# What Good is Pretending? Adding a Pretense Context to the Dimensional Change Card Sort

Tamara Spiewak Toub

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Betty M. Repacholi, Chair

Stephanie M. Carlson

Andrew N. Meltzoff

Program Authorized to Offer Degree:

Department of Psychology

University of Washington

## Abstract

What Good is Pretending? Adding a Pretense Context to the Dimensional Change Card Sort

Tamara Spiewak Toub

Chair of the Supervisory Committee:
Associate Professor Betty M. Repacholi
Department of Psychology

During the preschool years, children's pretend play increases in frequency and sophistication. Besides the joy obtained from these experiences, there might be benefits regarding children's social, emotional, and cognitive development. Prior literature shows both theoretical and empirical support for a relation between pretense and executive function (EF), in particular. Given the role of EF skills in childhood and beyond, this research focused specifically on the potential for pretense-based strategies to facilitate better EF performance for preschoolers. The primary aim was to investigate the effect of adding a pretense context to the Dimensional Change Card Sort (DCCS), a conflict-based EF task (Zelazo, 2006). In Study 1, 96 preschoolers (ages 40-47 months) were assigned to either a Fantasy Planet condition or a modified Standard condition, which did not involve pretense. In the Fantasy Planet condition, the experimenter encouraged children to pretend to travel to imaginary planets (Planet Shape and Planet Color) during the shape and color phases of the DCCS. Even after controlling for age, boys in the

Fantasy Planet condition were better able to make the switch from the shape rule to the color rule (or vice versa) than were boys in the modified Standard condition. This facilitative effect of the manipulation was not found for girls. Exploratory examination of data from secondary tasks did not explain these results. Study 2 was then designed to address two main questions: (1) What was it about the Fantasy Planet condition that facilitated better performance (for boys)? and (2) Why was this effect limited to boys? For this study, 64 preschoolers (ages 43-46 months) were assigned to either a Pretend Playground condition or a Non-Pretense Control condition. These conditions were designed to address specific alternative explanations from Study 1 and to incorporate a more gender-neutral theme for the pretense-based condition. Results from Study 2 showed no differences in DCCS performance based on sex, condition, or their interaction. Possible explanations for the set of the two studies' results are discussed. Among the suggestions for further research is the incorporation of a manipulation that involves a deeper pretense experience than what was adopted in these studies.

## Table of Contents

**List of Figures**

## List of Tables

**Acknowledgements**

This research would not have been possible without the generosity of the many participating families of preschoolers. I am deeply grateful for their time, energy, and interest.

In this dissertation process and throughout my graduate school career, I have been extremely fortunate to have Dr. Betty Repacholi as my primary advisor. Her mentorship has been instrumental in my development as a researcher, and I thank her for her attentiveness, respect, encouragement, and guidance. I have also been honored to work with Dr. Stephanie Carlson, who has remained a dedicated mentor despite her move away from UW. I am grateful for her expertise and support. My committee includes two other faculty whose work and input I value: Dr. Andrew Meltzoff, who contributed significant research design suggestions, and Dr. J. David Hawkins from the School of Social Work, who graciously serves as my GSR.

I have relied upon a team of hard-working research assistants for much of the behind-the-scenes coding and data entry. For their strong contributions to this dissertation work, I especially thank: Prysilla De La Torre, Travis Egan, Christina Gangan, Dan Kubitz, Shelley Leong, Tori Ly, and Niveda Shekar. I also very much appreciate the efforts of research assistants who worked on portions of the project that are less featured here: Sonia Giebel, Jeanna Koch, Ellen Lesh, Lisa Shelton, Youngmee Sul, and Greg Wyant, III.

Jennifer McBride provided invaluable assistance regarding participant recruitment through the CSPP, and I thank her immensely for all her efforts. I am also grateful for the assistance I received on a myriad of issues from other staff, faculty, and students, including: Rich Ball, Phil Burger, Margaret Cheng, Dr. Brian Flaherty, Jon Hauser, Michele Jacobs, Merly Jones, Doug Kalk, Jeanny Mai, Dr. Michael Passer, Justin Reichel, Marina Ryzhova, Catherine Schaefer, Kelly Scott, Dr. Jessica Sommerville, Julie Vaisarova, and CSSS staff.

This work was supported by the Bolles Fellowship and the Earl (Buz) and Mary Lou Hunt Endowed Fellowship awarded to the author through the UW Department of Psychology and by an NICHD grant (5R01HD051495-06) awarded to Dr. Stephanie Carlson.

Finally, I thank my family and friends for all their encouragement, patience, and understanding. In particular, my husband, Stephen Toub, has supported me in many, many ways, big and small. He listened to my detailed tales of both progress and disappointment, helped with random technical matters, assured me on a regular basis that I could do this, and sought out good prices on Play-Doh 4-packs (gifts for participants) whenever he was at the store.

**Dedication**

To my father Bruce and my late mother Lauren --
who instilled in me a hard work ethic, as well as an appreciation for the arts and for playing,
and who have encouraged me in each of my endeavors.

And to my late Grandpa Jerry --
who eagerly awaited the day he could call me "Doctor."

# Chapter 1: Introduction

If you spend a day observing a classroom of preschoolers playing freely, you are likely to see some children engaging in pretend play. They might pretend to be someone else, that a doll is a real baby, or that a block of wood is a car. Children's pretend play increases in both frequency and sophistication during the preschool years (e.g., Berk, Mann, & Ogan, 2006; Carlson & Beck, 2009; Fein, 1981; Haight, Black, Jacobsen, & Sheridan, 2006; Lillard, 1994; Piaget, 1962; Rubin, Fein, & Vandenberg, 1983; Vygotsky, 1933/1978). Aside from the pure pleasure derived from pretending, what might preschoolers gain from such activities? In addition, how might pretense be used proactively to promote children's healthy development?

Most broadly, this research program aims to investigate potential developmental benefits of pretense for preschoolers. In the following review of the literature, I will first explain what pretense involves and its developmental trends during childhood. Then, I will summarize some of the associations that have been made between pretense and children's emotional, social, and cognitive development. The current research focuses specifically on the relation between pretense and executive function (EF). By way of introduction, I will briefly explain the construct of EF and then present theoretical and empirical reasons to believe that there is a relation between pretense and EF. Finally, I will provide the hypotheses for Study 1 of this research program. After this introduction, I will describe Study 1 and then introduce and describe Study 2. With these two studies, I explore what happens to preschoolers' performance on a traditional and challenging EF task when it is presented within a pretense-based context.

## Pretend Play

**Defining pretend play.** A common simple definition of pretend play in the literature focuses on behavior that treats some element of the current reality in a non-literal (or "as if") fashion (e.g., Fein, 1981; Youngblade & Dunn, 1995). When discussing and studying pretend play, however, it could be helpful to define a more precise set of conditions that must be met for behavior to qualify. When Piaget (1962) described the initial differentiation of make-believe play from other types of play, he specified two requirements: that there be assimilation of a current object or situation into a schema originally regarding a different object or situation (i.e., a distinction between signifier and signified) and that it be engaged in purely for pleasure (not for adaptive learning or curiosity). More contemporary researchers have discussed similar elements and added others. For example, Leslie (1987) indicated that behavior qualifies as pretense if and only if it involves one of the following: object substitution (an object is used to stand for a different object), attribution of pretend properties to an object or situation, or imaginary objects. As acknowledged by Kelly and Hammond (2011), another possible type of pretending that is often added to Leslie's list is role-play. In addition to this focus on the manner of pretending, Leslie also emphasized that the pretender must know about the reality, distinguishable from the pretense.

Focusing less on the form of pretense and more on issues such as the pretender knowing about reality, Lillard (1994, 2001) enumerated six essential elements of pretend: (1) an animate pretender, (2) a reality about which to pretend, (3) a mental representation (usually different from the reality) guiding the pretense, (4) projection of pretense (the mental representation) onto reality, (5) awareness of the contents of real and pretend representations and their relation, and (6) intention to project the pretend onto the real. She argued that pretense often involves elements of non-seriousness and activity but that those are not essential (Lillard, 1994). Emphasizing the freedom from standard constraints involved in the pretense experience, Fein (1987) identified elements of *sequential uncertainty* (pretend sequences are nonlinear and improvisational), *affective relations* (emotional and often bizarre topics), and *self-mirroring* (consciously reflecting on oneself from a removed perspective), in addition to *denotative license* (freedom to use any object to substitute for any other) and *referential freedom* (freedom from the immediate environment), which are similar to elements discussed already.

Considering these sets of features collectively, behavior qualifies as pretend play if it meets the requirements regarding form (e.g., Leslie, 1987), mental representation and intention (e.g., Lillard, 1994, 2001), and style (e.g., Fein, 1987).

**Developmental trajectory of pretend play.** Children engage in pretend gestures starting around 12 months of age, and pretense grows more sophisticated into the preschool years and beyond (e.g., Fein, 1981). Although pretense is initially self-focused, there is a progression from self to other between the ages of 12 and 30 months (and beyond), with the child increasingly making the "other" an active participant in the pretense (e.g., by making a doll talk) (e.g., Fein, 1981; Lillard et al., 2007; Piaget, 1962; Rubin et al., 1983). Another developmental progression involves children's increasing sophistication regarding object substitutions (e.g., Berk et al., 2006; Fein, 1981; Lillard et al., 2007; Rubin et al., 1983; Vygotsky, 1933/1978). One main idea of this trend is that younger children require higher degrees of similarity between the object being used and that which it is representing, while older children are able to use objects as substitutes regardless of their resemblance. Eventually, children are able to create entirely imaginary objects. For example, Overton and Jackson (1973) asked children to perform pretend actions (e.g., brushing teeth) in the absence of any props. In this research, younger children tended to use their own body parts (e.g., finger) to substitute for the pretend objects (e.g., toothbrush), and their use of an imaginary object increased significantly between each of the age groups (3, 4, 6, and 8 years).

Children's play also becomes more social, rather than solitary, starting around age 3 years (Fein, 1981). In a longitudinal observational study, Howes and Matheson (1992) found that 79% of participants first engaged in cooperative social pretense (i.e., partners adopt complementary roles) by 30 to 35 months old, and 67% first engaged in complex social pretense (i.e., partners adopt roles and discuss the pretend roles and scenario) by 42 to 47 months old. This study illustrates the development of meta-communication activities that denote sophisticated pretense.

During social pretense, children must respond appropriately to their partner's behaviors, and there are developmental changes in these abilities, as well. At 15 months old (and more so at 24 months), infants have shown engagement in their mothers' enactments of pretend scenarios (i.e., by smiling and making related pretense gestures) (Lillard et al., 2007). As emphasized by Harris (1994, 2000), by 2 years old, children can recognize partners' pretend stipulations and respond appropriately. By this, Harris meant that children treat objects as their pretend identities, generalize to similar objects within the pretense context without over-generalizing, and build upon the stipulation with a reasonable response to the pretend action or to an effect the action could cause (e.g., pretend to wipe up liquid after a pretend cup of milk was overturned).

Thus, at 3 or 4 years old, preschoolers tend to have an established repertoire for pretend play, and they engage in pretense with increasing sophistication and frequency. Although pretend play is not the most dominant activity in childhood, engagement in pretense is common, with pretense during free-play at average rates as high as 20% of the observed free-play period (Field, Destefano, & Koewler, 1982). In addition, researchers have found that 13-65% of children (depending on children's age and operational definitions) have had imaginary companions (Gleason, 2005; Taylor, 1999; Taylor & Carlson, 1997; Taylor, Carlson, Maring, Gerow, & Charley, 2004). Pretending is a substantial part of early childhood.

**Adaptive benefits of pretend play.** Given that children engage in pretend play, why might they do so? Despite differing opinions regarding whether or not pleasure is a primary feature of pretend play (e.g., Lillard, 1994; Piaget, 1962; Vygotsky, 1933/1978), the pleasure gained from engagement in pretense has been proffered as a driving reason for children's involvement (e.g., Piaget, 1962; Taylor, 1999). Pretend play might have been selected by evolution, however, because of its function in facilitating children's development (Smith, 2002), rather than purely for pleasure's sake. There are areas of emotional, social, and cognitive development that might benefit from pretend play, making it an adaptive activity. Some benefits have been proposed on theoretical levels and others on empirical grounds, based on correlations between pretend play and these particular skills. Many of the correlates of pretend play share a common theme of decentration (involving children's developing understanding that there are many, interrelated dimensions to objects and situations) (Rubin et al., 1983). In a recent review of the literature, Lillard and colleagues (2012) discussed the state of the field regarding developmental benefits of pretending. They argued that there is not sufficient evidence to claim that pretending plays a crucial role in development and that further research must be conducted to distinguish between instances in which pretending is one of many possible paths to the development of a given outcome and instances in which it is a mere side-effect of a third variable that is truly responsible for both pretense and the outcome of interest. Potential emotional, social, and cognitive benefits of pretense are addressed here according to those main domains of development for the sake of organization, but many benefits cross the blurry boundaries, making categorization somewhat arbitrary.

***Emotional.*** It has been argued that the process of role-play during acting involves the adoption and managed display of a character's emotions, which could facilitate emotion regulation abilities (Goldstein & Winner, 2010b). A similar argument could be made for children's pretend play experiences. More generally, engagement in pretense might enable mastery of emotions, including fears and anxieties (e.g., Berk et al., 2006; Bretherton, 1989; Piaget, 1962). It might also provide a way to address desires that are not fulfilled in reality (e.g., Freud, 1956, as cited in Berk et al., 2006; Piaget, 1962; Vygotsky, 1933/1978). The paradox is that children most in need of assistance with emotions fixate on emotional themes during pretend play and are unable to resolve them (Bretherton, 1989). Similarly, children have been found to engage in less solitary pretend play when they are under stress (Kramer & Schaefer-Hernan, 1994, as cited in Haight et al., 2006). However, some children respond to and cope with trauma by creating imaginary companions (Taylor, 1999). Taylor suggests that having these companions might help these and other children to cope with loneliness, cope with real-life limitations, address fears, avoid blame (protect self-image), and address issues of incompetence. In addition, purposeful engagement in social pretend play (e.g., with parents' scaffolding) might help children cope with trauma or other emotional issues, depending on the way in which emotional content is handled and on the emotional theme of the scenes (Haight et al., 2006). Similarly, Berk, Mann, and Ogan (2006) discussed empirical findings that pretend play (with appropriate thematic content) helped decrease children's emotional stress after upsetting experiences. Recent individual differences research also shows empirical support for a relation between pretense and emotion through positive correlations between observed pretend play skills and parent-report of emotion regulation (Hoffmann & Russ, 2012). In their critique of empirical work regarding a relation between pretend play and emotion regulation in particular, Lillard and colleagues (2012) noted null findings and flaws in various studies, including a lack of appropriate controls. This suggests that, as those authors concluded, further research is needed to clarify any benefits for emotional development that pretending might offer.

***Social.*** When engaging in pretend play, children have the opportunity to explore social roles, including those that might be personally relevant in the future (Erikson, 1950, as cited in Berk et al., 2006). Pretense activities also offer companionship, even when partners are imaginary (Taylor, 1999), and the interactive nature of social pretend play involves cooperation, negotiation, and communication skills. It has been argued that engagement in pretend play helps in friendship development (Singer & Singer, 1990, as cited in Carlson & Beck, 2009). Indeed, some empirical findings relate pretend play to social competence. For example, in a longitudinal study of the toddler and preschool years, Howes and Matheson (1992) found that children who showed more or earlier emergence of cooperative and complex social pretense tended to score higher on social competence (e.g., prosocial behaviors, peer acceptance). In addition, engagement in more social pretend play at 30 to 35 months was associated with friendliness, being prosocial, gregariousness, sociability, fewer peer problems, less hesitancy and withdrawal, and more instrumental but not hostile aggression at 44 to 60 months (Howes & Matheson, 1992). It is possible that children with higher social skills engage in more pretend play. Similarly, pretend play itself could facilitate development of social skills. Especially without controlling for earlier social skills, causal relations are unclear. In addition, the relation between social pretend play and social skills is affected by both child's and play partner's sex (Colwell & Lindsey, 2005), and solitary (not social) pretense has been associated with poor peer relations and maladjustment (Nelson, Hart, & Evans, 2008). Such findings suggest that pretend play might facilitate social skills development under some conditions but not others. According to Lillard and colleagues (2012), the lack of consistency in the literature indicates that any relation between pretense and social development is not a crucial one, necessary for the development of social skills; however, these authors state that it is possible that pretending is one of a collection of possible routes towards social skill development.

