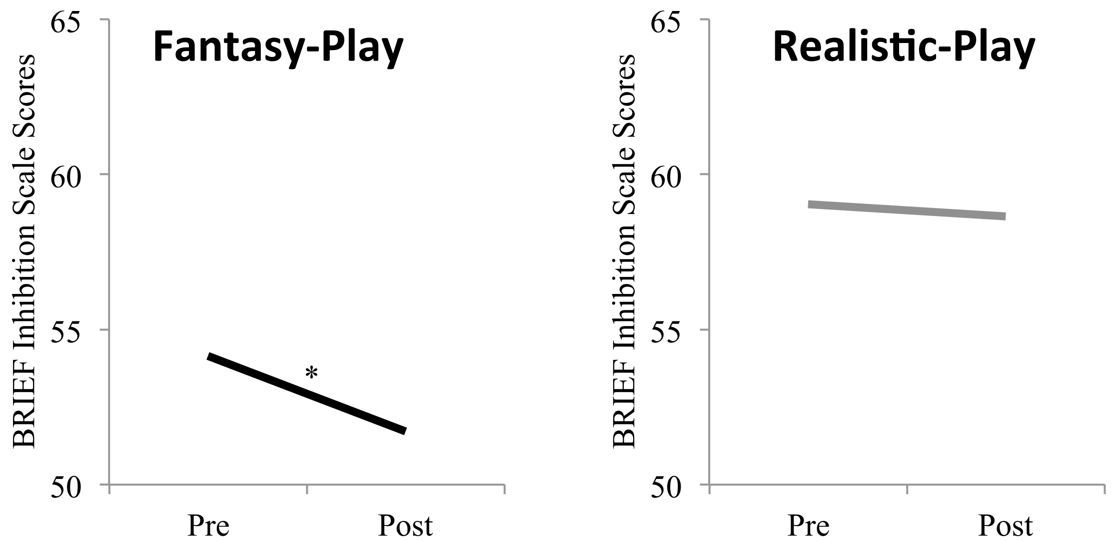

Figure 4 – Changes in BRIEF Inhibition Scale Scores. Changes in BRIEF Inhibition Scale scores from pre-intervention to post-intervention. **p* < .05. Teachers reported that the fantasy-play group made significant improvements in inhibitory control from pre- to post-intervention whereas the realistic-play group did not make any changes. Note that higher scores indicate poorer inhibitory control.

Fantasy vs. Control Only. It was also hypothesized that the fantasy-play group would show improvements over the control group after the intervention. Therefore, these groups were also analyzed separately using a series of one-way ANOVAs. Consistent with the 3-way group differences presented above, significant differences observed were on the Toy Phone Task total score, $F(1, 71) = 7.98$, $p = .006$, $\eta^2 = .10$, and Toy Phone Task length of conversation, $F(1, 69) = 9.135$, $p = .004$, $\eta^2 = .12$. Specifically, the fantasy-play group had higher total scores on this task (M = 3.59) than the control group (M = 2.56) as well as longer conversations (M = 1.34) than the control group (M = .70) after the intervention (See Figure 1). The fantasy-play group and control group did not differ on any other measures of FO or any measures of EF after the intervention.

In order to examine the nature of these significant differences, Toy Phone Task total scores and length of conversation scores were submitted to repeated measures ANOVAs with pre-intervention scores and post-intervention scores as the within-subjects factor. Results revealed that there was a significant increase in Toy Phone Task total scores from pre-intervention (M = 2.90) to post-intervention (M = 3.59) in the fantasy-play group, $F(1, 38) = 5.625$, $p = .02$, $\eta^2 = .13$. By contrast, in the control group there was neither a change in Toy Phone Task total scores from pre-intervention (M = 2.64) to post-intervention (M = 2.58), $F(1, 32) = .081$, $p = .78$, $\eta^2 = .003$, nor a change in Toy Phone Task length of conversation scores from pre-intervention (M = .72) to post-intervention (M = .72), $F(1, 31) = 0.00$, $p > .05$, $\eta^2 = .00$. Interestingly, there also was not a statistically significant change in Toy Phone Task length of conversation scores from pre-intervention (M = 1.05) to post-intervention (M = 1.32) in the fantasy-play group, $F(1, 36) = 2.61$, $p = .115$, $\eta^2 = .07$. This result is marginally significant, and with more power it is expected that a significant increase in Toy Phone Task length of conversation scores will be observed in the fantasy-play group. See Figures 5 and 6.

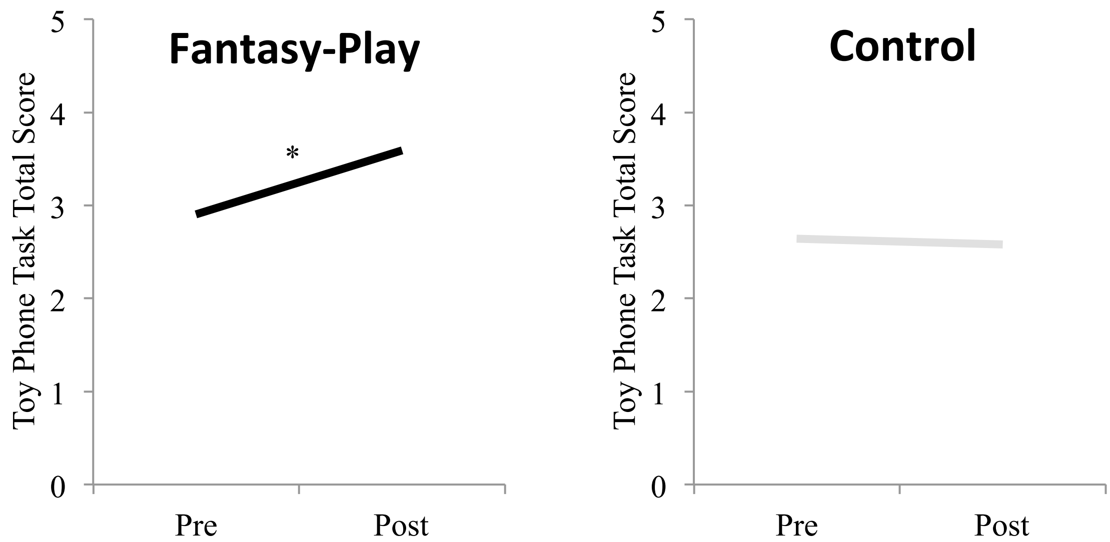

Figure 5 – Changes in Toy Phone Task Total Scores (Fantasy and Control). Changes in Toy Phone Task total scores from pre-intervention to post-intervention. *$p < .05$. Children in the fantasy-play group engaged in significantly more pretending behaviors from pre- to post-intervention whereas the realistic-play group did not make any changes in pretending behavior.

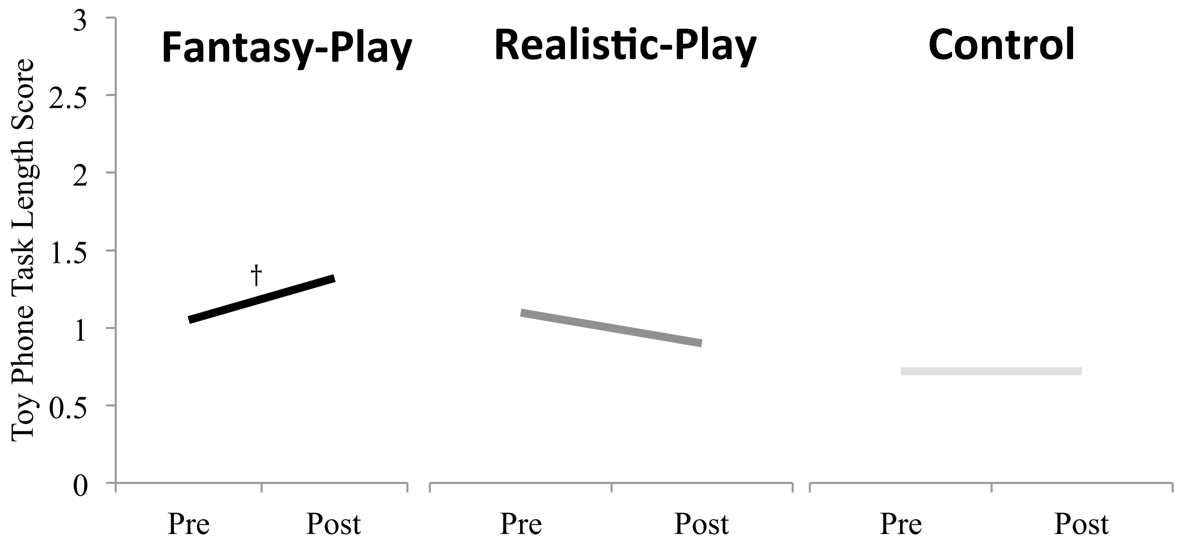

Figure 6 – Changes in Toy Phone Task Length Scores. Changes in Toy Phone Task length of conversation scores from pre-intervention to post-intervention. † $p = .12$. Children in the fantasy-play group tended to engage in longer pretend conversations from pre- to post-intervention whereas the realistic-play and control groups did not make any changes.