***Cognitive.*** Related to the possible social benefits discussed above, pretend play has been associated with the socio-cognitive understanding that people have their own mental states that drive their behaviors (often referred to as *theory of mind,* ToM). For example, children's engagement in pretend play has been found to predict later performance on ToM tasks (Youngblade & Dunn, 1995), although there are also conflicting findings (Rakoczy, Tomasello, & Striano, 2006; Taylor et al., 2004). In addition, children who engage in more pretend play also produce more internal state talk (Howe, Petrakos, Rinaldi, & LeFebvre, 2005), having imaginary companions was associated with higher ToM in 4- but not 3-year-old children (Taylor, 1999; Taylor & Carlson, 1997), and not engaging in impersonation of imaginary characters by age 5-8 years was associated with lower ToM (Taylor et al., 2004). A more recent study focusing on middle childhood indicated that parent-report of children's engagement in role-play and pretense was a significant predictor of children's ToM performance, even after controlling for SES, age, and verbal IQ (Goldstein & Winner, 2010a). In their critique of the literature, Lillard and colleagues (2012) found inconsistencies in findings, methodological flaws, and possible confounds that led them to conclude that there is not enough evidence to support an argument that pretend play facilitates ToM development. They acknowledged, however, that there are theoretical reasons to believe a relation might exist. A link between pretend play and ToM has been hypothesized based on a variety of commonalities, including the entertainment of non-existent, hypothetical scenarios (Harris, 1994), a meta-representational cognitive system (Leslie, 1987), a decoupling process that separates mental representations from the reality to which they relate (Leslie, 1987; Lillard, 2001), and a process of simulating the thoughts and mind of another person (Harris, 2000). In the context of role-play, in particular, children might practice ToM-related processes as they play their roles and coordinate with other players (Goldstein & Winner, 2010b; Lillard, 2001). Of course, as emphasized recently by Lillard and colleagues, if there is a relation between pretense and ToM, more work is necessary to determine the source of causality (i.e., the causal direction between those two variables or the identity of a third variable that facilitates both).

Lillard and colleagues (2012) applied similar reasoning to many other cognitive functions that have been suspected of benefiting from pretense. Other authors have made theoretical and empirical arguments in favor of a facilitative effect of pretense. For example, Vygotsky (1933/1978) emphasized that pretend play facilitates language development, due to their shared symbolic nature. Fein (1981) cited research relating pretend play to language, as well as to IQ and understanding of conservation. Recent empirical work by Hoffmann and Russ (2012) illustrated a number of positive relations between aspects of children's pretend play (e.g., organization, imagination, and affect) and aspects of their creativity, measured through tasks focused on divergent thinking and storytelling. The literature has also emphasized the relation between pretense and another cognitive element: executive function. It is that relation that is examined in depth in this dissertation work.

## Executive Function (EF)

**Defining EF.** What is "executive function" (EF)? This term (also presented as "executive functions" or "executive functioning") is often used to refer to a collection of higher-order cognitive abilities that are needed to regulate one's own thoughts and actions (e.g., Carlson, 2005). A lack of clarity regarding the term, the construct, and its measurement abounds, however, with controversial issues such as: whether EF is a unitary construct, a collection of separable processes, or both (e.g., Ionescu, 2012; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), how EF overlaps with yet also differs from “effortful control” and “self-regulation” (e.g., Beck, Carlson, & Rothbart, 2007; Liew, 2012; McClelland & Cameron, 2012; Zhou, Chen, & Main, 2012), and whether EF might be better conceptualized as a framework that ties together a variety of sub-functions relating to the common purpose of conscious problem solving (e.g., Zelazo & Müller, 2011). As these issues are well beyond the scope of this dissertation, suffice it to say that the term “EF” is used in this discussion to refer to the collection of cognitive skills used to control one’s thoughts and behaviors, especially when impulsivity, or a lack of reflection, would be counterproductive.

While preschoolers struggle with many elements of EF, these skills might play an important role in children’s later successes in personal, academic, and professional life. For example, positive relations between preschoolers’ inhibitory control and both theory of mind (Carlson & Moses, 2001) and emotion regulation skills (Carlson & Wang, 2007) suggest that EF relates to better social-cognitive and emotional functioning. Other research has shown positive correlations between preschoolers’ inhibitory control and their kindergarten academic abilities, such as math and letter skills (Blair & Razza, 2007). On a more longitudinal basis, researchers have found that individual differences in preschoolers’ performance on a delay of gratification task are predictive of later differences in Scholastic Aptitude Test (SAT) scores, abilities to cope with stress and frustration, and concentration skills, among other outcomes in adolescence (Mischel, Shoda, & Rodriguez, 1989). Therefore, it is crucial for both basic research and application purposes to understand how EF develops and might be facilitated for individuals (e.g., preschoolers) who have EF difficulties.

**A relation between pretense and EF.** Carlson and Beck (2009) highlighted many theoretical and empirical links between pretense and EF. They emphasized the potential for pretense to allow children to distance themselves from challenging situations, thereby enabling them to reflect upon the situation and control their responses and actions more easily. The authors also explained that pretense might facilitate the internalization of rules regarding appropriate behaviors and methods for controlling one's self. Many of the ideas they addressed are included and elaborated upon in this discussion.

According to Vygotsky (1933/1978), "A child's greatest self-control occurs in play" (p. 99). He emphasized a paradox involved in play: Children cannot play freely, according to their immediate impulses, because they have to abide by rules. Children engaging in pretend play must respect not only broad social norms but also the particular conditions established in the pretense framework, especially if there are other players. As his statement illustrates, Vygotsky viewed pretend play in terms of a zone of proximal development. He believed that children are able to engage in better self-regulation during pretend play than they can in other contexts and that their self-regulatory abilities during pretense forecast (and facilitate) their later self-regulatory abilities in other realms. As Berk, Mann, and Ogan (2006) explained, Vygotsky's reasoning was that pretense involves separation of mental representation from reality through the focus on imaginary situations, and that separation facilitates deliberate choice of actions rather than impulsivity (e.g., in object substitution, a child must treat the object as its imaginary identity, not as the real identity).

This notion of the separation of mental representations of pretense and reality is reminiscent of Piaget's (1962) description of children's "progressive differentiation between 'signifier' and 'signified'" (p. 101). This notion is incorporated into many models of the mechanisms underlying pretend play that are presented in the literature. According to Leslie's (1987) "decoupler" model, children "quarantine" their pretend representations from their true primary representations in order to avoid the confusion that would otherwise result from the two different and conflicting sets of representations. When given input associated with pretense, they can produce output that matches, since the real-world knowledge and representations are maintained separately and do not interfere as they otherwise might (Leslie, 1987). Similarly, the Twin Earth Model suggests that pretense involves a cognitive separation between pretend and real representations (Lillard, 2001). The maintenance of a disconnect between pretense and reality might facilitate children's ability to manage behaviors based on one versus the other.

This distancing between pretend and reality is considered to be a type of psychological or symbolic distancing (Carlson & Beck, 2009). One of the developmental trends in pretense recognized in the literature involves a progression from more to less realistic content (e.g., Fein, 1981), as noted earlier in the discussion of object substitution (see Overton & Jackson, 1973). As explained by Carlson and Beck, psychological distancing promotes creativity and flexibility. In fact, for older children, who have developed more advanced representational systems, similarity between an object and what it stands for inhibits pretense (Fein, 1981). This symbolic distancing might enable children to think creatively about possible responses to a stimulus and behave based on active self-regulation rather than impulse (Carlson & Beck, 2009).

Evidence from a variety of EF tasks has supported this idea that psychological distancing promotes self-regulation. For example, Mischel, Shoda, and Rodriguez (1989) used a Delay of Gratification (DoG) task in which children are left in a room with desirable treats and informed that if they wait until the experimenter returns to the room then they will receive a larger amount of treats; otherwise, they can only receive a smaller amount. One of the findings from this research was that children waited longer when they could see only a symbolic representation of the reward (rather than the actual reward) during the wait. A study by Carlson, Davis, and Leach (2005) using the Less is More task also illustrated the facilitative effect of symbolic distancing. To achieve the most personal gain on this task, a child must inhibit the more natural response of pointing to a container with more treats and point to a container with fewer treats instead. When preschoolers could see the two sets of real treats, this was challenging. However, when they could see only symbolic representations of the sets of treats (e.g., rocks representing treats in a 1:1 ratio), performance was significantly improved. Also, symbols more distant from reality (e.g., mouse for the smaller amount and elephant for the larger amount) led to better performance.

The ways in which these other types of symbols facilitate self-regulation might be applied to pretend play, as well. The separation between pretense and reality during pretend play might allow children to distance themselves from their situation and reflect more upon it. As discussed by Carlson and Beck (2009), role-play, in particular, might enable the consideration of a situation and the appropriate ways in which people (including the self) might respond.

This relates to another theoretical link between pretense and EF: private speech. It has been shown that children's private speech (self-directed talking aloud) is associated with better task performance, often occurs when a task is challenging, and is indicative of children's internalization of previously external sources of control and guidance (Berk et al., 2006). This type of speech has been described as the step between others externally and verbally directing one's behaviors and silently directing oneself (Carlson & Beck, 2009) . Therefore, engagement in private speech might enable children to consider their situation from a more distant perspective (similar to someone else's). That type of consideration is potentially facilitated in a similar manner during role-play, which requires the child to take on the perspective of someone else.

In sum, EF and pretense have theoretical relations based on the self-regulation that pretense requires and the distancing involved in pretense, which might allow for more self-reflection and regulatory abilities. The next section addresses empirical evidence supporting these relations.

**Empirical evidence.** One basic piece of evidence in support of a relation between EF and pretense is that they have similar developmental timelines. Many EF skills have been found to strengthen between ages 3 and 5 years (e.g., Carlson, 2005), which was mentioned earlier as a time for development of the more sophisticated forms of pretense (e.g., Howes & Matheson, 1992). There is evidence supporting a relation from individual differences research, as well. For example, 3-year-olds who engaged in more impersonation of imaginary characters exhibited fewer problems with self-control, according to their mothers (Taylor et al., 2004). In addition, 6- to 9-year-olds higher in fantasy engagement were able to wait longer than lower-fantasy peers in a delay task presented as a screening test for future astronauts (Singer, 1961, as cited in Carlson & Beck, 2009). More recently, 4- to 7-year-old children's observed spontaneous pretend play and measured structured pretend play were each positively correlated with their inhibitory control, although only the relation between structured pretend play and inhibitory control remained significant after controlling for mental age (Kelly & Hammond, 2011). In addition, research with certain subpopulations (e.g., children with conduct disorder) reveals that deficits in EF and pretense often co-occur (Carlson & Beck, 2009).

The research suggesting parallel developmental timelines and correlated individual differences in pretense and EF skills has often measured the two sets of skills separately – either at different time points or simply through consecutive tasks, with each task measuring one set of skills. In contrast, Nader-Grosbois and Vieillevoye (2012) recently reported data on both sets of skills that were collected simultaneously. In this study, observers watched dyads of typically developing children and of children with intellectual disabilities during both highly structured and less structured pretend play activities. Children were scored on the complexity of their pretense (e.g., roles, actions, and objects), as well as on their involvement and social participation in the pretense. During these same activities, children were rated on self-regulation abilities (e.g., objective identification and planning). Researchers found many strong and positive correlations between pretense skills and self-regulation skills for both groups of dyads, especially within the context of the more highly structured pretend play activity. It is noteworthy that there have been null findings regarding a relation between pretense and certain measures of EF (e.g., Hoffmann & Russ, 2012), as well. Therefore, although the collection of evidence regarding individual differences suggests that pretense might be linked in development to some EF skills, the specific nature of the relation is unclear.

A link of some kind is supported by findings from longitudinal studies, such as Elias and Berk's (2002) work with 3- and 4-year-olds. In their work, observed frequency and persistence of children's complex sociodramatic play at Time 1 predicted children's self-regulation during a clean-up activity 4 to 5 months later. This result was significant even after controlling for age, verbal ability, and self-regulation observed during clean-up at Time 1. This relation held for the total sample of children and for children who were highest (but not lowest) in impulsivity. Solitary dramatic play at time 1 negatively correlated with children's self-regulation during clean-up at time 2, even after controls. Researchers found no relation between pretense and self-regulation during a circle-time activity. Reportedly, a similar study failed to replicate these findings within a lower SES bracket (Berk et al., 2006). Collectively, these results again illustrate a relation between pretense and EF that may be limited to certain types of pretense, subtypes of EF, contexts, and demographics.

If pretense (engagement and skills) does relate to EF, can training programs emphasizing pretense increase children's EF? Several interventions have been found to facilitate EF development through the use of various activities, including physical activity, games, mindfulness, and classroom curriculum (Diamond & Lee, 2011). Most relevant here are studies investigating the effectiveness of the Tools of the Mind (Tools) curriculum, which includes an emphasis on pretend play. According to Barnett and colleagues (2008), developers Bodrova and Leong created this program based on Vygotskian notions regarding dramatic play and self-regulation. As part of this program, adults encourage children to cooperatively plan out and act out complex pretend scenarios. The assumption is that these pretend play activities demand self-regulation and therefore provide opportunities for children to practice and develop the necessary skills. Barnett and colleagues conducted a study of the Tools program in which they randomly assigned children to participate in either the Tools or a control curriculum. The groups were equivalent on demographic measures and various pre-test measures of cognitive abilities, including expressive and receptive vocabulary, reading, math, and problem-solving skills. These researchers found that children who began the Tools program in the fall were reported in the springtime to have fewer behavior problems and to need less management by teachers than children who had been part of a control program during that same time period. Unfortunately, pre-test data on these outcomes were not available, but condition assignment was randomized (Barnett et al., 2008). In other research by Diamond and colleagues (2007), after 1-2 years in a given program, children in the Tools program outperformed peers in a control group on direct behavioral measures of EF. There were no pre-test EF measures, but, again, condition assignment was randomized, suggesting that groups should have been similar at the outset (Diamond et al., 2007). Although the effects of the pretend play component were not isolated in these studies, it is possible that pretense involvement contributed to these better self-regulation-related outcomes. However, as others (Lillard et al., 2012) have acknowledged, it is also possible that any benefit from the Tools program derives from a different feature of the intervention. In addition, Lillard and colleagues questioned the effectiveness of the Tools program, citing three conference presentations that reported studies finding no effects of the Tools curriculum on EF outcomes.

Another collection of research regarding a hypothesized contribution of pretense to EF abilities involves studies in which fantasy is explicitly used to promote self-regulation. In a series of studies, preschoolers were better able to overcome their “empirical bias” and reason according to counterfactual premises (e.g., “all cats bark”) when they were cued to use their imagination and engage in make-believe (Dias & Harris, 1988, 1990). In these studies, children were essentially coached to use pretense as a strategy for managing the challenging task, and the data suggest that this was effective. Research on children’s spontaneous strategies during a DoG task found that 16% of children used pretense strategies at least once, even without any coaching (Carlson & Beck, 2009). In that study, both performance on DoG and the use of pretense increased with age, and the use of strategies (including pretense) was associated with longer waiting (although pretense did not contribute uniquely). The authors indicated that while children sometimes use pretense spontaneously in the service of EF, future research might address the possibility that children might be taught to use pretense more effectively than some of them naturally did. More recently, Karniol and colleagues (2011) differentiated between a pretense strategy of children cognitively transforming the reward object or the context in a DoG task and a pretense strategy of children cognitively transforming themselves. In their research, children tended to delay longer when they were given instructions to pretend to be Superman, especially when those instructions emphasized his relevant capabilities. Thus, there have been a number of studies investigating the facilitative effect of pretense-related strategies in the DoG task, which is a task emphasizing delay. Carlson and Moses (2001) provided evidence that inhibitory control tasks can be divided into two distinct categories: delay (requiring a person to delay their impulsive response) and conflict (requiring a person to activate a response that conflicts with the more natural response). To the best of my knowledge at the start of this dissertation work, the potential facilitative effect of pretense strategies had not yet been investigated in the context of an EF task that emphasizes conflict.

The recently published argument of Lillard and colleagues (2012) was that "evidence that pretend play assists EF is sparse at best" (p. 23). Part of their explanation involved the specificity of empirical findings, inconsistencies, and failures to replicate. In response, I would argue that specificity of empirical findings should not necessarily devalue the relations that are found. It is true that limitations (regarding when a relation exists and when it does not) in turn restrict the claims that one should make about a broad pretense-EF relation; however, the fact that a link is not always present does not mean that when the link is present it is too weak to be of interest or value. The authors seemed to dismiss results showing that higher fantasy children were better than lower fantasy children at remaining still in a task that incorporated a pretend scenario. They dismissed the findings because the pretense only seemed helpful for some children; however, this research could be interpreted as useful additional evidence that a pretense strategy can be effective under some circumstances. This view of the specificity of findings could similarly be applied to what the authors described as inconsistencies and failure to replicate. For example, the fact that some training studies found EF benefits of the Tools of the Mind program and other studies did not could be influenced by the lack of consensus in the field on how to define and operationalize EF. Such differences in approach and methodology within the relevant literature might be masking more precise patterns of pretense-EF relations that would be apparent if one were to limit the comparisons to sets of studies that defined and measured the construct(s) similarly. In the same way, a failure to replicate findings from one study in another study with a new population could be an indication of the specificity of the pretense-EF relation (i.e., its limited generalizability), rather than a sign that any pretense-EF relation captured in the original study is a weak or spurious one. It is up for debate where we ought to draw a line between interesting and useful specificity versus specificity so narrow that it is impractical to apply or interpret it. With that debate in mind, the more we identify cases in which pretense and EF are and are not related, the more we will understand whether, how, and when a pretense-EF relation operates.

**Current Research Program**

The current research program was thus designed to explore further the relation between EF and pretense that has been suggested in the literature. This exploration involved the use of parent report, observation of structured and less-structured tasks, and experimental manipulation to address the hypothesized positive relation between preschoolers' EF skills and their pretense abilities or engagement. The discussion here will be limited to the main focus of the research, which was an experimental manipulation of an EF task involving conflict. This manipulation specifically addressed the hypothesis that preschoolers who are given pretense-based instructions for an adapted version of the Dimensional Change Card Sort task (DCCS) (Zelazo, 2006) will perform better on the post-switch card sort trials than will preschoolers who are given more standard (non-pretense-based) instructions. The DCCS is a commonly used EF task in which children are presented with two boxes, each labeled with a different picture (e.g., blue star or red truck). Then, they are taught to sort cards depicting red stars and blue trucks into those boxes, based on a certain rule (e.g., sort by matching color). In its essential EF-related phase, children are then asked to sort by a conflicting rule instead (e.g., sort by matching shape). The task therefore requires children to exhibit EF skills during the post-switch phase by stopping themselves from sorting by the first rule and sorting by the second rule instead. As numerous studies have shown, many 3-year-olds typically fail this task, and performance improves with age across the preschool years (e.g., Carlson, 2005; Carlson & Moses, 2001; Dick, Overton, & Kovacs, 2005; Frye, Zelazo, & Palfai, 1995; Perner & Lang, 2002; Zelazo, 2006).

**Children's Struggles on the Standard DCCS**

There are a variety of explanations for younger children's poor performance on the DCCS, and this continues to be a topic of debate in the literature. A common theme, however, is that of children's struggle to overcome a tendency associated with the pre-switch rule in order to sort correctly during the post-switch phase. As illustrated in the examples discussed below, the nature of that struggle is controversial, with some theorists arguing that it is a matter of attentional focus or cognitive immaturity and others indicating that it is a matter of behavioral impulse.

During the pre-switch phase of the DCCS, children first sort based on a particular dimension displayed on the cards (e.g., color). Once the rules are switched, children must (a) stop sorting based on that particular dimension and (b) start sorting based on the dimension that they previously ignored (e.g., shape). According to an *attentional inertia* explanation, young children fail the DCCS when they have trouble with part (a): they cannot overcome the tendency to focus their attention on the first dimension, which interferes with their ability to sort according to the new dimension (e.g., Diamond, Carlson, & Beck, 2005; Kirkham, Cruess, & Diamond, 2003). Another explanation centering on the challenges of part (a) argues that children have difficulty stopping themselves from the behavioral action of sorting based on the pre-switch rule because they cannot inhibit their *prepotent response tendencies* (e.g., Carlson & Moses, 2001).

Children's problems might also stem from part (b): the *negative priming* argument focuses on the fact that the post-switch phase challenges children to cognitively re-activate the dimension that was previously inhibited during the pre-switch sorting game (because it had interfered with successful sorting in that phase) in order to sort correctly in the post-switch phase (e.g., Perner & Lang, 2002).

Essentially combining parts (a) and (b) above, the *re-description hypothesis* suggests that young children struggle with the understanding that the same object can and must be re-described in a different way during the post-switch phase (e.g., Kloo & Perner, 2005). This argument suggests that the combination of two conflicting notions is difficult for children. This idea is similarly highlighted in the original and revised versions of the *cognitive complexity and control theory* (CCC and CCC-r, respectively), espoused by Zelazo and colleagues (Frye et al., 1995; Zelazo, Frye, & Rapus, 1996; Zelazo, Müller, Frye, & Marcovitch, 2003). According to the CCC and CCC-r theories, younger children struggle with the DCCS because they do not yet have a mature ability to think in terms of a complex hierarchical rule structure. They can know about two sets of conflicting rules and act based on one or the other set of rules (e.g., sort by color or sort by shape), but they have trouble organizing the two sets of rules under a broader setting rule (e.g., if playing the color game, sort by color). This inability results in their failure to properly and actively choose the proper sorting rule during the post-switch phase.

Numerous prior studies have attempted to identify the best explanation for young children's failure on the DCCS. Many researchers have reported that children succeed on manipulated versions of the task that remove some proposed challenges (e.g., the requirement to overcome negative priming effects) but maintain others (e.g., an embedded rule structure), suggesting that the latter components of the DCCS cannot be responsible for children's struggles (e.g., Diamond et al., 2005; Kloo & Perner, 2005; Mack, 2007). The debate about the best specific explanation for young children's failure on the DCCS is tangential to the focus of the present research and beyond the scope of this discussion. The most relevant idea underlying many of the popular arguments is that children struggle during the post-switch phase because (for one reason or another) they fail to reflect appropriately on the relevant dimension in order to perform the correct associated sorting action.

# Chapter II: Study 1

## Overview

A pretense-based manipulation of the standard DCCS was implemented in Study 1 to determine if pretense might facilitate better performance in children 3 to 4 years old. This new condition, inspired by Dias and Harris (1990), created a pretense-based context for the card sort through the use of a make-believe tone of voice, a visual aid, and instructions to the child to pretend to take a rocket ship to a set of imaginary planets (Planet Color and Planet Shape). Based on the arguments that pretense involves the mental separation of conflicting representations (e.g., Leslie, 1987; Lillard, 2001), a pretense context might assist preschoolers with the post-switch challenge of negotiating conflict (i.e., regarding focus of attention, action schemas, descriptions of entities, or embedded rules). In a series of studies, Wyman, Rakoczy, and Tomasello (2009) found that 3-year-olds were able to switch between pretending an ambiguous object was identity A (e.g., a toothbrush) and pretending it was identity B (e.g., a spoon), depending on which of two previously associated contexts was currently relevant (e.g., pretense partner 1 or pretense partner 2). This shows that when young children are pretending, they are able to switch between conflicting representations of an object based on its context. Considering that finding, the Fantasy Planet manipulation for the DCCS might facilitate children's ability to switch between the conflicting rules or their associated representations. Overall, the psychological distancing argument about the cognitive processes involved in pretense (e.g., Carlson & Beck, 2009) suggests that a pretense component might facilitate better performance on the DCCS by allowing children to behave more purposefully instead of impulsively. As indicated earlier, this notion is compatible with many of the explanations proposed for why children struggle on the DCCS, since the majority of explanations involve failure to behave based on proper reflection.

Thus the primary hypothesis for Study 1 predicted that a pretense context should facilitate better DCCS performance for preschoolers. More specifically: Preschoolers should perform better when the DCCS is presented in a pretense-based context than when it is presented in a more standard manner without any pretense incorporated into it. In addition, the difference between conditions should be significant even after controlling for other variables. In particular, prior research suggests that EF is related to age and receptive vocabulary (e.g., Carlson, 2005; Carlson et al., 2005; Carlson & Moses, 2001; Hongwanishkul, Happaney, Lee, & Zelazo, 2005). Therefore, both of those variables were controlled in this research.

Three secondary measures of EF were included in this study, as well, partially for purposes tangential to this discussion. The tasks were Forward and Backward Digit Span (FDS and BDS, respectively), which emphasize working memory, and Bear-Dragon, which emphasizes inhibitory control. All of these tasks are recognized in the literature as EF tasks that are appropriate for ages 3 years and older (Garon, Bryson, & Smith, 2008). Working memory and inhibitory control often have been and continue to be considered among the key components of EF (Garon et al., 2008; Zelazo & Müller, 2011). Referencing Baddeley's (1981) model of working memory, Zelazo, Müller, Frye, and Marcovitch (2003) explained that a forward digit span task measures the *articulatory loop*, or memory span, and a backward digit span task measures the *central executive*, which is responsible for the more active processing. Davis and Pratt (1995) argued that the active processing capabilities of the central executive depend on the encoding and retrieval skills captured by the model's articulatory loop. Therefore, these researchers indicated that the central executive portion of working memory can be measured directly with BDS and indirectly with FDS. Carlson (2005) identified another distinction, indicating that FDS is a "pure measure of working memory" while BDS also depends on inhibitory control (p. 612). Although there is limited work relating the digit span tasks to the DCCS specifically, researchers have found significant relations between children's performance on working memory tasks (e.g., backward digit/word spans) and their performance on tasks requiring other components of EF (e.g., DCCS, Bear-Dragon) (e.g., Carlson, Moses, & Breton, 2002; Oh & Lewis, 2008). For example, Espy and Bull (2005) found that children who had longer memory spans (i.e., higher FDS scores) outperformed children with shorter spans on EF tasks that involved attentional control. EF tasks such as the DCCS arguably depend on both the memory span and active processing components of working memory. Therefore, FDS and BDS in this research will be discussed as additional EF measures to inform the interpretation of any group differences in DCCS performance.

Besides FDS and BDS, the other secondary EF task in this research was an adaptation of Bear-Dragon (Reed, Pien, & Rothbart, 1984). The Bear-Dragon task is often included in EF batteries because of its reliance on the conflict type of inhibitory control (e.g., Carlson & Moses, 2001; Carlson et al., 2002). While Bear-Dragon was originally included in this study as part of an exploration of parents' and children's EF strategies and was not intended to be part of the main focus, it is tentatively included in this discussion as a third measure of EF that has the potential to shed light on the interpretation of any group differences on the DCCS. However, as is discussed below, because the goal was to explore parents' and children's strategies, the task was administered in an unconventional manner. Therefore, conclusions depending on its validity as an EF measure must be considered tentative.

In addition to the main investigation into the effect of a pretense context on the DCCS, there were more exploratory questions about possible effects of sex or individual differences in children's pretense experiences. Data about individual differences in EF and pretense were gathered through parent questionnaires. Most data from the questionnaires are beyond the scope of this discussion, with the exception of basic demographic data and temperament data, which will be discussed briefly.

Parent-report of children's temperament was included as part of the exploration of individual differences because of prior evidence that temperament can relate to children's EF (e.g., Beck et al., 2007; Blair & Razza, 2007; Carlson & Moses, 2001; Hongwanishkul et al., 2005). For example, the Surgency and Effortful Control subscales of the Child Behavior Questionnaire (CBQ) have correlated with children's EF task performance, even after controlling for age (Beck et al., 2007). Also, researchers have begun to make distinctions between the types of EF tasks that tend to relate to CBQ subscales and those that do not. In particular, the collection of Inhibitory Control items on the CBQ has been related to performance on "conflict" EF tasks (which involve a conflict between the correct response and a more salient response) but not to performance on "delay" EF tasks (which require the temporary or total suppression of an impulsive response) (Carlson & Moses, 2001). Similarly, Effortful Control and Surgency subscales of the CBQ have related to children's performance on "cool" EF tasks (which test the cognitive components of EF in the absence of affect) but not to their performance on "hot" EF tasks (which involve affective salience), although all relations in that study were non-significant once age was partialled out (Hongwanishkul et al., 2005). Considering prior research, data on children's temperament were collected here with the expectation that they might inform interpretation of the DCCS data.

Also, during an early phase of Study 1, two age groups were created to explore potential differences between younger (40-43 months old) versus older (44-47 months old) preschoolers. Although this was also more exploratory in nature, one hypothesis arose out of the idea of a zone of proximal development: The pretense manipulation might only help children who are on the brink of being able to succeed on the task alone (e.g., older children).

**Method**

**Participants.** Many researchers have reported that 3-year-olds tend to perform poorly on the DCCS, while 4-year-olds perform better (e.g., Carlson, 2005; Carlson & Moses, 2001; Dick et al., 2005; Kirkham et al., 2003; Zelazo, 2006). Based on the associated literature, initial recruitment focused on children 40 to 44 months old. Preliminary data indicated that there might be stronger effects with an age range expanded upward. Therefore, the final sample included 48 "younger" children (ages 40-43 months, $M = 41.44$, $SD = 1.03$) and 48 "older" children (ages 44-47 months, $M = 45.35$, $SD = 1.06$), for a total of 96 children, ages 40 to 47 months ($M = 43.40$, $SD = 2.23$). Half of the children in each age group were girls. Participants of each age group and sex were randomly assigned to a DCCS condition (Fantasy Planet or modified Standard) and order (shape-color or color-shape). The between-subjects design therefore consisted of 16 cells based on the various combinations of age group, sex, condition, and order, and there were 6 children per cell.

Twelve additional children participated but were not included in the final sample due to failure to complete the DCCS ($n = 3$), low DCCS pre-switch scores ($n = 3$), experimenter error on the DCCS ($n = 1$), equipment failure ($n = 1$), or health concerns or developmental delays that parents reported during the session ($n = 4$). In addition, 33 other children participated in two

DCCS conditions that were abandoned partway through the study and will not be discussed here. These children are not included in the final sample of 96.

Children were generally recruited from the University of Washington (UW) Communication Studies Participant Pool (CSPP), which is a database of Seattle-area children whose parents have elected to be contacted about university research opportunities for them. Families in this database tend to be middle- to upper-middle-class socioeconomically. Any non-CSPP participants proactively contacted the researcher after learning of the project through other participants or via a flyer seen in the community. Parents reported on the children's race, ethnicity, health/development, and language exposure on the parent questionnaire. Most participants' racial background was reported to be White (80.2%) or multiracial (16.7%), with the remaining few described as Asian (2.1%) or other (1.0%). In terms of ethnicity, the majority of participants were described as non-Hispanic (91.7%) versus Hispanic. Study inclusion criteria included a lack of known developmental delays or significant health problems, and English language abilities were also required. At least one parent/guardian also had to speak and understand English well enough to participate in the consent process, parent questionnaires, and (for a subset of families) parent task participation that is not part of the main study reported here.