Realistic vs. Control Only. In order to examine if post-intervention differences existed between the realistic-play and control conditions, a series of one-way ANOVAs were calculated. Although it was hypothesized that if any differences existed between the realistic-play and control conditions the realistic-play group would score higher on measures of EF, no significant differences were found between these two groups on any measures of FO or any measures of EF after the intervention.

Individual Differences in Related Variables' Influence on Primary Analyses

In addition to collecting data before and after the intervention relating to FO and EFs, data was collected throughout the intervention on several aspects including how engaged each child was, how peer-oriented or solitary their play was, and how fantastical/representational (i.e., style of play) a child's play was on average throughout the entire intervention. After each play session, research assistants rated each child on these three dimensions (see Appendix B). Because children in the control condition did not participate in any intervention, there is no data for the control group on these dimensions. Therefore, all analyses relating to these dimensions only considered the fantasy-play and realistic-play conditions. Interestingly, these dimensions related to the outcome variables in meaningful ways.

The Influence of Engagement. Each day of the intervention, children received a 2 if they were highly engaged, a 1 if they were somewhat engaged, and a 0 if they were not engaged in the play that day. Scores were averaged across all 25 days of the intervention. No children received an average score of 0. Interestingly, the data indicated that a child's level of engagement throughout the intervention, regardless of the intervention condition they were in (i.e., fantasy-play vs. realistic-play), impacted many outcome variables. For example, when collapsing across conditions, children who were highly engaged in the intervention scored higher on the Grass/Snow Task (measure of inhibition; M = 12.63) than children who were only somewhat engaged throughout the intervention period (M = 8.29), $F(1, 18) = 8.741$, $p = .008$, $\eta^2 = .17$ (Welch correction used). In other words, children who were more engaged in the intervention showed better inhibitory control post-intervention than children who were less engaged in the intervention (see Figure 7).

In addition to this significant finding, there were other outcome variables that were approaching significance in this domain. Specifically, children who were highly engaged in the intervention made fewer errors post rule switch on the Card Sort task (measure of attention shift; M = 1.74) than children who were somewhat engaged (M = 3.81), $F(1, 22) = 2.818$, $p = .09$, $\eta^2 = .05$ (Welch correction used; see Figure 7). In addition, children who were highly engaged in the intervention were marginally more likely to report having an imaginary companion (M = .385) compared to children who were somewhat engaged in the intervention (M = .177), $F(1, 34) = 3.157$, $p = .085$, $\eta^2 = .04$ (Welch correction used). Overall, children who were highly engaged in the intervention, whether they were in the fantasy-play group or the realistic-play group, performed better on a range of EF outcome measures.

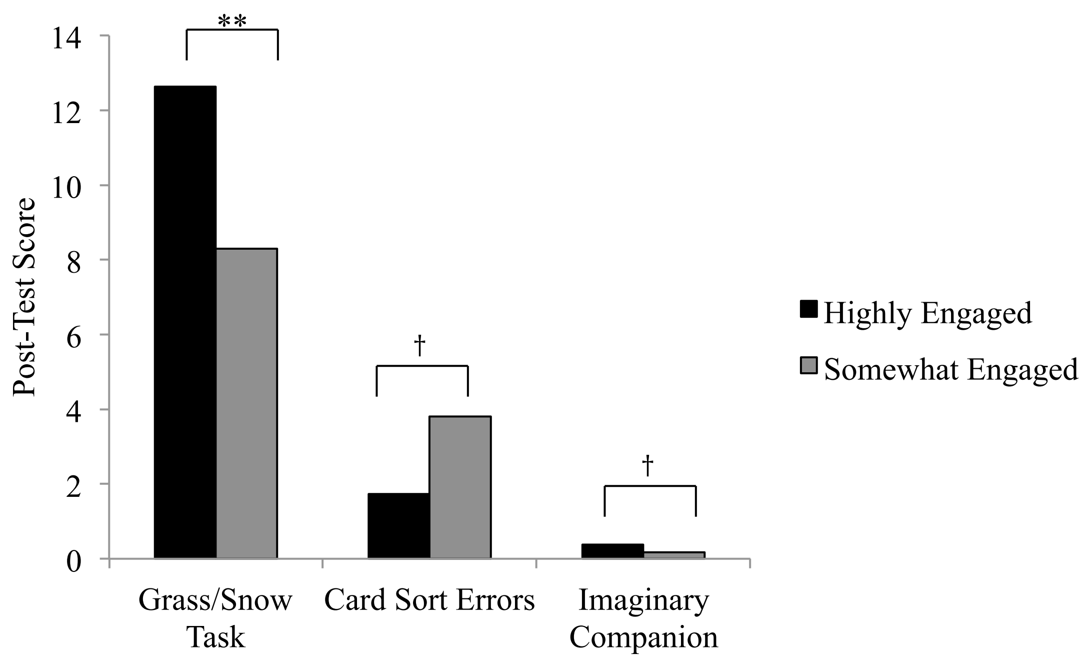

Figure 7 – FO and EF Differences considering Engagement. Differences between highly engaged children and somewhat engaged children collapsed across fantasy-play and realistic-play groups after the intervention. $^{**}p < .01$. † $p < .10$.

Given this information, it is important to re-examine post-intervention differences between the fantasy-play and realistic-play groups only considering children who were consistently, highly engaged throughout the intervention (i.e., engagement score of 2). Note that there were no differences between groups on pre-test measures when only considering children who were highly engaged. To examine post-test differences between groups for highly engaged children, a series of one-way ANOVAs were calculated to examine differences on FO and EF measures between the fantasy-play group and the realistic-play group after the intervention, only considering children who were rated as highly engaged throughout the intervention period. These two groups differed on many measures of FO/pretense including Taylor and Carlson's Impersonation Interview, $F(1, 47) = 6.478$, $p = .014$, $\eta^2 = .12$ (Welch correction used), and "length of conversation" score on the Toy Phone task, $F(1, 47) = 5.147$, $p = .03$, $\eta^2 = .10$. Specifically, the fantasy-play group reported impersonating more animals, people, and other entities (M = 2.21) than the realistic-play group (M = 1.46). Likewise, the fantasy-play group carried on longer pretend conversations (M = 1.57) than the realistic-play group (M = .96). In addition, differences between the fantasy-play group and the realistic-play group on the Toy Phone Task total score were marginally significant, $F(1, 50) = 3.573$, $p = .065$, $\eta^2 = .07$. Participants who were highly engaged in the fantasy-play group scored higher on the Toy Phone Task (M = 3.92) than participants who were highly engaged in the realistic-play group (M = 3.11). See Figure 8.