**Procedure.** Participation involved one visit to a child-friendly lab at the UW-Seattle campus during one of three waves of data collection. After the parental consent process and a brief introduction to the experimenter and the testing room set-up, the child and the experimenter went into a testing room and engaged in a series of tasks together. Parents viewed the activities from a nearby room through a window or via live video feed. For the first two waves of data collection ($n$ = 64), the session lasted approximately 60-75 minutes and included the following activities (in order[1]): Peabody Picture Vocabulary Test (PPVT-IV), DCCS, Pretend Actions, FDS and BDS, Free Play, Clean-Up, Bear-Dragon, and Forbidden Toy. Parents participated in the last four activities. During the final wave of data collection (*n* = 32), parents were not involved in the sessions, which were limited to approximately 30 minutes and included only the following activities (in order): PPVT-IV, DCCS, FDS, BDS, and Bear-Dragon. Pretend Actions, Free Play, Clean-Up, and Forbidden Toy will not be mentioned further in this dissertation as they were not part of the primary focus of the research.

All parents were asked to complete the very short form of the Child Behavior Questionnaire (CBQ), which is an established shorter version of the 195-item parent-report of child's temperament (Putnam & Rothbart, 2006), and a separate parent questionnaire. Only the demographic data from the latter questionnaire are used in this study, although the exploratory portions also addressed parent and child attitudes and tendencies regarding pretense and executive function. These questionnaires were given to parents in advance, and they were completed prior to or during the testing visit. The primary investigator conducted all testing sessions. In exchange for participation, families' parking was paid for out of research funds, parents were debriefed, and children received a small gift (e.g., toy or UW t-shirt).

[1] For all waves of Study 1 and for Study 2 there was a pre-determined task order; however, in some instances the tasks were re-ordered in an effort to accommodate the child's wishes. For all participants, the DCCS was administered without pausing for other tasks, but some children needed a break in the middle of the PPVT or refused to play the Digit Span or Bear-Dragon game during an early stage of the task administration. In those cases, the experimenter set the particular task aside and tried again later in the session.

**Measures.**

***Main components.***

*PPVT-IV.*

The Peabody Picture Vocabulary Test (4th edition; Dunn & Dunn, 1997) is a standard measure of receptive vocabulary that involves a book of pictures. Using this test, the experimenter showed the child a page of four pictures, said a word, and asked the child to point to the picture that showed the word the best. After the practice phase, the experimenter provided only general encouragement and no feedback. Words were presented in sets of 12, which got progressively more difficult. The task was continued until a child got a minimum of 8 words in a single set incorrect. Individual scores were computed based on the number of items answered correctly and the highest item answered. Raw scores were then converted to standardized scores based on the child's age in months, using tables provided by the developers of the PPVT.

*DCCS.*

At the beginning of the DCCS task (adapted from Zelazo, 2006), the experimenter placed a card stand on the table and explained that she would display cards on this stand before handing them to the child to sort [2]. She then introduced the child to two black recipe boxes with openings on the top. One box was labeled with a target card depicting a blue star on a white background, and the other box was labeled with a target card depicting a red truck on a white background. The initial instructions for the game depended on the condition and order associated with the child's assigned cell (see Appendix A for sample scripts). In the Fantasy Planet condition, the experimenter encouraged the child to pretend to take a rocket ship to Planet Shape (or Planet

---

[2] The card stand was essential for one of the original alternative DCCS conditions, which is not discussed here. It was retained in all conditions for Study 1 for the sake of consistency, even after that original condition was abandoned partway through the study. It was not used in Study 2.

Color). As she described this fantasy context, she displayed a visual aid (see Appendix B), depicting three planets (Earth, Planet Shape, and Planet Color) and pointed to the appropriate travel trajectory and planets. She put the visual aid away before explaining the rules of the game. In the modified Standard condition, the experimenter explained that she had learned the game from a friend and wanted to play it with the child on the table in front of them.

After introducing the game, regardless of condition and order, the experimenter explained the rules (e.g., "All the stars go here, and all the trucks go there") and gestured to the appropriate boxes. She next demonstrated the correct placement for one card (a red star) and repeated the rules (e.g., "If it's a star, it goes here; if it's a truck, it goes there") and gestures. The child was then given a practice trial (a blue truck) and was corrected as needed. Then, a series of six pre-switch test trials began, with no feedback provided. The six test cards included three of each kind (red stars and blue trucks), in a set order, with no more than two matching cards in a row. Before each test trial, the experimenter repeated the game's rules, with an additional game reminder on the fourth trial (see Appendix A for additional details). The experimenter then placed a new test card on the stand, labeled it by the appropriate dimension (e.g., "Here's a red one" or "Here's a star"), and handed it to the child.

Following the sixth pre-switch test trial, the experimenter transitioned into the post-switch game, which involved sorting by the other dimension (color/shape). Again, the script differed depending on condition; however, the general procedure for introducing the post-switch game was similar to the procedure for the pre-switch instructions (but without the demonstration and practice trial). Then, six post-switch test trials progressed similarly to the six pre-switch test trials.

All scripts were designed to be similar in duration and frequency of words regarding shape/color and parallel in form. During test trials, the experimenter remained as neutral as possible, responding to child inquiries with only vague instructions (e.g., to place the card in *a* box) and avoiding looking towards either box. Of primary interest was the child's decision regarding into which box to place the card. If the child changed his or her mind after initially putting the card into one box, the final answer was counted as long as the child did not obtain information in between the original and final response (e.g., if the child clearly looked into the box or heard the experimenter say the rules again, then the original decision was counted). All children included in the final sample sorted at least 5 out of the 6 pre-switch cards correctly. The number of post-switch cards each child sorted correctly was used as that child's DCCS score in the analyses (max = 6).

For the DCCS, all sessions were coded by two independent coders who had not been told the hypothesis: One coder knew the general study design and coded with sound, and the other coder was blind to the study design details (i.e., number and nature of DCCS conditions) and was kept blind to condition by coding without sound. There was excellent agreement on the pre-switch score (range for final sample participants: 5-6, Cohen's $\kappa = .90$) and on the post-switch score (range: 0-6, Cohen's $\kappa = .96$). For individual post-switch trials, Cohen's $\kappa$ ranged from .96 to 1.00. There were a total of four post-switch trials (four different children) that had discrepancies between coders, and these discrepancies were resolved by a third coder who coded with sound but was naïve to the study's research questions, design, and hypotheses.

***Secondary measures.***

*Forward and Backward Digit Span.*

This set of games (based on Davis & Pratt, 1995; see also Carlson, Moses, & Breton, 2002) addresses aspects of working memory (see scripts in Appendix C). In the Forward Digit Span task (FDS), the experimenter said a string of numbers and asked the child to repeat the numbers in the same order. In introducing the game, the experimenter used an Ernie doll (Sesame Street character) to demonstrate. A practice string of two numbers was used, with the experimenter correcting the child as needed (max = 7 tries total). Then, the experimenter began test trials, starting with a new 2-digit string. If the child repeated that string correctly, the experimenter continued with a 3-digit string, and so forth, with increasing string lengths. If the child failed to repeat a string correctly, the experimenter gave additional strings of the same length until the child repeated it correctly or got a total of three trials incorrect at the same level (whichever happened first). The child's score for FDS was the highest level (i.e., number of digits in a string) successfully repeated. The Backward Digit Span task (BDS) was similar, but children were instructed to say the experimenter's string of numbers in the reverse order (e.g., if the experimenter says, "1, 2," the child should say, "2, 1"). There was a demonstration and practice trial, similar to FDS, and the experimenter used a variety of techniques to explain the rules repeatedly if correction was needed. After a maximum of 7 tries, test trials began, regardless of success, with a similar procedure to that of FDS. The child's score for BDS was the highest level successfully repeated backwards. As a reliability check, half of the participants ($n$ = 48) were double-coded, with ICC = .99 for FDS and ICC = 1.00 for BDS.

*Bear-Dragon.*

This task (an adaptation of Reed et al., 1984; see also e.g., Carlson & Moses, 2001) is similar to the popular children's game Simon Says. Bear-Dragon, as it is commonly used, is a measure of EF, since it requires children to follow the instructions of the bear puppet (e.g., "touch your nose") but inhibit a similar response to the instructions of the dragon puppet (see script in Appendix D).

In the warm-up phase, the experimenter asked the child to touch his or her own head, nose, etc., and demonstrated these actions for the child. After this warm-up, the experimenter introduced the child to a "nice" bear puppet and a "naughty" dragon puppet, explaining that the game involves doing what the bear says to do (with the experimenter using a higher-pitched voice) and not doing what the dragon says to do (with the experimenter using a lower-pitched voice). There was a practice trial for each puppet and then a verbal rule-check regarding each puppet, with corrections by the experimenter as needed. After the child indicated understanding of the rules, test trials began, with a set order of alternating instructions given by the bear and dragon puppets. After five test trials, the experimenter reminded the children of the rules, regardless of their performance, and then proceeded to the last five trials. For each trial, children received a score ranging from 0 to 3 points, with higher scores indicative of better performance. For each of the five bear trials, children received a 0 if they did not move, 1 if they made a wrong movement, 2 if they made a partial correct movement, and 3 if they made a full correct movement. For each of the five dragon trials, children received a 0 if they made a full commanded movement, 1 if they made a partial commanded movement, 2 if they flinched or made a wrong movement, and 3 if they did not move. Each child received a total proportion score, calculated as the number of points earned on all trials divided by total possible points (i.e., allowing for individual trials to be deemed un-codeable due to child's positioning and camera angle, etc.). Data were considered complete only for those children who had at least three codeable trials for each puppet (*n* = 88).

As was mentioned above, the task was adopted as part of the secondary purposes of the research. The initial goal was to gather exploratory data regarding the strategies spontaneously adopted or promoted by children and parents during this type of task. As part of this exploration, the task was adapted for the first two waves of data collection (*n* = 64) in that the parent/guardian was seated beside the child and was encouraged to help the child play the game correctly. For the final wave of data collection (*n* = 32), the parent was not involved, and only the children's performance on the task was coded. Due to the nature of the task administration, as was mentioned earlier, the data from this task were not going to be included here, so formal reliability checks were not completed. The data are only mentioned here in a post-hoc effort to tentatively address EF differences between groups.

*Very short Child Behavior Questionnaire (CBQ).*

The very short form of the CBQ (Putnam & Rothbart, 2006) was used in this research to collect parent-report of children's temperament. This 36-item questionnaire provides statements about young children's reactions to various situations and requires parents to rate the degree to which each statement is true for their child (scale of 1 to 7). The very short form of the CBQ has three established subscales: Surgency, Negative Affect, and Effortful Control. Children received mean scores for each subscale based on parents' responses to the related items. The Surgency subscale includes items relating to children's impulsivity and activity level. An example item is: "Often rushes into new situations." The Negative Affect subscale includes items capturing children's tendencies towards negative emotions (e.g., anger, sadness) and difficulties being soothed. For example, parents rate the applicability of the phrase: "Gets quite frustrated when prevented from doing something s/he wants to do." Finally, the Effortful Control subscale includes items regarding children's inhibitory and attentional control, as well as their perceptual sensitivity and low-intensity pleasure. An example item from this subscale is: "Approaches places s/he has been told are dangerous slowly and cautiously."

**Results**

The variables of interest for the primary research questions included children's age and their scores on the DCCS and PPVT. Of secondary interest were FDS, BDS, Bear-Dragon, and the CBQ, which might inform interpretation of results from the primary analyses. All 96 participants completed the DCCS and had CBQ data, but some children did not complete the PPVT ($n$ = 10), FDS ($n$ = 10), BDS ($n$ = 25) and/or Bear-Dragon ($n$ = 8) (e.g., due to child refusal or, for Bear-Dragon, too many uncodeable trials). Children without scores on a specific task were excluded from analyses involving that task. Because the main question involved comparisons between children in the modified Standard condition for the DCCS and children in the Fantasy Planet condition for the DCCS, descriptive statistics are displayed in Table 1 for these two sub-groups separately.

Table 1.

*Descriptive Statistics for Each Condition for Study 1.*

| | Standard | | | Fantasy Planet | | |
|---|---|---|---|---|---|---|
| **Measure** | ***n*** | ***M (SD)*** | ***95% CI*** | ***n*** | ***M (SD)*** | ***95% CI*** |
| Age | 48 | 43.42 (2.21) | [42.77, 44.06] | 48 | 43.38 (2.27) | [42.72, 44.03] |
| DCCS post-switch (0-6) | 48 | 2.92 (2.49) | [2.19, 3.64] | 48 | 3.52 (2.73) | [2.73, 4.31] |
| PPVT (raw) | 43 | 79.12 (19.58) | [73.09, 85.14] | 43 | 78.77 (23.81) | [71.44, 86.09] |
| PPVT (standardized) | 43 | 119.37 (14.79) | [114.82, 123.93] | 43 | 119.42 (18.55) | [113.71, 125.13] |
| Bear-Dragon | 43 | .84 (.19) | [.78, .90] | 45 | .88 (.18) | [.83, .94] |
| FDS | 44 | 3.70 (0.88) | [3.44, 3.97] | 42 | 3.62 (0.76) | [3.38, 3.86] |
| BDS | 36 | 1.33 (0.59) | [1.14, 1.53] | 35 | 1.11 (0.40) | [0.98, 1.25] |
| Surgency (CBQ) | 48 | 4.60 (0.87) | [4.35, 4.85] | 48 | 4.66 (0.78) | [4.44, 4.89] |
| Negative Affect (CBQ) | 48 | 3.84 (0.94) | [3.56, 4.11] | 48 | 4.03 (0.83) | [3.79, 4.27] |
| Effortful Control (CBQ) | 48 | 5.28 (0.77) | [5.06, 5.51] | 48 | 5.17 (0.75) | [4.96, 5.39] |

*Note.* CI = confidence interval.

**Relations between DCCS scores and other variables.** Correlations among most variables are presented in Table 2. Children's DCCS scores were positively correlated with their rounded age (in months) and with two of the secondary measures of EF: BDS and Bear-Dragon. Given the significant relations between age and both DCCS and BDS, the correlation between DCCS and BDS was further examined after partialling out age. The resulting partial correlation between DCCS and BDS was non-significant, $r = .19$, $p > .10$. The relation between age and Bear-Dragon was marginal. As a conservative test, the relation between DCCS and Bear-Dragon was also examined with age partialled out, and it remained significant, $r = .33$, $p < .01$. DCCS was not significantly related to any of the subscales of the CBQ, which is not included in the correlation matrix, *p*s > .05. For a non-parametric approach, Spearman rho similarly indicated that the DCCS was only significantly related to age and Bear-Dragon, $r_s = .23$, $p = .03$ and $r_s = .27$, $p = .01$, respectively. As planned then, age was used as a covariate in the main analyses.

Table 2.
*Pearson r Correlation Matrix for Study 1.*

| | Measure | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 1 | DCCS | —— | .23* | .14 | .12 | .15 | .24* | .36** |
| 2 | Age | | —— | .12 | - .05 | .17 | .30* | .20 |
| 3 | PPVT *raw* | | | —— | .98** | .27* | .16 | .30** |
| 4 | PPVT *standardized* | | | | —— | .24* | .09 | .26* |
| 5 | FDS | | | | | —— | .19 | .14 |
| 6 | BDS | | | | | | —— | .15 |
| 7 | Bear-Dragon | | | | | | | —— |

* $p < 0.05$ level (2-tailed). ** $p < 0.01$ level (2-tailed).

**Did children in modified Standard and Fantasy Planet conditions differ in meaningful ways?** The quasi-random assignment process was adopted to ensure that the modified Standard and Fantasy Planet groups would not differ in any meaningful ways on any variables other than the DCCS (see Table 1). To test this assumption, a series of one-way ANOVAs was conducted to determine whether there were any significant differences between the condition-based groups on age, PPVT, FDS, or CBQ subscales. In addition, ANCOVAs factoring out age were used to test for differences in BDS and Bear-Dragon, since these two measures were significantly (or marginally) related to age.

There were no significant differences, although the results for BDS approached significance, with the modified Standard group scoring somewhat higher than the Fantasy Planet group, $F(1, 68) = 3.28$, $p = .08$, $\eta_p^2 = .05$. These results showed no indication that the condition-based groups were significantly different from each other on any of these measures, which suggests that any difference found on the DCCS was likely due to the experimental manipulation.

**DCCS performance overview.** All children in the final sample sorted at least 5 DCCS pre-switch cards correctly, with the majority ($n = 90$) scoring a perfect 6. As expected, performance on the post-switch portion of the DCCS was not as consistently strong: 31.3% ($n = 30$) sorted none of the cards correctly, 13.5% ($n = 13$) sorted 5 out of 6 correctly, and 34.4% ($n = 33$) sorted all 6 cards correctly. The remaining 20.8% ($n = 20$) sorted 1, 2, 3, or 4 out of 6 cards correctly.

**Did DCCS score depend on any of the design factors?** A preliminary 3 (testing wave) x 2 (condition) x 2 (order) ANOVA was conducted to determine if data could be collapsed across testing wave and testing order, which were assumed to be unimportant for addressing the primary research questions. The only differences that were revealed in this analysis were differences between testing waves, and those differences were not significant after age was entered as a covariate (since the age range was expanded after the first wave of data collection), all *p*s > .05. Therefore, testing wave and DCCS order were subsequently dropped from the analyses. To determine if age group was an important factor to include in the analyses, a basic 2 (sex) x 2 (age group) x 2 (condition) ANOVA was conducted, *p*s > .05. With no significant effects involving age group, this factor was also dropped for the remainder of the analyses.

The primary analysis was therefore a 2 (sex) x 2 (condition) ANCOVA with age (in months) as a covariate. In this analysis, there was a significant Sex × Condition interaction, $F(1, 91) = 5.47$, $p = .02$, $\eta_p^2 = .06$ (see Figure 1). Simple effects analyses revealed that boys in the Fantasy Planet condition ($M = 4.17$, $SD = 2.48$, 95% CI [3.12, 5.21]) scored significantly higher on the DCCS than did boys in the modified Standard condition ($M = 2.38$, $SD = 2.45$, 95% CI [1.34, 3.41]), $F(1, 45) = 6.63$, $p = .01$, $\eta_p^2 = .13$. There was no significant difference between girls in the Fantasy Planet condition ($M = 2.88$, $SD = 2.86$, 95% CI [1.67, 4.08]) and girls in the modified Standard condition ($M = 3.46$, $SD = 2.47$, 95% CI [2.42, 4.50]). There was a trend for girls to score higher than boys in the modified Standard condition and an opposite trend (for boys to score higher than girls) in the Fantasy Planet condition, but only the difference within the Fantasy Planet condition even approached significance ($p = .08$).

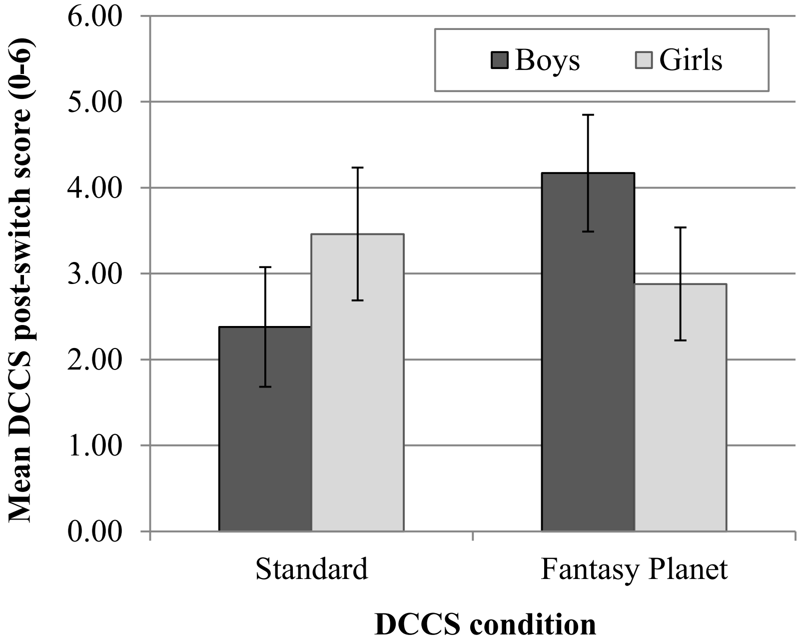



*Figure 1.*
Mean DCCS post-switch score based on condition and sex (Study 1).
Error bars represent standard errors.

Technically, ANCOVA requires a normal distribution and a continuous outcome variable, and yet DCCS scores are integer values between 0 and 6 and tend to be bimodally distributed. As is common in the literature, children in this sample tended to sort all or almost all cards in the same way: 38.5% of participants ($n$ = 37) scored 0-1 and 47.9% ($n$ = 46) scored 5-6. DCCS scores were therefore also analyzed using a non-parametric approach. First, original scores reflecting the number of post-switch cards sorted correctly were converted to pass/fail scores. The conversion formula that was adopted was provided in Zelazo's (2006) task protocol and is commonly used in the literature (e.g., Diamond et al., 2005; Hongwanishkul et al., 2005; Honomichl & Chen, 2010; Jacques, Zelazo, Kirkham, & Semcesen, 1999; Kirkham et al., 2003; Muller, Zelazo, Lurye, & Liebermann, 2008; Qu & Zelazo, 2007; Towse, Redbond, Houston-

Price, & Cook, 2000): Scores of 5 and 6 were categorized as "pass" and scores of less than 5 were categorized as "fail."

Table 3 shows the distributions of scores for each combination of sex and condition, with a vertical demarcation separating children who "passed" from children who "failed." These dichotomous data were subjected to logistic regression, which is an approach adopted by others in the field (e.g., Carlson, 2005; Kloo & Perner, 2005) and which enables analysis of the interaction term and consideration of age as a covariate. This analysis also revealed a significant Sex × Condition interaction, even with age also in the model, Wald statistic (1) = 2.11, $p = .04$.

Table 3.

*DCCS Post-switch Performance by Condition and Sex (Study 1)*

| | **DCCS Score** | | | | | | |
|---|---|---|---|---|---|---|---|
| Sex | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| | | | | **Standard Condition** | | | |
| Boys | 8 (33.33) | 5 (20.83) | 1 (4.17) | 2 (8.33) | 1 (4.17) | 2 (8.33) | 5 (20.83) |
| Girls | 6 (25.00) | 1 (4.17) | 2 (8.33) | 1 (4.17) | 2 (8.33) | 5 (20.83) | 7 (29.17) |
| | | | | **Fantasy Planet Condition** | | | |
| Boys | 5 (20.83) | 1 (4.17) | 0 (0.00) | 1 (4.17) | 1 (4.17) | 4 (16.67) | 12 (50.00) |
| Girls | 11 (45.83) | 0 (0.00) | 1 (4.17) | 1 (4.17) | 0 (0.00) | 2 (8.33) | 9 (37.50) |

*Note.* Table values reflect frequencies (with percentages in parentheses). The dashed vertical line is a pass/fail demarcation: scores less than 5 were categorized as "fail," and scores of 5 and 6 were categorized as "pass."

To summarize, both parametric and non-parametric analyses supported a difference in DCCS scores between the Fantasy Planet condition and the modified Standard condition in the hypothesized direction for boys but not for girls. Boys in the Fantasy Planet condition scored higher on the DCCS than boys in the modified Standard condition scored (and they were more likely to pass the DCCS).

**Further exploration of the boys.** As reported above, there were no differences between conditions on the variety of non-DCCS measures, but that investigation did not take sex into account. Given that on the DCCS there was a significant difference between conditions for boys but not for girls, it was worth exploring any other differences there might have been between the sexes on the rest of the measures. In addition, further exploration was warranted regarding any differences between boys in the modified Standard condition and boys in the Fantasy Planet condition because those results might inform the interpretation of the DCCS results.

As a preliminary step, a series of 2 (sex) x 2 (condition) ANOVAs was conducted to see if there were any main effects of sex or Sex × Condition interaction effects on age, PPVT, FDS, or parent-reported temperament. Parallel ANCOVAs were conducted to examine effects on BDS and Bear-Dragon, with age as a covariate. There were no significant differences between boys and girls in terms of age or the experimental tasks, although on the raw[3] PPVT scores the boys ($M$ = 83.51, $SD$ = 23.47, 95% CI [76.11, 90.92]) scored marginally higher than the girls did ($M$ = 74.78, $SD$ = 19.22, 95% CI [69.00, 80.55]), $F(1, 82) = 3.62$, $p = .06$, $\eta_p^2 = .04$. There were sex differences on all three subscales of the parent-report of temperament (very short form of the CBQ): As shown in Table 4, boys were reportedly higher on Negative Affect and Surgency, with girls reportedly higher on Effortful Control. The Sex × Condition interaction approached significance for FDS, $F(1, 82) = 3.30$, $p = .07$, $\eta_p^2 = .04$, and for Bear-Dragon, $F(1, 83) = 3.53$, $p = .06$, $\eta_p^2 = .04$, but the interaction term was not significant for any of the tested variables.

[3] Results were similar for the standardized PPVT scores.

Table 4.

*Boys' and Girls' Mean Scores on Very Short CBQ Subscales (Study 1)*

| | Boys | | Girls | | $F(1, 92)$ | $p$ | $\eta_p^2$ |
|---|---|---|---|---|---|---|---|
| **CBQ Subscale** | ***M (SD)*** | ***95% CI*** | ***M (SD)*** | ***95% CI*** | | | |
| Surgency | 4.84 (0.74) | [4.63, 5.06] | 4.42 (0.86) | [4.17, 4.67] | 6.55 | .01 | .07 |
| Negative Affect | 4.18 (0.92) | [3.91, 4.44] | 3.69 (0.80) | [3.46, 3.92] | 7.58 | < .01 | .08 |
| Effortful Control | 4.94 (0.78) | [4.71, 5.17] | 5.52 (0.62) | [5.34, 5.70] | 15.68 | < .001 | .15 |

*Note*. CI = confidence interval. Results for the main effect of sex from a 2 (sex) x 2 (condition) ANOVA are also listed.

The lack of any significant Sex × Condition interactions meant that separate examination of the boys' performance in the modified Standard condition and the Fantasy Planet condition was not statistically justified. However, given the significant difference between those two groups on the DCCS, such analyses help to tentatively test for any other differences that might contribute to the interpretation of those results. Therefore, a series of exploratory one-way ANOVAs was conducted on the data from the boys only, to test for differences between conditions (see Table 5). There were no differences or trends for CBQ subscale scores. In contrast, these tests revealed that boys in the Fantasy Planet condition scored higher on Bear-Dragon than did boys in the modified Standard condition, $F(1, 41) = 4.33$, $p = .04$, $\eta_p^2 = .10$. In addition, there were trends on the FDS and BDS in the opposite direction: Boys in the modified Standard condition tended to score higher on FDS than boys in the Fantasy Planet condition scored, $F(1, 40) = 3.05$, $p = .09$, $\eta_p^2 = .07$. Boys in the modified Standard condition also tended to score higher on BDS compared to boys in the Fantasy Planet condition, $F(1, 30) = 3.13$, $p = .09$, $\eta_p^2 = .10$.

Table 5.

*Boys' Mean Scores on Secondary EF Tasks by Condition (Study 1)*

| | Standard | | | Fantasy Planet | | |
|---|---|---|---|---|---|---|
| **Measure** | ***n*** | ***M (SD)*** | ***95% CI*** | ***n*** | ***M (SD)*** | ***95% CI*** |
| FDS | 22 | 3.86 (0.71) | [3.55, 4.18] | 20 | 3.45 (0.83) | [3.06, 3.84] |
| BDS | 17 | 1.53 (0.72) | [1.16, 1.90] | 15 | 1.13 (0.52) | [0.85, 1.42] |
| Bear-Dragon | 21 | .77 (.20) | [.68, .87] | 22 | .89 (.17) | [.82, .97] |

*Note*. CI = confidence interval.

Although more exploratory in nature, these latter results suggest possible differences between boys in the modified Standard condition and boys in the Fantasy Planet condition on the secondary EF measures. It should be noted that these analyses did not include corrections for multiple or post-hoc comparisons. More stringent tests might not yield any significant differences between boys in the two conditions. The only difference tentatively reported here was on the Bear-Dragon task, with unknown validity and reliability due to the unconventional administration.

In addition, the relevance of these patterns in the boys' data to the interpretation of the DCCS depends partially on how strongly related these EF tasks were to the DCCS. If any of the secondary EF tasks correlated strongly with the DCCS, then they could tentatively be used to measure pre-existing EF differences, although with the caveats mentioned in the discussion below. If they did not correlate with the DCCS, after partialling out age when appropriate, then these suggestions of differences between conditions for boys, again, might not be as relevant.

Recall from the original presentation of correlations for the full sample that the DCCS was only correlated with Bear-Dragon, after age was taken into account (see Table 2). This suggests that FDS and BDS were not measuring EF skills related to DCCS. Correlations for boys only, separated by condition, are presented in Table 6. As shown in the table, a positive relation between DCCS and Bear-Dragon (and between DCCS and BDS) was apparent for boys in the modified Standard condition but not for boys in the Fantasy Planet condition.

Table 6.
*Pearson r Correlation Matrix for Boys in Study 1.*

| | Measure | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | DCCS | —— | .04 | .11 | .60* | .56** |
| 2 | Age | .37 | —— | .22 | .22 | .33 |
| 3 | FDS | - .02 | .19 | —— | .26 | .32 |
| 4 | BDS | .22 | .38 | .18 | —— | .48 |
| 5 | Bear-Dragon | .27 | .14 | .28 | .17 | —— |

*Note.* Values above the diagonal apply to boys in the modified Standard condition ($n = 24$) and values below the diagonal apply to boys in the Fantasy Planet condition ($n = 24$).
* $p < 0.05$ level (2-tailed). ** $p < 0.01$ level (2-tailed).

**Summary.** Even after taking age into account, there was a difference between DCCS conditions in the expected direction for boys (based on both parametric and non-parametric analyses), but there was no similar difference for girls. There were no differences between conditions or Sex × Condition interactions on children's age, performance on PPVT or the secondary EF tasks, or the parent-report of temperament. The only significant sex differences found were on the temperament scale. Although further analyses were not statistically warranted, exploratory tests indicated that boys in the Fantasy Planet condition scored higher on Bear-Dragon than did boys in the modified Standard condition; however there were trends in the opposite direction for boys' FDS and BDS scores. In addition, the relation between the DCCS and the secondary EF tasks was inconsistent, with the pattern that was apparent at the level of the full sample not holding for boys in both conditions.

**Discussion**

The primary research question in Study 1 focused on whether a pretense-based context would facilitate better preschooler performance on a challenging EF task. Specifically, would performance on the DCCS be better for children in the Fantasy Planet condition than for children in the modified Standard condition? The next questions sought to determine whether this difference, if found, would be related to children's sex or age group and whether this difference would be significant after accounting for factors known to relate to EF (i.e., age and receptive vocabulary).

Results from both parametric and non-parametric analyses indicated support for the initial hypothesis that children in the Fantasy Planet condition would outperform children in the modified Standard condition on the DCCS; however, this support was limited to boys. There were no significant differences in DCCS performance based on age group, although scores were correlated with individual age (in months). Analyses controlled for age, and there were no significant differences between conditions or Sex × Condition interaction effects on any variables other than the DCCS. Therefore, no other tested variable can definitively account for boys' differences between conditions on the DCCS.