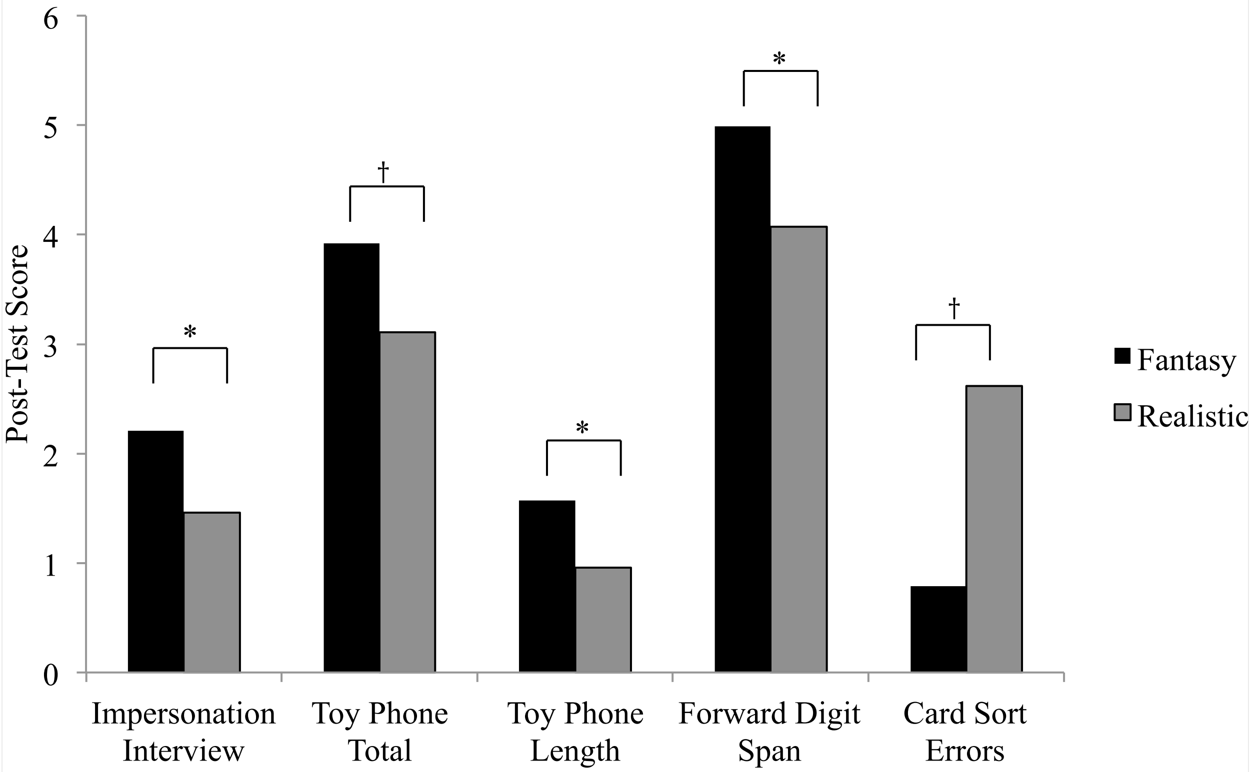

Figure 8 – Fantasy vs. Realistic considering Engagement. Differences between fantasy-play and realistic-play on measures of FO and EF after the play intervention only considering children who were highly engaged. **p* < .05. † *p* < .10.

Differences also existed between the fantasy-play group and the realistic-play group on measures of EF among participants who were highly engaged in the intervention. Specifically, the fantasy-play group remembered significantly more digits on the Forward Digit Span task (i.e., better working memory; M = 4.99) than the realistic-play group (M = 4.07), $F(1, 46) = 6.081$, $p = .017$, $\eta^2 = .11$ (Welch correction used; see Figure 8). In addition, the fantasy-play group scored lower on the inhibition subscale of the BRIEF-P (M = 52.25) than the realistic-play group (M = 58.96), $F(1, 39) = 4.623$, $p = .038$, $\eta^2 = .09$ (Welch correction used), indicating better teacher reported inhibitory control in children who were highly engaged in the fantasy play group after the intervention compared to children who were highly engaged in the realistic play group (see Figure 9). Moreover, there was also a marginally significant finding on errors post rule switch during the Card Sort task, $F(1, 39) = 3.229$, $p = .080$, $\eta^2 = .06$ (Welch correction used). Specifically, participants in the fantasy-play group who were highly engaged throughout the intervention made fewer errors after the rule switch (i.e., better attention shift; M = .79) than highly engaged individuals in the realistic-play condition (M = 2.62; see Figure 8).

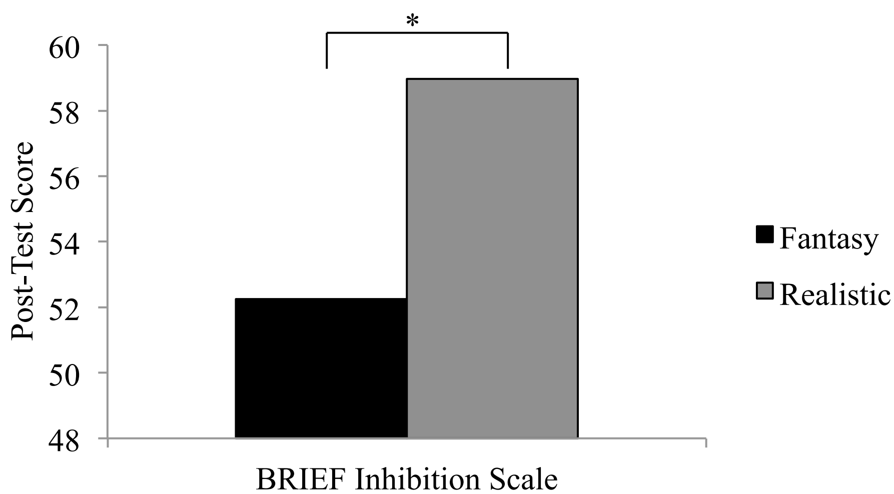

Figure 9 – Fantasy vs. Realistic on BRIEF-P considering Engagement. Differences between fantasy-play and realistic-play on teacher ratings of inhibitory control after the play intervention only considering children who were highly engaged. *$p < .05$.

Influence of Peer-Oriented vs. Solitary Play. Each day of the intervention children were also rated on how much they engaged in play with their peers vs. by themselves. Children received a score of 2 if they consistently engaged in peer-oriented play, a score of 1 if they engaged in a mix of peer-oriented and solitary play (i.e., sometimes engaged in play with peers and sometimes played by themselves), and a score of 0 if they consistently engaged in solitary play that day. Again, scores were averaged across all 25 days of the intervention and no children received an average score of 0.

Interestingly, the data indicate that a child's level of peer-oriented play throughout the intervention, regardless of the intervention condition they were in (i.e., fantasy-play vs. realistic-play), was related to many outcome variables (see Figure 10). For example, children who mostly engaged in peer-oriented play during the intervention remembered more digits on the Forward Digit Task (i.e., measure of working memory; M = 4.55) than children who's play was less peer-oriented (M = 3.52), $F(1, 67) = 5.320$, $p = .024$, $\eta^2 = .07$. In addition, children who consistently engaged in more peer-oriented play throughout the intervention scored marginally higher on the Grass/Snow Task (i.e., better inhibitory control; M = 12.30) than children who engaged in less peer-oriented play (M = 8.75), $F(1, 14) = 3.991$, $p = .066$, $\eta^2 = .10$.

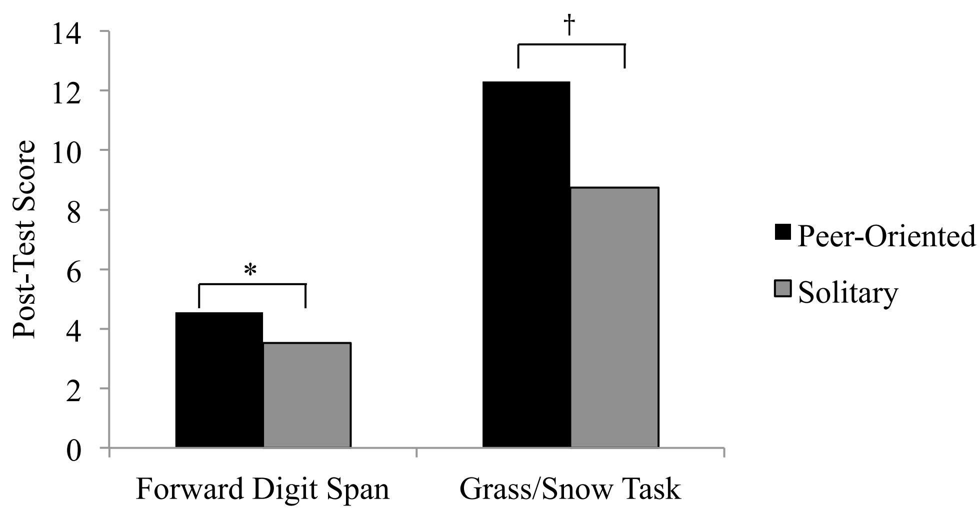

Figure 10 – FO and EF Differences Considering Influence of Peer-Oriented Play. Differences between children who consistently engaged in peer-oriented play and children who engaged in a mix of peer-oriented and solitary play collapsed across fantasy-play and realistic-play groups after the intervention. $*p < .05$.