There was some suggestion that the boys in the modified Standard condition and boys in the Fantasy Planet condition might have differed on EF abilities, based on the secondary EF tasks. If the condition assignments did not successfully create equivalent groups for the boys, then the difference in DCCS scores might not have been due to the experimental manipulation. Although this is possible, the patterns of results for the three secondary EF measures are inconclusive. There were no Sex × Condition interactions for any of these measures, so further analyses were purely exploratory. These further analyses showed conflicting patterns, with boys in the modified Standard condition scoring lower on the Bear-Dragon but tending to score higher on the FDS and BDS, compared to boys in the Fantasy Planet condition. There is no conclusive evidence that boys in one condition were better at EF than boys in the other condition independent of the experimental manipulation.

In addition, the use of any of these three secondary EF tasks as a measure of baseline EF is not fully legitimate for a few reasons. Technically, a "baseline" measure should be administered before an experimental manipulation. In this case, that means that a baseline measure should have come before the DCCS, yet all three secondary EF tasks were presented after the DCCS. It is therefore possible that the experimental manipulation influenced children's performance on the FDS, BDS, and/or Bear-Dragon, even though those tasks were administered in the same way regardless of condition. In addition, a baseline measure should measure the same construct that is measured by the outcome variable of interest (in this case, the DCCS). While the FDS and BDS are both arguably measures of the working memory component of EF, the FDS is only an indirect measure (as described by Davis & Pratt, 1995), working memory is only one component of EF, and the DCCS involves more than working memory. Indeed, the correlations between the DCCS and FDS and between the DCCS and BDS were not significant in this sample, after controlling for age. When looking at the boys in each condition, the only significant correlation was between DCCS and BDS for boys in the modified Standard (but not the Fantasy Planet) condition. Therefore, it is not necessarily meaningful to say that the boys in the two conditions might have differed on the FDS or BDS, let alone the fact that this difference trended in the opposite direction of the DCCS findings.

Although the Bear-Dragon task correlated with the DCCS for the full sample and for boys in the modified Standard condition, there was no significant relation for boys in the Fantasy Planet condition. Also, the construct validity of the Bear-Dragon task in this experiment is suspect because it was administered in an unconventional manner. Although the pattern on this task suggests that the boys in the Fantasy Planet condition might have had better EF than the boys in the modified Standard condition had, that argument is weakened by the possibilities that the DCCS manipulation impacted subsequent Bear-Dragon performance and that Bear-Dragon in this study might not capture the same EF construct that is measured by the DCCS.

Even in the broader literature, although Digit Span (especially BDS) and Bear-Dragon are each often presented as measurements of EF or its components (e.g., inhibitory control, working memory) (e.g., Carlson, 2005; Carlson & Moses, 2001; Carlson et al., 2002; Davis & Pratt, 1995; Kochanska, Murray, Jacques, Koenig, & et al., 1996), only a few studies directly address correlations between the DCCS and FDS, BDS, or Bear-Dragon, and there are some conflicting findings (see Carlson & Moses, 2001; Matte-Gagné & Bernier, 2011; Oh & Lewis, 2008). An example of research suggesting only weak relations is recent work by Carlson, White, and Davis-Unger (2012), which showed a significant correlation between DCCS and BDS when controlling for age but then a non-significant result after partialling out additional measures. Carlson and colleagues also reported that the relation between DCCS and Bear-Dragon only approached significance with age partialled out (and was simply not significant after other variables were also taken into account). There are other suggestions, however, that these measures have been found to relate to one another. Carlson (2005) presented a scale of EF tasks that indicated their comparative difficulty for preschoolers. Her results indicated that 3-year-olds who pass DCCS are likely to also pass Bear-Dragon, since it is considered easier. In turn, Carlson showed that

BDS is more challenging than DCCS, so children who pass BDS are likely to pass DCCS and Bear-Dragon, as well. Adding to the picture of how these EF tasks interrelate, some researchers have conducted principal component analyses and presented composite EF scores that combine data from a collection of measures based on the factor loadings. For example, Beck and colleagues (2007) had a one-factor solution from their EF tasks and thus combined them all, including Bear-Dragon and DCCS. Carlson and colleagues (2002) created an Inhibitory Control battery score (including Bear-Dragon) and a Working Memory battery score (including BDS) and found a positive relation between those batteries. Despite the non-significance of correlations reported above, Carlson, White, and Davis-Unger also computed a Conflict EF composite score, including DCCS, BDS, and Bear-Dragon, among other tasks. Overall, the field often suggests that the Bear-Dragon and digit span tasks, at least in their traditional form, should correlate with DCCS. The lack of consistent correlations in the present research adds another reason why it might not be entirely appropriate to use any of the secondary EF tasks in this study to interpret the DCCS results.

Finally, to the best of my knowledge, studies testing a between-subjects manipulation of the DCCS in the literature do not tend to include any baseline EF measure (e.g., Kirkham et al., 2003; Kloo & Perner, 2005; Mack, 2007; Muller et al., 2008). The authors do not often address the possibility that their groups were unequal in EF abilities because the randomization process is assumed to have created comparable groups. The remainder of this discussion will therefore set aside that possibility for Study 1 and focus on the interpretation that the experimental manipulation in this study facilitated better DCCS performance for boys in the Fantasy Planet condition.

**Remaining questions.** The above findings support the hypothesis that the Fantasy Planet condition would facilitate better performance than would the modified Standard condition, but only for boys. Two main questions arose based on these findings, and each question will be addressed in turn below: (1) What was it about the two conditions that was responsible for the observed difference in children's (boys') performance? (2) Why was the difference observed only in boys?

***Question 1.*** Given that there was a difference between the modified Standard and Fantasy Planet conditions (for boys), was the difference due to pretense itself (i.e., cognitive mechanisms involved in pretending), or was it due to another variable that might be similarly effective if manipulated in a manner independent of pretense? Other candidates include the Fantasy Planet visual aid, the experimenter's emphasis regarding change versus sameness, and children's engagement, motivation, or mood.

*Fantasy Planet visual aid.*

The Fantasy Planet condition involved the experimenter's use of a visual aid (see Appendix B) during task instructions, yet the modified Standard condition did not involve any visual aid. In Dias and Harris's (1988) work using a pretense manipulation with a counterfactual syllogisms task, make-believe manipulations were similarly effective regardless of whether or not visual cues were present. The same researchers later found that three methods of make-believe delivery (classic make-believe intonation, a pretend setting, and imagery instructions) were no better in combination than as individual methods in terms of promoting better counterfactual reasoning, although imagery instruction was the method most strongly associated with children's better justifications to support their reasoning (Dias & Harris, 1990). These findings suggest that benefits of a pretense context in the Fantasy Planet condition for the DCCS similarly might not depend on the visual aid component; however, there are arguments to suggest that the visual aid might have supported better performance through other mechanisms unrelated to pretense.

For example, the visual aid provided an additional and non-verbal cue regarding the transition between the pre- and post-switch sections of the DCCS task. According to Mack (2007), young children might require more than the basic verbal indication of the rule change that traditionally occurs during the DCCS. In one of Mack's manipulations, the rule change was accompanied by the temporary removal and subsequent replacement of the images labeling the sorting boxes. This manipulation was associated with better DCCS performance compared to a control group. The researcher argued that the brief image removal and replacement served as a visual marker that provided a supplementary indication that one section of the game had ended and a new section was beginning. It is possible that the raising and lowering of the Fantasy Planet visual aid during the post-switch instructions in Study 1 similarly provided children in that condition with an additional and non-verbal cue that facilitated their adoption of the new rule. This visual cue was not present in the modified Standard condition.

In addition, although the experimenter referred to both color and shape during the initial introduction of the sorting boxes in both conditions (i.e., "Here is a blue star, and here is a red truck"), the visual aid in the Fantasy Planet condition might have made the existence of two dimensions more salient for children at the beginning of the task. During the pre-switch instructions, the experimenter only referred to and gestured towards the relevant planet; however, both planets were depicted on the displayed visual aid. Perhaps this helped prepare children in the Fantasy Planet condition for the existence of two planets, games, and sets of rules. It is possible that children in this condition therefore might have started with a Shape-not-Color mentality in the pre-switch, rather than a simple Shape mentality. Negative priming theories (e.g., Perner & Lang, 2001) would suggest that this would make it harder to switch (i.e., not-Color would make it harder to switch to Color later), but there is also support for an argument that pre-training emphasizing the two dimensions facilitates better performance on the task itself (Mack, 2007). Presentation of the visual aid at the beginning of the task might have implicitly provided children in the Fantasy Planet condition with a pre-training exposure to the existence of two rules. Children in the modified Standard condition received no such pre-training exposure.

The visual aid presented yet a third issue, which relates to the visual separation of task-relevant dimensions. Prior research has demonstrated that children perform better on the DCCS when the color and shape dimensions are visually separated on the test cards rather than being presented in the standard integrated format (Diamond et al., 2005; Kloo & Perner, 2005). Although the test cards used in Study 1 were the standard integrated cards, it could be argued that children performed better in the Fantasy Planet condition because that condition included separate visual depictions of the two dimensions (through the two associated planets). There were no separate visual depictions for children in the modified Standard condition. Kloo and Perner found that, although separation of the dimensions on test cards facilitated better performance among preschoolers, separation on target cards (i.e., on the boxes) had no effect. That finding suggests that the separation on the visual aid is unlikely to have been responsible for better performance in the Fantasy Planet condition here; however, it is worth acknowledging the possibility that the visual aid played a role.

Given the arguments outlined above, Study 2 was designed to rule out the possibilities that the difference between conditions (for boys) was due to the visual aid in the Fantasy Planet condition providing a visual marker cuing the rule-change, pre-training regarding two dimensions, or visual separation. By removing the visual aid in Study 2, it would be possible to determine if the condition effect was independent of those three factors.

*Emphasis on change versus sameness.*

The transition between the pre-switch and the post-switch phases of the DCCS generally involves instructions to stop playing one game and to start playing another game. There is sometimes a three-part verbal indication of a game change: (1) a declaration that a "new" or "different" game is starting, (2) a negative reference to the first game (i.e., to "not" playing that game), and (3) a specific reference to the name of the new game (which inherently refers to the newly relevant dimension). In fact, the published protocol requires the experimenter to adhere to the following script: "Now we're going to play a new game. We're not going to play the [color] game anymore. We're going to play the [shape] game" (Zelazo, 2006, p. 299). This is generally followed by an explanation of the new game's sorting rules. In their early uses of the DCCS, Frye, Zelazo, and Palfai (1995) reported slightly different scripts, but there was a similar requisite emphasis on the change when the experimenter switched between games (e.g., "The color game is different from the shape game," Experiment 1, p. 490, and "Now we are going to switch. We are not going to play the shape game any more. We are going to play the number game," Experiment 2, p. 511). Generally, when specific phraseology has been reported in studies involving the DCCS, it has included this type of emphasis on change attained either through the use of all three components cited above (e.g., Dick et al., 2005; Munakata & Yerys, 2001; Qu & Zelazo, 2007; Yerys & Munakata, 2006; Zelazo et al., 1996) or a subset [e.g., no reference to a "new game" (Diamond et al., 2005; Jacques et al., 1999; Kirkham et al., 2003; Muller et al., 2008) or no negative reference to the old game (Kloo & Perner, 2005; Kloo, Perner, Aichhorn, & Schmidhuber, 2010; Mack, 2007; Marcovitch, Boseovski, & Knapp, 2007; Perner & Lang, 2002)].

Both conditions in Study 1 involved the three-part reference to the change in the game; however, when other elements of the two scripts are taken into account, there is an imbalance in the emphasis on change versus on sameness. Thus, another alternative explanation for the (boys') performance difference between the Fantasy Planet and modified Standard conditions in Study 1 is that the instructions in the former condition further emphasized change while the instructions in the latter condition ultimately emphasized sameness. Specifically, in the Fantasy Planet condition, the experimenter referred to taking a rocket ship to a "different faraway planet" when it was time to transition to the second rule. Therefore, in addition to the incorporation of the three-part references to a different game, this script involved explicit references to two different entities (i.e., planets) in association with the two different games. In the modified Standard condition, however, although there was a three-part reference to playing a different game, the experimenter also inadvertently emphasized sameness in two ways: (1) by referring to learning the second game from the same friend who taught her the first game and (2) by suggesting that she and the participant continue to play on the same table where they were already set up. Thus, in the modified Standard condition, there was no set of two different entities for children to associate with the two different games that would parallel the two planets in the pretense condition. Instead, both games were associated with the same friend and the same location. This semantic difference between the conditions' scripts was an unintentional side effect of an effort to provide insignificant filler to ensure that the modified Standard condition instructions were similar in duration to those of the Fantasy Planet condition. The distinction between an emphasis on change versus an emphasis on sameness and its potential significance was identified only in retrospect and was later addressed in Study 2.

*Engagement, motivation, and mood.*

If the difference between children's (boys') scores in the Fantasy Planet condition and children's (boys') scores in the modified Standard condition was related to the inclusion of pretense, it is still unclear whether it was due to cognitive underpinnings of pretense itself or another characteristic of pretense. For example, it is possible that the pretense context caused the task to be more engaging for children in the Fantasy Planet condition and that a non-pretense manipulation that similarly increased engagement would have a similar effect. There have been suggestions in the literature that the incorporation of fantasy into children's tasks might facilitate interest and motivation (e.g., Bergin, 1999; Matheson & Spranger, 2001; Parker & Lepper, 1992) and that the relative attractiveness of different tasks might influence performance (e.g., Ionescu, 2012; Nguyen, Kemp, & Want, 2011; Parker & Lepper, 1992). Thus, the pretense manipulation might have led to differing levels of motivation or interest associated with each condition, and that might have had an impact on (boys') DCCS performance. Unfortunately, it is difficult to behaviorally measure these nebulous constructs of engagement and motivation, and it is equally difficult to design a non-pretense control condition that matches the Fantasy Planet condition on the degree of the engagement and motivation that it invites.

Similarly, the mood associated with the task might have been affected by the experimental manipulation, and that could have had implications for children's (boys') performance. Qu and Zelazo (2007) reported that the use of stimuli with positive affective valence (i.e., happy faces) was related to better DCCS performance than the use of standard stimuli, whereas performance with sad or neutral stimuli was not significantly different from performance with standard stimuli. This positive emotion effect was apparent regardless of whether the stimuli were sorted based on emotion or the emotion was a contextual factor, with cards sorted on unrelated features (e.g., sex and age). The authors argued that positive mood has been linked to improved cognitive flexibility and children's positive mood might have been elevated as a result of the happy faces. Similarly, although Study 1 did not involve any manipulation of the sorting cards themselves, it is possible that the Fantasy Planet context created more positive affect for children (or at least boys) than the modified Standard condition elicited. The contextual description of the modified Standard condition script was arguably more neutral (i.e., referring in passing to a friend and to the playing space on the table).

As is the case with the engagement argument, it is difficult to address the mood argument without sophisticated measurements of children's moods and the design of a non-pretense control condition that facilitates a comparable mood. However, as will be discussed in a later section, some of these issues might be addressed by creating pretense-based and control conditions that involve themes that might be similar to each other in terms of their attractiveness.

***Question 2.*** Regardless of the reasons behind the performance difference between the conditions in Study 1, the difference was apparent; however, the difference only existed for boys. Why was the manipulation effective for boys but not for girls? This is not the only research showing no evidence of a relation between pretend play and EF in girls: A recent study with only female participants also found no support for the relation (Hoffmann & Russ, 2012). However, it is unclear whether sex differences in Study 1 and the null findings in Hoffman and Russ's work were due to valid differences between the sexes or to aspects of each study's design.

*Sex differences on various measures.*

The Sex × Condition interaction might relate to sex differences that are unrelated to the manipulation itself. For example, on the parent-report of temperament, boys in this sample scored higher on Negative Affect and Surgency and lower on Effortful Control than did girls. Hongwanishkul and colleagues (2005) used the CBQ (195 statement version) with preschoolers in their research, and they reported the same sex difference on Effortful Control. This subscale of the CBQ has been reported to relate to performance on various EF tasks (Beck et al., 2007). The patterns found here for both Effortful Control and Surgency were also found in a meta-analysis of sex differences in temperament that incorporated studies involving a variety of measures (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006). When sex differences in EF are reported in the literature, it is often (but not always) the boys who struggle compared to the girls (Carlson & Moses, 2001; Carlson & Wang, 2007), so while there are important distinctions between Effortful Control and EF (see Liew, 2012; Zhou et al., 2012), the current CBQ findings appear consistent with a trend for boys to struggle more with self-control (even though there were no other sex difference on EF measures). It is possible that that the Fantasy Planet manipulation facilitates better performance for children with lower self-control abilities (who tended to be boys in this study). Future work with a more diverse sample (e.g., broader range of SES) might enable EF abilities and sex to be less intertwined, which would allow for more precise conclusions to be drawn.

*Sex-related biases in study design.*

Another possible underlying explanation for the sex difference in the results from Study 1 derives from the content of the pretense-based manipulation. In the literature on self-generated themes in preschoolers' pretend play, researchers have reported sex differences in type and content, with boys tending to choose themes or roles that are more fantastic (e.g., superheroes) and girls tending to choose themes or roles that are more reality-based and perhaps anticipatory of their own future roles (e.g., housekeeping) (Black, 1989; Connolly & Doyle, 1984; McLoyd, 1980; McLoyd, Warren, & Thomas, 1984). Given that the script for the Fantasy Planet condition described taking an imaginary rocket ship to imaginary planets, the content could be considered more masculine due to the focus on a more fantastic, not-so-everyday, activity and setting.

Aside from the fantastic nature of the pretense, the specific focus on outer space could be considered gender-typed, as well. In the literature focusing on sex differences in young children, researchers have categorized themes relating to outer space as "for boys" (Pulaski, 1973) and discussed empirical findings demonstrating that related themes (e.g., "aliens") and items (e.g., "space helmet" as a dress-up prop) are more common choices amongst boys than amongst girls during playful activities and story-telling (Boyatzis & Eades, 1999; Maccoby, 1998; Pulaski, 1973). These specific references to outer space provide further support for the argument that the Fantasy Planet condition in Study 1 was likely to be more engaging for boys than for girls.

*Girls had less room for "improvement."*

It is also possible that the Sex × Condition interaction was related to the girls' strong performance overall. The range of possible scores on the DCSS post-switch phase is only 0 to 6. In the modified Standard condition, girls scored high enough ($M$ = 3.46, $SD$ = 2.47) that the girls in the Fantasy Planet condition would have had to perform very well (an average of approximately 4.88 out of 6, with many high scorers) to create a difference that was statistically significant. While it was statistically possible, this required the manipulation to raise girls to a fairly high level for their age. For the boys, performance in the modified Standard condition ($M$ = 2.38, $SD$ = 2.45) was low enough that there was more room for the boys in the Fantasy

Planet condition to score significantly higher. This suggests that a condition effect might be apparent for girls at a younger age, when their performance in the modified Standard condition would be lower than it was in Study 1, or with a task that is more challenging than the basic DCCS. However, it should also be noted that the trend in results for Study 1 was for girls in the Fantasy Planet condition to perform worse than did the girls in the modified Standard condition.

Thus, overall, there were a number of issues remaining after Study 1 that required further investigation.

## Chapter III: Study 2

### Overview

Study 2 was another between-subjects design with a DCCS manipulation that compared a pretense-based condition to a control condition. This study was designed to address the two main questions that grew out of the results from Study 1. The primary focus was on question 1: What element of the Fantasy Planet manipulation in Study 1 was responsible for the condition effect (for boys)? Was the manipulation effective because of the pretense component or because of something else? Specifically, the second study addressed the possible alternative explanations of the visual aid and the verbal emphasis on change versus sameness. To a lesser extent, this study also addressed the possibility that the manipulation was effective because it increased engagement, motivation, and/or mood (particularly for boys) through the use of the outer space theme.

The visual aid from the Fantasy Planet condition was removed for the new pretense condition, and the content of the pretense was changed to a more gender-neutral theme: Pretend Playground. In addition, a carefully designed Non-Pretense Control condition was implemented. The script for this new condition contained references to change (and not to sameness) that were similar to the references in the scripts for the Fantasy Planet condition in Study 1 and for the Pretend Playground condition in Study 2 (see Appendix E for examples of Study 2 scripts). In this new Non-Pretense Control condition, the experimenter taught the participants the game that one friend taught her (for the pre-switch) and then taught them another game that her other friend taught her (for the post-switch).

If the incorporation of a pretense context facilitates better performance on the DCCS (and the specifics of the pretense are not important) then children in the Pretend Playground condition should score higher than children in the Non-Pretense Control condition score. (Possible sex differences are addressed below.) In contrast, if it is not the incorporation of a pretense context but the inclusion of a visual aid or the verbal emphasis on change that facilitates better DCCS performance then children in the two new conditions should score similarly (because neither condition involves a visual aid and the scripts are matched in terms of an emphasis on change).

The second study also addressed question 2: Why was the condition effect present for boys but not for girls? This issue was primarily addressed by replacing the outer space theme for the pretense-based condition with a more gender-neutral theme. If the gender biases of the pretense content in Study 1 were responsible for the condition effect appearing only for boys then this more neutral theme might eliminate the sex difference. This could lead to a condition effect for both sexes (or neither). As indicated above, the possibility that differences in engagement, motivation, and/or mood were responsible for the condition effect appearing for boys was also tentatively addressed by this change in thematic content. The new conditions were designed to remove potential gender biases and to be equally engaging across sex. If the Fantasy Planet condition facilitated better performance for boys because the masculine thematic content elevated their engagement, motivation, and/or mood (and not because of pretense itself) then there should be no significant difference for boys between the two gender-neutral conditions in Study 2. Alternatively, if there is a facilitative effect of pretense that applies only to boys (but not to girls) regardless of the thematic content of the pretense--a genuine sex difference in the relation between pretense and EF--then a Sex × Condition interaction should be replicated in Study 2.

Through other changes, Study 2 also addressed the possibility that girls in Study 1 were approaching ceiling level in their performance on the DCCS. The oldest age (47 months) that was included in Study 1 was excluded for Study 2, which also had a narrower age range overall. Because DCCS performance was related to age in the first study, excluding the oldest children was a way to reduce the highest performance levels. Also, the possibility of a ceiling effect was addressed in a more exploratory manner through the administration of Level 5 of the new EF Scale for Preschoolers (Carlson, 2012). The EF Scale involves a collection of variations on the traditional DCCS task, with each level designed to be more challenging than the prior level. The DCCS task generally used in the literature and embraced for Study 1 and for the main focus of Study 2 is equivalent to Level 4 of the EF Scale. By also incorporating Levels 3 and 5, it was anticipated that Study 2 would capture more of the sample's variability in EF skills and eliminate the possibility of a ceiling effect.

In addition, Study 2 included the PPVT and secondary EF tasks, for control and exploratory purposes, respectively. For the sake of consistency with Study 1, two of the same EF tasks were adopted again: FDS and Bear-Dragon. The BDS was not included because it proved to be too difficult for most children in Study 1, and the concern about a floor effect is supported in the literature (Davis & Pratt, 1995). To address some of the issues raised by Study 1, FDS and Bear-Dragon were administered prior to the DCCS and in a consistent and traditional manner throughout data collection. Although some of the caveats from Study 1 still apply, it was assumed that the FDS and Bear-Dragon data could again tentatively inform the interpretation of any group differences on the DCCS. The same approach was taken with the CBQ.

In sum, a condition effect (for boys, girls, or both), after controlling for age and receptive vocabulary, would support the hypothesis that the inclusion of pretense facilitates better performance on the DCCS even when there is no visual aid for the pretense manipulation and the emphasis on change versus sameness is experimentally controlled. Also, if there is no Sex ×

Condition interaction when the thematic content of the pretense is gender-neutral, this would support the hypothesis that the Study 1 findings were due to the gender-biased nature of the pretense manipulation.

**Method**

**Participants.** Exploration of the data from Study 1 suggested that the most promising subset of ages to pursue for the second study was 43 to 46 months old. For this subset of children in Study 1, there were no correlations between DCCS and other tasks or between DCCS and age, and DCCS results mirrored those of the full sample: There was a significant Sex × Condition interaction effect with a significant simple effect for boys and a non-significant effect for girls, as well as non-significant effects for the differences between sexes within each condition. This narrower age range excluded the oldest children (47-month-olds), which partially addressed the concern that girls in Study 1 were approaching ceiling on the DCCS.

The final sample for Study 2 included 64 participants ($M_{age}$ = 44.38 months, *SD* = 1.05 months), half of which were boys. An additional 15 children were excluded due to health concerns or developmental delays that parents reported after participation had begun (*n* = 3) or due to problems regarding the main level of the DCCS, including: child refusal (*n* = 4), parental interference (*n* = 1), procedural error (*n* = 5), and equipment failure (*n* = 2). As this was another between-subjects design, children were randomly assigned (based on sex) to condition (Non-Pretense Control or Pretend Playground) and dimension order (shape-color or color-shape). Cell membership was counterbalanced with one exception[4]; otherwise, there were 8 boys and 8 girls in each cell.

---

[4] One boy in the Pretend Playground condition was accidentally given the color rule first when he was supposed to be given the shape rule first (i.e., there were 16 boys in the Pretend Playground condition, and 9 of them were taught color before shape).

As in Study 1, recruitment was generally through the UW participant pool, consisting mainly of middle- to upper-middle-class families living in the greater Seattle area. Most participants' parents identified the children's racial background as White (78.1%) or multiracial (17.2%) with the remaining parents selecting Asian (1.6%) or other (1.6%). One parent left this question blank (1.6%). As with Study 1, the majority of participants were identified as non-Hispanic (92.2%). The exclusion criteria were the same as those described for Study 1.

**Procedure.** The procedure for this second study was similar to the procedure for Study 1, except that there were fewer tasks involved and the administration order was different. The testing session lasted approximately 25-40 minutes and took place at a child-friendly lab at UW-Seattle. Again, the primary investigator was the experimenter for all children. Tasks (in order of administration) included: PPVT-IV, Bear-Dragon, FDS, and DCCS. To supplement descriptions provided for Study 1, any additional information about each task is provided below. The parent questionnaires that were used in Study 1 were again used in Study 2, with the added option for parents to complete these questionnaires via a secure online survey before their lab visit. Compensation for participation was the same as it was in Study 1 with the addition of reimbursement for newly implemented toll expenses.

**Measures.**

*PPVT-IV.*

This measure of receptive vocabulary was the same in both Study 1 and Study 2.

*Bear-Dragon.*

This task was administered in the traditional manner, without the parental involvement that was included in the initial waves of Study 1. Extra steps were taken to ensure that most if not all trials were codeable (e.g., making sure that the child was not already touching a given body part at the time the puppet started to instruct the child to do so). The procedure was otherwise the same as it was in Study 1. Half of the participants were double-coded, and reliability was strong overall. Coders agreed on all 32 children's scores on three of the bear trials (with all children scoring a 3 on these trials). On the remaining two bear trials, coders had three disagreements. Cohen's kappa values for those two bear trials were .48 and .79, which were acceptable given that there was very little variability, with the majority of children scoring a 3. The three disagreements on the bear trials led to a Cohen's kappa value of .54 for the total bear score. As for the dragon trials, coders' agreement was strong, with only 5 disagreements on individual trials (out of 155 total trials). Cohen's kappa values for the 5 dragon trials ranged from .84 to 1.00. Agreement on the total score for dragon trials was acceptable (Cohen's $\kappa = .78$). Since there were some trials that were not codeable (e.g., due to the camera getting moved temporarily), total proportion scores were again used for this task. These proportion scores expressed the number of points (0-3 per trial) the child received out of all the possible points the child could have received on the codeable bear and dragon trials (i.e., out of 30 if all 10 trials were codeable).

*FDS.*

The administration of the FDS was the same for both studies. As a reliability check, half of the participants were double-coded, and the reliability was strong (ICC = .98).

*DCCS.*

The experimenter began the task by introducing the child to the two boxes: one with a blue star label and one with a red truck label. The script for the DCCS depended on the child's condition and dimension order (see Appendix E). For children in the Pretend Playground condition, the experimenter suggested that children pretend to be at a faraway playground. With verbal description only (i.e., without the use of a visual aid), the experimenter instructed children to pretend that they have driven in their car "all the way over to the Shape/Color Playground," where they "see lots and lots of shapes/colors everywhere, like stars/blue and trucks/red." The experimenter then explained the rules of the associated card game (e.g., "All the stars go here, and all the trucks go there"). As in Study 1, the experimenter demonstrated how to sort one card (a red star) and then administered a practice trial in which the child was asked to sort one card (a blue truck). There were then six pre-switch trials, with rule reminders preceding all trials and a game reminder before Trial 4. The experimenter labeled each card by the appropriate dimension as she handed it to the child, and she gave no feedback regarding the child's sorting. To introduce the post-switch game, the experimenter suggested that the child pretend to be at a "different faraway playground." She asked the child to pretend to have driven their car "all the way over to the Color/Shape Playground," where they "see lots and lots of colors/shapes everywhere, like blue/stars and red/trucks." The experimenter then explained the sorting rules for the new game. Six post-switch trials followed, structured in the same way as the pre-switch trials.

The basic outline of the DCCS was the same for the Non-Pretense Control condition, but the details differed: The experimenter introduced the Shape/Color game as a game that she learned from a certain friend and that she wanted to teach to the child[5]. During the transition to the post-switch game, the experimenter explained that they would be playing a game that she learned from a certain "different friend," which she would then teach the participant. The experimenter's friends' names were Jonathan/Julia (shape) and Samuel/Samantha (color), with the friends'

[5] The Study 1 modified Standard condition's scripted references to playing on the table were excluded from the Non-Pretense Control script for Study 2. This was to address the possible alternative explanation in Study 1 involving the conditions' scripts' emphases on change versus sameness.

gender matched to that of the participants. Thus, one game was called "Jonathan's Shape Game," while the other was called, "Samuel's Color Game" (for example).

Scripts for the different conditions were again designed to be parallel in form and similar to each other in duration and frequency of words regarding the sorting dimensions. The experimenter's behavior was neutral, as in Study 1, but with an added component: The experimenter recorded the child's sorting responses during the session. This in-session coding was included to facilitate exploratory administration of adapted versions of Levels 3 and 5 of the EF Scale for Preschoolers (Carlson, 2012), which is described in further detail below. All participants' videos were subsequently coded without sound by an additional coder who was blind to condition. The two coders agreed on 63 children's pre-switch scores (6 out of 6), but they disagreed on whether the remaining child's score was a 5 or a 6, Cohen's $\kappa = .66$. Children's post-switch performance (0-6 total score for correct sorts or fail/pass score, with scores of 5 and 6 considered to be "passing") was the outcome of primary interest. For the individual post-switch trial scores, there was complete agreement for 3 trials (Cohen's $\kappa = 1.00$ for each) and the remaining 3 trials also had strong reliability, Cohen's $\kappa = .93$ to $.97$. The total post-switch score (Level 4), was the main variable used in analyses, and for this there was strong reliability: Coders agreed on all but 3 children's total scores, Cohen's $\kappa = .92$. A naïve third coder (watching with sound but ignorant of the purpose and background for this study) resolved all discrepancies.

As indicated earlier, additional levels of the DCCS-based EF Scale for Preschoolers (Carlson, 2012) were incorporated into Study 2 in an exploratory fashion. The EF Scale involves seven levels of difficulty, with Level 4 being the most similar to the traditional DCCS task as outlined by Zelazo's (2006) protocol. The scale manual dictates at which level an experimenter should start, based on the participant's age. Then, the experimenter moves up or down, administering harder or easier levels when children pass or fail, respectively. For children ages 43 to 46 months, the official starting place is Level 3; however, because the main focus of this study was a continuation of the investigation in Study 1, which manipulated the traditional DCCS, all children started on Level 4. After this portion was over, children were moved up or down one level, as a more exploratory piece of the research.