Because the previous data indicate that there are benefits to engaging in peer-oriented play, it is important to re-examine post-intervention differences between the fantasy-play and realistic-play groups only considering children who were consistently engaging in peer-oriented play (i.e., score of 2). Note that there were no differences between groups on pre-test measures when only considering children who were consistently engaged in peer-oriented play. Therefore, a series of one-way ANOVAs were calculated to examine differences on FO and EF measures between the fantasy-play group and the realistic-play group after the intervention, only considering children who were rated as consistently engaged in peer-oriented throughout the intervention period. These two groups differed on many measures of pretense including the "length of conversation" score on the Toy Phone task, $F(1, 49) = 4.679$, $p = .035$, $\eta^2 = .09$. Specifically, the fantasy-play group carried on longer pretend conversations (M = 1.50) than the realistic-play group (M = .93; see Figure 11). Differences also existed between the fantasy-play group and the realistic-play group on measures of EF among participants who consistently engaged in peer-oriented play throughout the intervention. Specifically, the fantasy-play group remembered significantly more digits on the Forward Digit Span task (i.e., better working memory; M = 5.15) than the realistic-play group (M = 4.03), $F(1, 45) = 8.138$, $p = .007$, $\eta^2 = .14$ (Welch correction used; see Figure 11). In addition, the fantasy-play group tended to score lower on the inhibition subscale of the BRIEF-P (M = 52.82) than the realistic-play group (M = 59.20), $F(1, 45) = 3.702$, $p = .06$, $\eta^2 = .08$, indicating better teacher reported inhibitory control in children who engaged in peer-oriented play in the fantasy-play group after the intervention compared to children who engaged in peer-oriented play in the realistic-play group (see Figure 12).

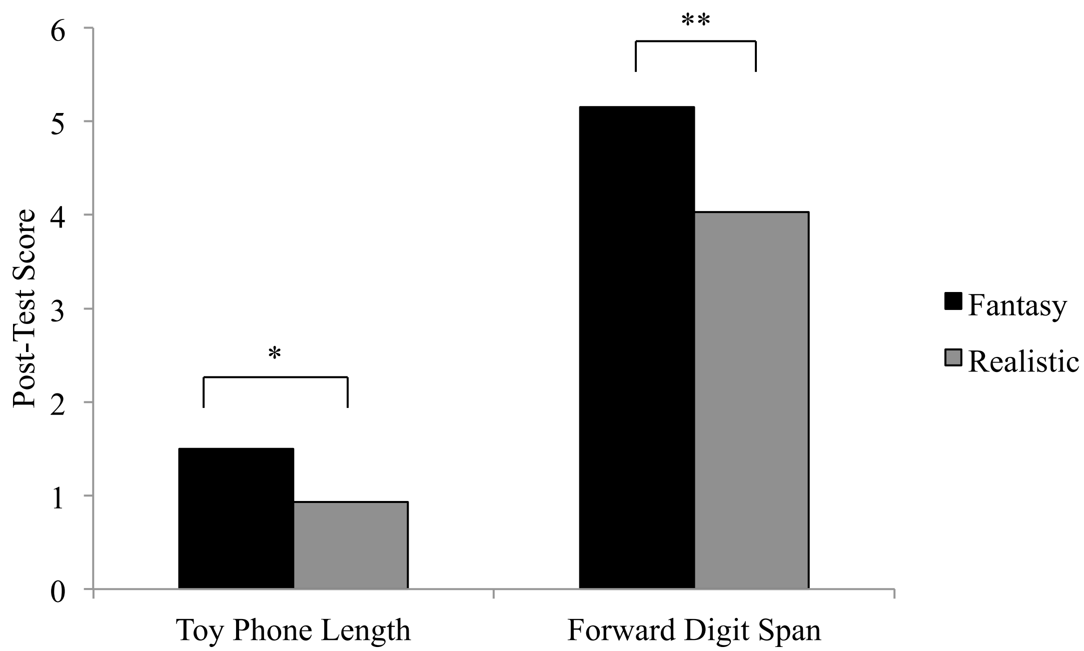

Figure 11 – Fantasy vs. Realistic considering Peer-Oriented Play. Differences between fantasy-play and realistic-play conditions on measures of FO and EF after the play intervention only considering children who consistently engaged in peer-oriented play. **p* < .05. ***p* < .01.

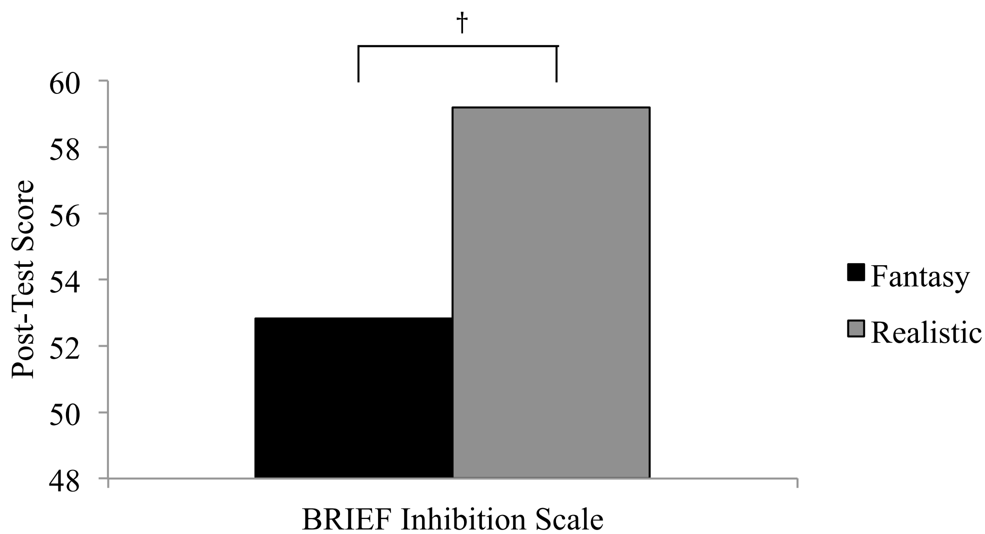

Figure 12 – Fantasy vs. Realistic on BRIEF-P considering Peer-Oriented Play. Differences between fantasy-play and realistic-play conditions on teacher ratings of inhibitory control after the play intervention only considering children who consistently engaged in peer-oriented play. † *p* < .10.

The Influence of Style of Play. Finally, each day children's style of play was rated on a three-point scale. Children received a 2 if their play was highly fantastical (e.g., involving fairies, dragons, princesses), a 1 if their play was representational (i.e., imaginative but not necessarily fantastical such as pretending to cook or pretending to be animals), and a 0 if their play was reality based (i.e., non-imaginative). Scores were averaged across all 25 days of the intervention. As anticipated, all children in the realistic-play group had an average score of 0. Because there was no variability in their scores, children in the realistic-play group will not be considered in the following analyses. It is also important to note that within the fantasy-play condition, there was no relationship between children who scored high on pre-intervention measures of FO/pretense and their average style rating. In other words, all children in the fantasy-play condition were equally likely to engage in highly fantastical play regardless of their initial trait FO.

Within the fantasy-play group, the data indicated that a child's style of play (i.e., fantastical or representational) throughout the intervention was related to many outcome variables. Specifically, children in the fantasy-play condition who were more fantastical throughout the intervention period remembered significantly more digits on the Forward Digit Span task (i.e., better working memory; M = 5.22) than children in the fantasy-play condition who engaged in more representational play (M = 3.96), $F(1, 36) = 5.321$, $p = .027$, $\eta^2 = .13$ (see Figure 13).

There were a few other FO/pretense outcome variables that approached significance when considering how fantastical a child's play was throughout the intervention. The following findings provide corroboration for the observers' ratings of style. Specifically, children who engaged in more fantastically oriented play throughout the intervention reported impersonating more animals, people, and other entities on Taylor and Carlson's Impersonation Interview (M = 2.14) than children who engaged in more representational play (M = 1.47), $F(1, 29) = 3.516$, $p = .07$, $\eta^2 = .09$ (Welch correction used). In addition, children who engaged in more fantastical play tended to carry on longer conversations during the post-intervention Toy Phone Task (M = 1.62) than children who engaged in more representational play (M = 1.06), $F(1, 35) = 3.530$, $p = .069$, $\eta^2 = .09$. See Figure 13.

In addition, EF scores of children in the fantasy-play group who consistently engaged in representational play were compared to children in the realistic-play group. Interestingly, no statistically significant differences were observed between these groups of children. In other words, children in the fantasy-play condition who engaged in more representational play throughout the intervention performed at similar levels to the realistic-play group after the intervention. Thus, the fantasy component of the fantasy-play condition seems to be critical to the benefits in EFs observed in this group.

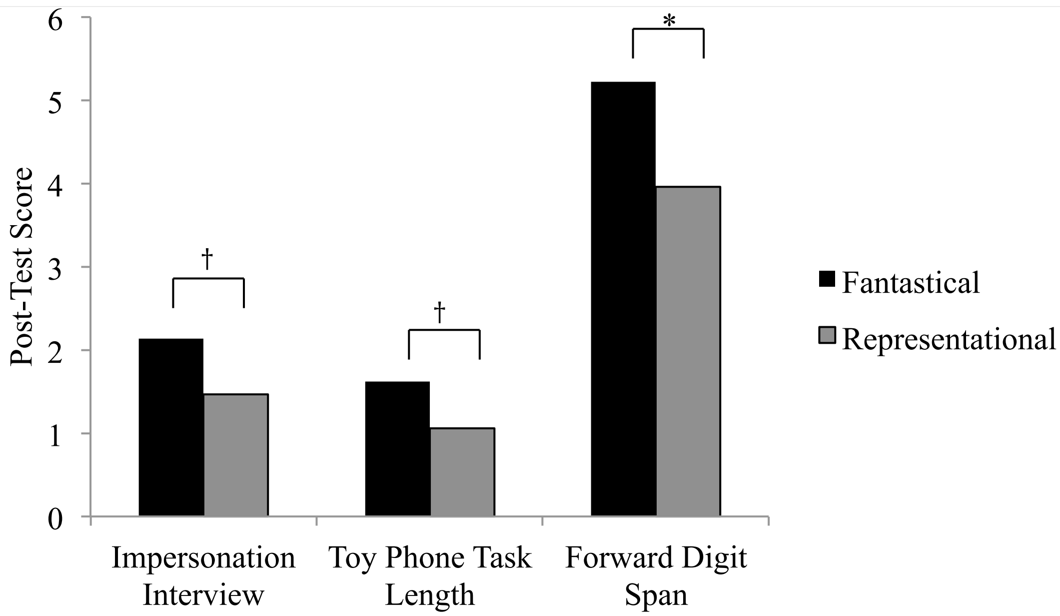

Figure 13 - FO and EF Differences – Fantastical vs. Representational. Differences between children in the fantasy-play group who consistently engaged fantastical play and children in the fantasy-play group who consistently engaged in representational play. *$p < .05$. † $p < .10$.

DISCUSSION

The purpose of the present study was to experimentally determine whether or not the repeated act of engaging in FO behaviors actually facilitates the development of EFs beyond what would be observed due to simple maturation or engaging in non-fantasy play. It was hypothesized that engaging in a five-week fantasy-play intervention would facilitate the development of EFs such that there would be a significant difference between intervention conditions (i.e., fantasy-play, realistic-play, and control) on post-test measures of EFs. Specifically, it was expected that the fantasy-play condition would show an increase in EF abilities beyond that of children in the realistic-play or no intervention conditions. In addition, if there were differences between the realistic-play and no intervention conditions, it was anticipated that the realistic-play condition would show greater increases in EF abilities than the no intervention condition because the act of engaging in play may inadvertently exercise cognition to some degree. In addition, FO was not hypothesized to change over the course of the intervention period because FO is hypothesized to be a stable trait, operationalized as an aspect of the openness personality construct (McCrae, 1987, 1993; Pierucci et al., 2013).

The results of the current study indicated meaningful differences between the fantasy-play condition and the realistic-play after the intervention, supporting the original hypotheses. Specifically, the fantasy-play group performed better than the realistic-play group on measures of working memory and teacher reported inhibitory control after the 5-week play intervention. These results were consistent even after considering how engaged a child was and how peer-oriented a child's play was throughout the intervention. Of note, the data indicated that performance actually improved from pre-intervention to post-intervention in the fantasy-play condition whereas performance did not change from pre-intervention to post-intervention in the realistic-play condition.

Because previous research indicates that the domains of EF are highly integrated in early childhood (Carlson et al., 2002; Diamond & Taylor, 1996; Hughes et al., 2010; Wiebe et al., 2008; Wiebe et al., 2011), it would be expected to see consistent post-test changes across all EF measures in the present study. However, it should be noted that we observed ceiling effects on several measures (i.e., Day/Night Task, Grass/Snow Task, Card Sort Task) preventing us from seeing any changes throughout the course of the intervention. With adequate power and more developmentally appropriate measures, it is expected that we would observe consistent improvements on all measures of EF in the fantasy-play condition as well as no changes on any EF measures in the realistic-play or control condition after the intervention. Thus, the data from the present study support the findings of previous literature that EFs are more unitary than distinct in early childhood.

The following results add to existing literature that points to a relationship between FO and EF in important ways. Specifically, Pierucci and colleagues (2013) found relations between FO and better cognitive inhibition skills. Our data helps to elucidate the directionality of this relationship. Specifically, we found that the act of engaging in FO play facilitates the development of inhibitory control. In addition, Pierucci et al. (2013) found better attention shift abilities among children who engaged in more FO cognitions (i.e., thinking about fantastical themes). Although we did not find any difference between groups on measures of attention shift, children were consistently performing at high levels on this measure pre- and post-intervention.

Therefore, it is possible that ceiling effects in our data prevented us from observing any group differences in this domain. We know of only one direct measure of children's attention shift in the literature, but often see ceiling effects with this measure. Future research should address the need for scalable measures of attention shift in childhood.

Interestingly, some of the results of the current study actually contradict findings in previous literature. Fortunately, these contradicting results offer additional explanations for previous findings that were once perplexing. For example, previous research has shown that interest in fantastical toys and games unexpectedly relates to poorer working memory performance (Pierucci et al., 2013). By contrast, data from the present study indicated that engaging in FO play actually enhances working memory performance. Previous studies that have found a negative relationship between FO and working memory performance have relied on a backward digit span task, which may not be developmentally appropriate for preschool aged children (Carlson, 2005; Pierucci et al., 2013). This task is thought to greatly tax EF resources in this age group (Carlson, 2005) contributing to the floor-effects obtained in Pierucci and colleagues (2013) data. By using a more developmentally appropriate forward digit span task, we observed that engaging in FO play actually facilitates working memory development.

Similarly, Pierucci and colleagues (2013) found that a preference for FO toys and games was related to poorer cognitive inhibition. However, our data indicate that engaging in FO play actually resulted in improved inhibitory control. The fundamental difference between the present study and results cited in previous literature is that in the current study, inhibitory control is related to behavioral differences in FO whereas Pierucci and colleagues (2013) report a relationship between inhibitory control FO *preferences*. In fact, Pierucci et al. (2013) highlight the need to determine if reported preferences for FO toys and games actually relates to engaging in fantastical cognitions and behaviors in everyday life. Given this information, it seems likely that there is a behavioral aspect of FO that is a necessary component for enhanced cognitive development.

In addition to highlighting the importance of engaging in FO behaviors, the results of the current study also emphasize that certain types of imaginative behaviors are more beneficial than others. Specifically, we found that children in the fantasy-play group who typically engaged in more fantastical play throughout the intervention (e.g., pretending to be fairies, pretending to play with dragons) scored better on several measures of FO/pretense and EF when compared to children in the fantasy-play group who typically engaged in more representational or sociodramatic play (e.g., pretending to be a mom, pretending to be an animal). Although all of these children were consistently engaging in imagination, there were clear differences in their post-test EF and FO scores when their imaginative behaviors involved non-realistic, or fantastical elements (e.g., princesses, aliens) than when their imaginative behaviors involved more realistic themes that they might experience in everyday life (e.g., pretending to be at a restaurant). Interestingly, these results are consistent with previous literature that shows that pretense (i.e., pretending to be an animal or another person) is not correlated with measures of EF (Pierucci et al., 2013). Therefore, it appears that the fantasy element to the fantasy-play intervention is what was driving the observed improvements in EF abilities.

The two control conditions used in the present study (i.e., realistic-play and control) were not designed to encourage pretense separately from fantastical play. Although the findings above provide great insight into what is driving the observed cognitive benefits in the fantasy-play condition, the results are correlational. Therefore, future experimental studies should include a fourth condition that specifically encourages imaginative play that is sociodramatic and representational, but does not have a fantasy component.

In addition to finding that there are behavioral aspects of FO that are necessary components for enhanced cognitive development, the present study also builds upon previous literature by clarifying the directionality of the observed relationship between FO and EF. For example, Carlson and colleagues (2014) found a relationship between EFs and the understanding of pretend versus reality as well as the ability to perform various pretend actions. However, due to the nature of their study, these authors were not able to determine if the ability to engage in these pretend behaviors facilitates the development of EF, or whether having more advanced EF provides individuals with the capacities needed to engage in pretend behaviors. Because Carlson et al. (2014) did not have any direct measure of *observed* pretend play, these authors assumed the latter, suggesting that EF skills provide individuals with the aptitude for engaging in pretense. The data of the present study indicates that the former is true, such that engaging in FO behaviors facilitates the development of EF skills. In fact, our fantasy-play condition was made up of individuals with varying levels of EF skills prior to the intervention. The fact that we observed improvements in many domains of EF (i.e., working memory and inhibitory control) in the fantasy-play condition after engaging in five weeks of FO play not only indicates that FO play facilitates EF development, but also contradicts the idea that a certain level of EF skill is necessary in order to engage in FO behaviors.

In addition to finding post-intervention differences in EF across the three groups of the present study, the results also revealed consistent differences between groups on one measure of pretense (i.e., Toy Phone Task). Specifically, our data indicate that individuals in the fantasy-play condition scored higher on the Toy Phone Task after the intervention than the realistic-play group and control group. It should be noted that we did not anticipate any changes in FO after the intervention, as FO appears to be a stable trait that is unlikely to change overtime (Pierucci et al., 2013). However, finding that the fantasy-play group scored higher on the Toy Phone Task measure post-intervention does not necessarily contradict the original hypothesis. Although we listed the Toy Phone Task as a measure of FO/pretense, no aspects of the task are fantastical in nature. Rather, this task, which simply asks children to pretend to call a friend they like to play with, is strongly based in reality and therefore is better classified as a measure of pretense than a measure of FO. The fact that the fantasy-play group scored higher on this task after the intervention serves as a manipulation check indicating that we did increase pretending behaviors in this group throughout the course of the intervention period. By contrast, pretending behaviors did not increase in the realistic-play or control conditions providing evidence of the uniqueness of the games played in the fantasy-play conditions.

Previous literature has laid out three possibilities for the influence of pretend play on cognitive development (Lillard et al., 2013). Firstly, Vygotsky (1978) would argue that pretend play is a *crucial* element to cognitive development. In other words, pretend play is a necessary component through which cognitive abilities develop. The second possibility is that pretend play is *equifinal* (cf. Smith, 2010). According to this view, pretend play helps to promote cognitive development, but it is not the only route through which cognitive abilities can develop. Finally, Piaget (1962) suggests that pretend play may be *epiphenomenal* such that it simply a byproduct of some other factor that is actually linked to cognitive development. According to this particular theoretical perspective, pretend play does not make any direct contributions to cognitive development. Rather, pretend play simply coincides with some other factor that is directly related to cognitive development. In such cases, pretend play may be mistakenly thought to cause the development of cognitive abilities.

Up to this point, no studies have been able to determine which theoretical perspective best describes the relationship between pretend play and cognitive development (Lillard et al., 2013). In fact, Lillard and colleagues (2013) discuss several major shortcomings of prior studies on pretend play that have prevented us from determining which theoretical perspective is most appropriate. First of all, many of the studies previously conducted on pretend play and cognitive development are correlational in nature. Thus we have not been able to make conclusions about the causality and directionality of the relationships observed. Secondly, many previous studies did not control for experimenter bias by making sure that the experimenters were unaware of the hypotheses of the study and blind to the condition of each child. Finally, studies that have shown increases in cognitive and social development after pretend play training have failed to replicate their results.

Fortunately, the current study addresses many of these limitations. Specifically, the presented study utilized a randomized control experimental design allowing us to make conclusions about the causality and directionality of the relationship between FO and EF. In addition, several steps were taken to minimize experimenter bias in the present study. All experimenters and teachers were blind to the hypotheses of the study as well as the condition of each child. Because we have addressed these limitations, we are able to contribute to the debate on which theoretical perspective is best based on the results from the present study.

Since the present study found a direct causal link between FO play and improvements in EF, we can rule out the epiphenomenal view point because we found that pretend play appears to make direct contributions to cognitive development. However, without replicating the results of our study and including other control conditions (e.g., other EF training mechanisms), we cannot conclude that FO play is crucial to EF development. In other words, as prior research demonstrates, there are likely additional mechanisms through which EF can be scaffolded (e.g., bilingualism; Bialystok, 1999; Bialystok & Craik, 2010; Carlson & Melzoff, 2008). Therefore, our data suggests that the relationship between FO play and EF development is equifinal such that engaging in FO play is one of many ways to directly enhance EF development.

Although the present study helps to elucidate discrepancies in theoretical perspectives, it is important to address several potential limitations of the current study. Of first concern is the nature of the sample used in the present study. Specifically, children who participated in the current study come from predominantly middle-class households and varied little in ethnicity. Therefore, our results are not generalizable outside of this narrow sample. Future studies should incorporate children living in at risk populations who exhibit serious deficits in cognitive, attentional, and behavioral control. Because our data indicate that engaging in FO play creates an enriched environment through which EFs can be scaffolded, it is anticipated that children in at-risk environments would show even greater gains in EF after participating in a FO play intervention. In addition, the present study had a relatively small sample size resulting in under-powered results. Small sample sizes are a recurring limitation in play intervention research (Lillard et al., 2013) and should be addressed in future replication studies. It should also be noted that some of the EF measures used in the present study (i.e., Card Sort Task, Day/Night Task, Grass/Snow Task) were not sufficiently difficult enough for children in our sample, as the majority of participants made few errors on these measures resulting in ceiling effects at pre- and post-test. These ceiling effects may have prevented us from observing any changes in these measures from pre- to post-intervention. Thus, using more robust and developmentally appropriate measures in future studies may reveal even more benefits of FO play on EF development. Unfortunately, there are very few direct assessment measures of EF for children in the literature. Researchers are currently working on creating and validating new measures.

Despite these limitations, the present study has several important implications. To begin, the present research is the first to experimentally investigate the relationship between FO and EFs, and thus provides important insights into the causality and directionality of this relationship that were previously lacking. In finding that FO facilitates EF development, the present study has the unique ability to stimulate a multitude of future research. Specifically, results from the current study could advance a research line investigating fantasy-play curriculum development and dissemination. In addition, the present study could lead to longitudinal studies of the effects of engaging in FO play on children's cognitive development.

In conclusion, a growing body of evidence indicates the importance of EF development throughout childhood and adolescence as EFs lay the foundation for cognitive and social development (Blair & Razza, 2007; Moffitt et al., 2011). Unfortunately, EF abilities do not fully develop via brain maturation alone, but rather require rich environmental experiences, putting some individuals, such as children raised in poverty with no preschool preparation for kindergarten, at risk for underdevelopment of EFs at an early age. Therefore, it is imperative to identify environmental experiences that facilitate the development of EFs early in life in order to ensure that every child has an equal opportunity to become a successful and contributing member of society. The present research has identified fantastical pretend play, the most ubiquitous natural environmental experience in childhood, as a facilitator of EFs and cognitive control. The results of this study indicate that encouraging fantasy play in a child's everyday life will facilitate development in many domains. These findings have the potential to improve school-readiness in preschool children, and therefore close achievement gaps throughout the nation. Because fantasy play is easily implemented into existing classroom curriculums at little to no cost, the findings from the present study will have even broader impacts on curriculum development, especially in at-risk populations such as Head Start. Children who are well prepared for school are much more likely to be occupationally successful. Indeed, a conservative estimate for the return on investment for preparing children for school is 7:1 based on several longitudinal school-readiness interventions such as the Chicago and High/Scope Perry Preschool Projects (Bruner, 2004). Thus, legislators and education specialists will be very interested in and benefit greatly from learning about the effects of a low-cost intervention involving fantastical play to prepare children for school success.

REFERENCES

Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London Series B, Biological Science, 302*, 311-324.

Baddeley, A. (1992). Working memory. *Science, 255*, 556–559. doi:10.1126/science.1736359

Bergman Nutley, S., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: A controlled, randomized study. *Developmental science, 14*(3), 591–601. doi:10.1111/j.1467-7687.2010.01022.x

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70,* 636-644.

Bialystok, E. (2011). Coordination of executive functions in monolingual and bilingual children. *Journal of Experimental Child Psychology, 110*, 461–468.

Bialystok, E. & Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science, 19,* 19-23.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. J*ournal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 859–873. doi: 10.1037/0278-7393.34.4.859

Bialystok, E., & Martin, M.M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*, 325–339.

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*, 647-663.

Bodrova, E., & Leong, D. J. (2007). *Tools of the Mind: The Vygotskian approach to early chidhood education.* New York: Merrill/Prentice Hall.

Boerger, E. B., Tullos, A., & Woolley, J. D. (2009). Return of the Candy Witch: Individual differences in acceptance and stability of belief in a novel fantastical being. *British Journal of Developmental Psychology, 27*, 953–970. doi:10.1348/026151008X398557

Bruner, C. (2004). Many happy returns: Three economic models that make the case for school readiness. Retrieved from http://www.finebynine.org/uploaded/file/SECPTAN_MHR_final.pdf

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28,* 595-616.

Carlson, S. A., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*, 282–298. doi:10.1111/j.1467-7687.2008.00675.x.

Carlson, S. M. & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032-1053.

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. *Infance and Child Development, 11,* 73-92. doi:10.1002/icd.298

Carlson, S. M., White, R. E., & Davis-Unger, A. C. (2014). Evidence for a relation between executive function and pretense representation in preschool children. *Cognitive Development, 29,* 1-16.

Carlson, S. M., & Zelazo, P. D. (2008). Symbolic throught. In M. M. Haith, & J. B. Benson (Eds.), *Encyclopedia of infant and early childhood developent* (Vol. 3) (pp. 288-297). London: Elsevier.

Center on the Developing Child at Harvard University (2011). *Building the Brain's "Air Traffic Control" System: How Early Experiences Shape the Development of Executive Function: Working Paper No. 11.* Retrieved from www.developingchild.harvard.edu

Davis, H. L., & Pratt, C. (1996). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*, 25–31.

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology, 38,* 352-362. doi:10.1037//0012-1649.38.3.352

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4–12 years old. *Science, 333*(6045), 959–964. doi:10.1126/science.1204529

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "do as I say not as I do." *Developmental Psychology, 29,* 315-334.

Dunn, J. R. (2010). Healh beahvior vs the stress of low socioeconomic status and health outcomes. *Journal of the American Medical Association, 303,* 1199-1200. doi:10.1001/jama.2010.332

Dunn, L. M. & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test- Fourth Edition*. Minneapolis, MN: NCS Pearson, Inc.

Eakin, L., Minde, K., Hechtman, L., Ochs, E., Krane, E., Bouffard, R., …Looper, K. (2004). The marital and family functioning of adults with ADHD and their spouses. *Journal of Attention Disorders, 8,* 1-10.

Estes, D., Wellman, H. M., & Woolley, J. (1989). Children's understanding of mental phenomena. In H. Reese (Ed.), Advances in Child Development and Behavior (pp. 41–86). New York, NY: Academic Press.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10,* 483–527. doi: 10.1016/0885-2014(95)90024-1.

Gerstadt, C.L., Hong, Y.J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 ½ - 7 years old on a Stroop like day-night test. *Cognition, 53,* 129-153.

Gilpin, A. T. (2009). *Mechanisms for overcoming reality status biases* (Unpublished doctoral dissertation). University of Texas, Austin.

Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & Espy, K. A. (2002). Confirmatory factor analysis of the behavior rating inventory of executive function (BRIEF) in a clinical sample. *Child Neuropsychology.Special Issue: Behavior Rating Inventory of Executive Function (BRIEF), 8*(4), 249-257. doi:10.1076/chin.8.4.249.13513

Golumb, C., & Kuersten, R. (1996). On the transition from pretense play to reality: What are the rules of the game? *British Journal of Developmental Psychology, 14*, 203–217. doi:10.1111/j.2044-835X.1996.tb00702.x

Halford, G. S., Maybery, M. T., & Bain, J. D. (1988). Set-size effects in primary memory: An age-related capacity limitation. *Memory and Cognition, 16,* 480–487. doi:10.3758/BF03214229.

Henry, L. A. (2001). How does the severity of a learning disability affect working memory performance? *Memory, 9,* 233-247.

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2010). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology, 35,* 20-36. doi:10.1080/87565640903325691

Johnson, M.H. (1993). Cortical maturation and the development of visual attention in early infancy. In M. H., Johnson (Ed.), *Brain development and cognition: A Reader* (pp. 167-193). Oxford: Blackwell.

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2013). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin, 139*(1), 1–34. doi:10.1037/a0029321

McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, *52*, 1258–1265.

McCrae, R. R. (1993). Openness to experience as a basic dimension of personality. *Imagination, Cognition, and Personality, 13*, 39–55.

Miyake, A., Friedman, N., Emerson, M., Witzki, A., 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*, 49–100.

Miyake, A., & Freidman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*, 8-14. doi:10.1177/0963721411429458

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), 2693–8. doi:10.1073/pnas.1010076108

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp. 93-148). Hove, UK: Erlbaum.

Piaget, J. (1962). Play, dreams, and imitation in childhood (G. Gattegno & F. M. Hodgson, Trans.). New York, NY: Norton.

Pierucci, J. M., O'Brien, C. T., McInnis, M. A., Gilpin, A. T., & Barber, A. B. (2013). Fantasy-oriented constructs and related executive function-development in preschool: Developmental benefits to executive functions by being a fantasy-oriented child. *International Journal of Behavioral Development*, 1-8. doi:10.1177/0165025413508512

Sharon, T., & Wooley, J. D. (2004). Do monsters dream? Young children's understanding of the fantasy/reality distinction. *British Journal of Developmental Psychology, 22,* 293-310.

Singer, D. G., & Singer, J. L. (1981). *Television, imagination, and aggression: A study of preschoolers.* Hillsdale, NJ: Erlbaum.

Smith, P. K. (2010). *Children and play.* West Sussex, England: Wiley- Blackwell.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59*(4), 745–59. doi:10.1080/17470210500162854

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Tahiroglu, D., Mannering, A.M., & Taylor, M. (2011). Visual and auditory imagery associated with children's imaginary companions. *Imagination, Cognition and Personality, 31,* 99-112. doi:10.2190/IC.31.1-2.i

Taylor, M. (1999). *Imaginary companions and the children who create them.* New York, NY: Oxford University Press.

Taylor, M., & Carlson, S. M. (1997). The relation between individual differences in fantasy and theory of mind. *Child Development, 68*, 436–455. doi:10.1111/j.1467-8624.1997.tb01950.x

Taylor, M., Cartwright, B. S., & Carlson, S. M. (1993). A developmental investigation of children's imaginary companions. *Developmental Psychology, 29,* 276–285. doi:10.1037/0012-1649.29.2.276.

Taylor, M., Sachet, A.B., Maring, B.L., & Mannering, A.M. (2013). The assessment of elaborated role-play in young children: Invisible friends, personified obejcts, and pretend identities. *Social Development, 22,* 75-93. doi:10.1111/sode.12011

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12*(1), 106–13. doi:10.1111/j.1467-7687.2008.00745.x

Vygotsky, L.S. (1967). Play and its role in the mental development of the child. *Soviet Psychology, 5,* 6-18.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory analyses to understand executive control in preschool children: I. Latent structure. *Developmental Psychology, 44,* 575–587.

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

Woolley, J. D., Boerger, E. A., & Markman, A. B. (2004). A visit from the Candy Witch: Factors influencing young children's belief in a novel fantastical being. *Developmental Science, 7*(4), 456–468. doi:10.1111/j.1467-7687.2004.00366.x.

Woolley, J. D., & Tullos, A. (2008). Imagination and fantasy. In M. Haith & J. Benson (Eds.), *Encyclopedia of Infant and Early Childhood Development* (pp. 117–127). Oxford, UK: Elsevier Press.

Woolley, J. D., & Wellman, H. M. (1990). Young children's understanding of realities, nonrealities, and appearances. *Child Development, 61*, 946–961. doi:10.2307/1130867

Wright, I., Waterman, M., Prescott, H., & Murdoch-Eaton, D. (2003). A new stroop-like measure of inhibitory function development: Typical developmental trends. *Journal of Child Psychology and Psychiatry, 44*, 561–575. doi:10.1111/1469-7610.00145

Zelazo, P. D. (2006). The dimensions change card sort (DCCS): A method of assessing exectuive function in children. *Nature Protocols, 1,* 297-301. doi:10.1038/nprot.2006.46

Zelazo, P. D., Muller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68*(3, Serial No. 274).

APPENDIX

Appendix A

Fantasy Orientation Coding

Measure	Code
Parent/Teacher Questionnaires	
Believes in Fantastical Figures	0 = child knows fantastical figure is pretend .5 = child does not know if fantastical figure is real or pretend 1 = child believes fantastical figure is real
Favorite Books, Games, TV Shows, Videogames	0 = reality based 1 = low fantasy based 2 = high fantasy based
Child's Level of FO	1 = strong interest in reality (e.g., plays sports) 2 = some interest in fantasy but mostly interested in reality 3 = equal interests in fantasy and reality play/media 4 = mostly interested in fantasy but some interest in reality 5 = strong interest in fantasy (e.g., often engages in pretense, enjoys fantastical books)
Child FO Measures	
Imaginative Play Predisposition	0 = Reality-oriented response 1 = Representational/low fantasy response 2 = High fantasy response
Imaginary Companion Interview	0 = Does not have imaginary companion 1 = Has an imaginary companion
Impersonation Activities	0 = Does not pretend to be an animal/person/something else 1 = Does pretend to be an animal/person/something else

Appendix B

Intervention Coding

Aspect of Behavior being Assessed	Code
Engagement	
"How engaged was this child during the play session today?"	0 = Not Engaged 1 = Somewhat Engaged 2 = Very Engaged
Peer-Oriented vs. Solitary	
"How solitary was this child's play?"	0 = Mostly played by themselves 1 = Interacted with other kids and played by themselves 2 = Interacted with other kids for most of the play session
Style	
"How fantastical/realistic was this child's play today?"	0 = Very reality based (no imagination involved) 1 = Somewhat fantastical, somewhat reality based (imaginative but included more realistic imagination like pretending to be a mom or an animal) 2 = Very fantastical (included entities like dragons, aliens, princesses, etc. in play)

Appendix C

Table of FO Measures

Measure	Description	Code	Reliability	Validity
		Teacher FO Questionnaire		
FO Questionnaire	7 questions assessing child's belief in fantastical figures	0 = knows it's pretend .5 = does not know if real/pretend 1 = believes it's real	Unknown	Unknown
	2 questions assessing child's favorite books, games, TV shows, videogames	0 = reality based 1 = low fantasy based 2 = high fantasy based		
	Child's level of FO	1 = strong interest in reality 2 = some interest in fantasy but mostly interested in reality 3 = equal interests in fantasy and reality play/media 4 = mostly interested in fantasy but some interest in reality 5 = strong interest in fantasy		
	Imaginary companion	0 = No imaginary companion 1 = Has an imaginary companion		
		Child FO Measures		
Imaginative Play Predisposition	4 questions assessing favorite game, toy, story, TV show	0 = Reality-oriented response 1 = Representational/low fantasy 2 = High fantasy response	Inter-rater reliability = 73% to 93% (Sharon & Wooley, 2004; Taylor & Carlson, 1997)	Unknown
	3 questions assessing what child likes to do by themselves/with their friends and what they think about before going to sleep	0 = Reality-oriented response 1 = Representational/low fantasy response 2 = High fantasy response		
Imaginary Companion Interview		0 = No imaginary companion 1 = Has an imaginary companion	Inter-rater reliability = 73% to 93% (Sharon & Wooley, 2004; Taylor & Carlson, 1997)	Unknown
Impersonation Activities	3 questions assessing if child ever pretends to be an animal, person, or something else	0 = Does not pretend to be 1 = Does pretend to be	Inter-rater reliability = 73% to 93% (Sharon & Wooley, 2004; Taylor & Carlson, 1997)	Unknown

Appendix D

Table of EF Measures

Measure	EF Assessed	Description	How Measure is Scored	Reliability	Validity
Teacher EF Questionnaire					
BRIEF-P	Global measure of EF	63 questions of inhibition, attention shift, emotional control, & working memory	Scored according to standardized BRIEF-P manual	Alpha = .80-.98 Retest = .81-.88	High construct validity (Gioia et al., 2002)
Child EF Measures					
Forward Digit Span	Working Memory	Child instructed to repeat a string of digits	Highest number of digits successfully repeated	Retest = .78 (Henry, 2001)	Convergent validity with other working memory measures (Henry, 2001)
Animal Stroop Task	Cognitive Inhibition	Participants shown pictures where animal's face does not match its body (e.g., pig head on a cow body) and must name the animal by its body	Number of errors/reaction times	Alpha = .56-.93 (Wright et al., 2003)	Sensitive to development and inhibitory control differences (Wright et al., 2003)
Day/Night Task	Cognitive Inhibition	Participants say "day" when the experimenter points to the picture of night and say "night" when the experimenter points to the picture of day	Number of correct and incorrect responses	Inter-rater reliability = .90 Alpha = .56-.93 (Wright et al, 2003, Carlson & Moses, 2001; Gerstadt et al.,1994)	Correlated with other established measures of inhibitory control (Carlson & Moses, 2001)
Standard Dimensional Change Card Sort Task	Attentional Shift	Sort the cards by shape. After correctly sorting five cards consecutively, participants instructed to sort by color. Order counterbalanced	Number of errors in first game, number of errors in the second game, and average number of error in both games	Inter-rater reliability = 100% (Carlson & Moses, 2001)	Correlated with other established measures of inhibition (Carlson & Moses, 2001)
Word Span Task	Working Memory	Participants instructed to repeat the words said by the experimenter	0 = omits a word, substitutes a word, or inserts a word 1 = repeats the correct words in an incorrect order 2 = no errors	Inter-rater reliability = 91%-100% (Edwards, 2006)	Convergent validity with other measures of short-term memory (Edwards, 2006)
Grass/Snow Task	Cognitive Inhibition	Participants instructed to point to the grassy field when the experimenter says "snow", and	Number of correct and incorrect responses	Alpha = .56-.93 (Wright et al, 2003, Carlson & Moses, 2001; Gerstadt et	Correlated with other established measures of

Appendix E

Demographics Data

	Fantasy-Play	Realistic-Play	Control	Total
Gender				
Males	49% (n=19)	52% (n=17)	44% (n=18)	48% (n=54)
Females	51% (n=20)	48% (n=16)	56% (n=23)	52% (n=59)
Age	3.82	4.00	3.88	3.90
Vocabulary	106.74	110.3	108.97	108.67
Ethnicity				
Caucasian	95% (n=37)	88% (n=29)	85% (n=35)	90% (n=10)
African-American	2.5% (n=1)	6% (n=2)	10% (n=4)	6% (n=7)
Other	2.5% (n=1)	6% (n=2)	5% (n=2)	4% (n=5)
Parent Education	M=4-year college or bachelors degree	M=4-year college or bachelors degree	M=4-year college or bachelors degree	M=4-year college or bachelors degree
Income	M=$90,000-$99,999	M=$80,000-$89,999	M=$90,000-$99,999	M=$80,000-$89,999

Office for Research

Institutional Review Board for the
Protection of Human Subjects

THE UNIVERSITY OF
ALABAMA
RESEARCH

March 3, 2014

Ansley Gilpin, PhD
Dept of Psychology
College of Arts and Sciences
Box 870348

Re: IRB#: 14-OR-060 "Effects of Fantasy-Oriented Sociodramatic Play on the Development of Executive Functions"

Dear Dr. Gilpin:

The University of Alabama Institutional Review Board has granted approval for your proposed research.

Your application has been given expedited approval according to 45 CFR part 46. Approval has been given under expedited review category 7 as outlined below:

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies

Your application will expire on February 26, 2015. If your research will continue beyond this date, complete the relevant portions of the IRB Renewal Application. If you wish to modify the application, complete the Modification of an Approved Protocol Form. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, complete the appropriate portions of the IRB Request for Study Closure Form.

Please use reproductions of the IRB approved stamped consent/assent forms to obtain consent from your participants.

Should you need to submit any further correspondence regarding this proposal, please include the above application number.

Good luck with your research.

358 Rose Administration Building
Box 870127
Tuscaloosa, Alabama 35487-0127
(205) 348-8461
FAX (205) 348-7189
TOLL FREE (877) 820-3066