Therefore, children who failed the main DCCS described above (Level 4) were administered Level 3 of the EF Scale, with adaptations made to the scripts to fit the conditions established in this study (See Appendix F for sample scripts). Although the instructions for this level were briefer than those for Level 4, they included references to the same condition-based premises (i.e., pretending to drive to the Shape Playground or to playing the experimenter's friend Jonathan's/Julia's Shape Game). Level 3 of the scale involves cards depicting black shapes on colored backgrounds, which are considered to be easier to sort due to the separation of the shapes and colors (as opposed to the integration of shape and color into the same picture on the cards in Level 4) (Diamond et al., 2005; Kloo & Perner, 2005). For Level 3, the experimenter replaced the blue star and red truck box labels with a black heart on a pink background and a black flower on a yellow background, respectively. She re-introduced the children to the two boxes, using the new labels associated with the current dimension, and showed them examples of the sorting cards, labeling them by the appropriate dimension, as well. In keeping with the scale manual, the pre-switch dimension was always matched to the dimension used in the recent post-switch section of Level 4, which the child previously failed. The experimenter demonstrated by sorting one of each of the test cards (black hearts on yellow backgrounds and black flowers on pink backgrounds). She then asked the child two rule check questions: (1) "Can you show me where the flowers/pink ones go in [name of game]?" and (2) "Can you show me where the hearts/yellow ones go in [name of game]?" If the child answered incorrectly, the experimenter corrected the child up to two times and then proceeded with 5 pre-switch trials. If the child sorted at least 4 cards correctly, the experimenter explained the rules for the other dimension and proceeded with 5 post-switch trials. If the child failed the pre-switch section (i.e., sorted fewer than 4 cards correctly), the task was terminated prior to the post-switch section. Cards were given in a fixed order for each section, with 2-3 cards of each type per section and no more than two similar cards in a row.

Children who passed the main DCCS (Level 4, sorting 5 or more of the 6 post-switch cards correctly) were administered Level 5, which introduces more challenges. For this level, the experimenter explained that she would hand the child a card and instruct the child about which game to play on that trial. The script's references to each game were based on the child's assigned condition (See Appendix F). The experimenter demonstrated by first stating the shape game and its rule and sorting one of each type of card by shape and then stating the color game and its rule and sorting one of each type of card by color. There were four rule-check questions in which the experimenter asked the child what game the child should play given each of two condition-based commands from the experimenter (e.g., "Play Jonathan's/Julia's Shape Game" and "Play Samuel's/Samantha's Color Game" or "Pretend we're at the Shape Playground" and "Pretend we're at the Color Playground") and how to play each of the two games. As with the rule check for Level 3, the experimenter corrected the child up to two times before proceeding. There were then 10 trials (5 shape, 5 color) in which the experimenter held out a card and gave a command regarding which game the child should play for that trial. Equal numbers of the two types of cards were administered in a fixed order and with a fixed command on each trial.

As with the main DCCS level (Level 4), the experimenter coded Levels 3 and 5 during the session, and all videos were later coded by a second coder who coded on mute (i.e., was blind to condition). Coders were in complete agreement for all scores on Level 3 (pre- and post-switch), Cohen's κ's = 1.00. For Level 5, there was complete agreement on all 5 color trials and on 3 out of 5 shape trials, Cohen's κ's = 1.00. There was good agreement on the remaining 2 shape trials, as well, Cohen's κ = .88 and .94. Overall, coders disagreed on only 3 children's Level 5 scores (disagreeing on one shape trial for each child), Cohen's κ = .89 (total shape) and 1.00 (total color). Again, a third coder, who was naïve to the purpose of the study and who coded with sound, resolved the few discrepancies.

*CBQ.*

Parents completed the same measures that were used in Study 1, including the very short form of the CBQ (in addition to the parent questionnaire regarding demographics, pretense, and EF that will not be discussed here any further).

**Results**

All 64 participants in the final sample for Study 2 completed the main level of the DCCS, and the usable[6] sample sizes for the other measures were as follows: 60 for PPVT, 61 for Bear-Dragon, 62 for FDS, and 54 for the exploratory portions (Levels 3 and 5) of the DCCS. All children had CBQ data. As in Study 1, children without usable scores on a specific measure were excluded from analyses involving that measure.

---

[6] Most cases of missing data were due to child refusal to complete the particular task. Other reasons were failure to pass the rule-check (e.g., for dragon trials on Bear-Dragon) or procedural error (e.g., on the more exploratory levels of the EF Scale, as the experimenter adapted to the differences in administration compared to the previously administered DCCS).

The primary measure in this study was the DCCS (Level 4). Age and PPVT score were examined as possible covariates, and those two variables, as well as Bear-Dragon, FDS, and CBQ scores, were used to confirm that there were no significant pre-existing differences between the two conditions that might have impacted the DCCS results. In addition, Levels 3 and 5 of the DCCS were examined on a more exploratory basis. As the main interest in this study was the comparison of the condition-based groups, Table 7 shows average scores for each condition on each measure.

As in Study 1, DCCS performance on Level 4 was analyzed using both post-switch scores (0-6 out of 6) and dichotomous pass/fail scores (children passed if they sorted at least 5 out of 6 cards correctly on the post-switch). For the exploratory analyses of other DCCS levels, scores were converted to a "highest level passed" on the DCCS (scores ranged from 2 to 5), as is recommended in the EF Scale manual (Carlson, 2012). For levels 3 and 5, children had to sort at least 4 cards (out of 5) correctly for each of the two dimensions in order to be categorized as passing that level. Children who failed Level 3 ($n$ = 9, including 2 children who failed the pre-switch phase) were given a default score of 2, although Level 2 of the DCCS was not actually administered to participants in this study. On the opposite end of the spectrum, children who passed level 5 were granted a score of 5, since it was the highest level included in this study, but it should be kept in mind that some of these participants might have been able to pass at a higher level, too.

Table 7.

*Descriptive Statistics for Each Condition for Study 2.*

| | Non-Pretense Control | | | Pretend Playground | | |
|---|---|---|---|---|---|---|
| **Measure** | ***n*** | ***M (SD)*** | ***95% CI*** | ***n*** | ***M (SD)*** | ***95% CI*** |
| Age | 32 | 44.19 (1.00) | [43.83, 44.55] | 32 | 44.56 (1.08) | [44.17, 44.95] |
| DCCS post-switch (0-6) | 32 | 3.78 (2.69) | [2.81, 4.75] | 32 | 3.81 (2.83) | [2.79, 4.83] |
| DCCS highest level passed (2-5) | 27 | 3.59 (0.93) | [3.22, 3.96] | 27 | 3.52 (0.94) | [3.15, 3.89] |
| PPVT (raw) | 30 | 80.63 (20.86) | [72.84, 88.42] | 30 | 78.43 (18.07) | [71.69, 85.18] |
| PPVT (standardized) | 30 | 119.53 (15.59) | [113.71, 125.35] | 30 | 117.03 (13.51) | [111.99, 122.08] |
| Bear-Dragon | 31 | .92 (.15) | [.86, .97] | 30 | .89 (.19) | [.81, .96] |
| FDS | 31 | 3.77 (1.00) | [3.41, 4.14] | 31 | 3.94 (.85) | [3.62, 4.25] |
| Surgency (CBQ) | 32 | 4.60 (0.90) | [4.27, 4.92] | 32 | 4.67 (0.82) | [4.37, 4.96] |
| Negative Affect (CBQ) | 32 | 3.73 (0.64) | [3.50, 3.96] | 32 | 3.88 (0.72) | [3.62, 4.14] |
| Effortful Control (CBQ) | 32 | 5.33 (0.74) | [5.06, 5.59] | 32 | 5.29 (0.53) | [5.11, 5.48] |

*Note.* CI = confidence interval.

**Relations between DCCS scores and other variables.** Correlations among variables are presented in Table 8. Of utmost importance are the relations between DCCS (Level 4) and possible covariates: age (months) and PPVT. As predicted based on the small age range selected for Study 2, there was no significant correlation between children's DCCS performance and their age. The correlations between children's post-switch scores on the main DCCS (Level 4) and their PPVT scores (raw and standardized) only approached significance. However, the highest level passed on the DCCS (considering performance on levels 3 through 5) was significantly positively related to (raw and standardized) PPVT scores. Similar results were found through the non-parametric approach to these correlations (Spearman rho). Therefore, DCCS analyses were run both with and without standardized PPVT scores as a covariate.

A secondary purpose in examining the correlation matrix was to see the relations between the DCCS and the secondary EF tasks: Bear-Dragon and FDS. The correlations between DCCS and Bear-Dragon were strong, which is consistent with the argument that they measure similar or related aspects of EF. In contrast, FDS scores were not significantly correlated with either the DCCS or Bear-Dragon. These results suggest that Bear-Dragon would be more informative than FDS regarding interpretation of any differences on the DCCS. None of the CBQ subscales was significantly correlated with DCCS performance, *p*s > .05.

Table 8.
*Pearson r Correlation Matrix for Study 2.*

| | Measure | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 1 | DCCS *post (0-6)* [a] | ——— | .87** | - .02 | .21 | .21 | - .08 | .36** |
| 2 | DCCS *highest level (2-5)* [b] | | ——— | .05 | .30* | .31* | .05 | .33* |
| 3 | Age | | | ——— | .30* | .21 | .06 | - .02 |
| 4 | PPVT *raw* | | | | ——— | .99** | .36** | .40** |
| 5 | PPVT *standardized* | | | | | ——— | .37** | .41** |
| 6 | FDS | | | | | | ——— | .11 |
| 7 | Bear-Dragon | | | | | | | ——— |

[a] = Scores on DCCS Level 4. [b] = Highest level passed, based on testing on Levels 3-5.
† $p < .10$ (2-tailed). * $p < 0.05$ level (2-tailed). ** $p < 0.01$ level (2-tailed).

**Did children in Non-Pretense Control and Pretend Playground conditions differ in meaningful ways?** The random assignment process should have ensured that the children in the Non-Pretense Control condition were not significantly different from those in the Pretend Playground condition on non-DCCS measures. To test this, univariate and multivariate analyses were conducted to examine group differences in age, PPVT, FDS, Bear-Dragon, and CBQ. All results were non-significant, confirming that children in the two conditions were similar to one another (see Table 7). This suggests that any differences found between the groups on DCCS performance can be attributed to the experimental manipulation.

**DCCS performance overview.** The main focus of Study 2 was on Level 4 of the DCCS. For that level, all children sorted all 6 pre-switch cards correctly, and the distribution of scores for the post-switch portion showed the bimodal pattern that is common for this task: 31.3%

($n$ = 20) sorted none of the cards correctly and 53.1% ($n$ = 34) sorted all 6 cards correctly. An additional 7.8% ($n$ = 5) passed by sorting 5 cards correctly. The remaining five children (7.8%) failed by sorting only 1, 2, 3, or 4 cards correctly.

The more exploratory levels of the DCCS had usable data for 54 children. Of the 22 children brought down to Level 3, 9 children also failed this easier level (and were granted a default score of 2), while the other 13 children were able to sort well on both the pre-switch and the post-switch phases and earn a score of 3. Of the 32 children brought up to Level 5, 7 children also passed this harder level (scored 5), while it was too challenging for the other 25 children (who scored 4).

**Did DCCS score depend on any of the design factors?** The main analysis examined DCCS post-switch scores (Level 4) with a 2 (sex) x 2 (condition) x 2 (order) ANOVA. There were non-significant results for all factors and interactions, and this was paralleled in the ANCOVA that incorporated standardized PPVT score as a covariate, all *p*s > .10. In particular, the difference between condition-based groups was non-significant, ANOVA $F(1, 56) < 0.01$, $p = .99$, $\eta_p^2 < .001$ and ANCOVA $F(1, 51) = 0.24$, $p = .63$, $\eta_p^2 = .01$.

Given the bimodal nature of the results (and the fact that DCCS post-switch score is not a strictly continuous variable), a non-parametric approach was also adopted, based on the pass/fail scores for Level 4 (see Table 9). Logistic regression results were consistent with the parametric results, showing no significant main effects of sex, condition, or order, and no significant interaction effects, all *p*s > .10. Most relevant, the main effect of condition was non-significant both with and without standardized PPVT as part of the model, Wald statistic (1) = .14, $p = .89$ and Wald statistic (1) = .26, $p = .79$ respectively.

Table 9.

*DCCS Level 4 Post-switch Performance by Condition (Study 2)*

| | DCCS Score (Level 4) | | | | | | |
|---|---|---|---|---|---|---|---|
| Condition | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| Non-Pretense Control | 9 (28.13) | 1 (3.13) | 1 (3.13) | 1 (3.13) | 1 (3.13) | 3 (9.38) | 16 (50.00) |
| Pretend Playground | 11 (34.38) | 0 (0.00) | 0 (0.00) | 0 (0.00) | 1 (3.13) | 2 (6.25) | 18 (56.25) |

*Note.* Table values reflect frequencies (with percentages in parentheses). The dashed vertical line is a pass/fail demarcation: scores less than 5 were categorized as "fail," and scores of 5 and 6 were categorized as "pass."

Using the data from Levels 3 and 5 of the EF Scale, a similar question was explored: Did the highest level passed[7] depend on condition (or any of the other design factors)? I re-ran the same ANOVA and ANCOVA analyses used above but this time with the highest level passed (range: 2 to 5) as the outcome variable of interest. There was no significant difference between the highest level passed in the Non-Pretense Control group and in the Pretend Playground group (see Table 7), ANOVA $F(1, 46) = 0.14$, $p = .71$, $\eta_p^2 < .01$, ANCOVA $F(1, 43) = 0.62$, $p = .44$, $\eta_p^2 = .01$. There were no significant effects involving other design factors either, all *p*s > .10.

**Summary of primary results.** In Study 2, there was no evidence to suggest a relation between children's performance on the DCCS (whether measured by post-switch score, fail/pass score, or the exploratory score reflecting the highest level passed) and the condition to which they were randomly assigned. In fact, there was no relation between DCCS performance and any of the design factors (i.e., sex, condition, or order).

[7] Similar analyses were also performed using total scores for Levels 3-5, based on correspondence with the EF Scale developer (S.M. Carlson, personal communication, August 27, 2012). Total scores were calculated by summing the number of cards that children correctly sorted over the course of the two tested levels. Children who passed Level 4 also received full credit for all 10 trials at the easier Level 3, under the assumption that they would have performed well had they been tested on that level. Therefore, children who passed Level 4 could score between 20 and 32 while children who failed Level 4 could score between 0 and 20. The ANOVA and ANCOVA based on total scores yielded no significant effects.

**Exploratory cross-study comparisons.** Given the similarities and relation between Study 1 and Study 2, I also conducted some exploratory analyses that compared conditions across studies. For comparison purposes, I selected a subset of children from Study 1 ($n$ = 49) to match the age range of the 64 children in Study 2 (i.e., 43-46 months old). Using data from these 113 children and four conditions, a 2 (sex) x 4 (condition) x 2 (order) ANOVA revealed no significant differences. When the pass/fail data were subjected to logistic regression, there were again no significant results.

Due to the significance of the Sex × Condition effect in Study 1, I initiated further analyses of the condition effects for each sex on a purely exploratory basis, given the fact that it was not statistically warranted. DCCS post-switch performance for boys and girls in each condition is displayed in Figure 2 and Table 10. The 4 (condition) x 2 (order) ANOVA for 43- to 46-month-old boys showed an omnibus effect of condition, $F(3, 49) = 3.01$, $p = .04$, $\eta_p^2 = .16$. Next, I conducted post-hoc pair-wise comparisons implementing the relatively stringent Bonferroni correction because these tests were exploratory and not statistically justified. These pair-wise comparisons showed that boys in the Fantasy Planet condition scored significantly higher than did boys in the Standard condition, $p = .04$. This result matches what was found in the complete sample of Study 1 boys (i.e., ages 40-47 months old). There were no other significant pair-wise comparisons for boys. Also, there were no significant effects for girls in an identical ANOVA.

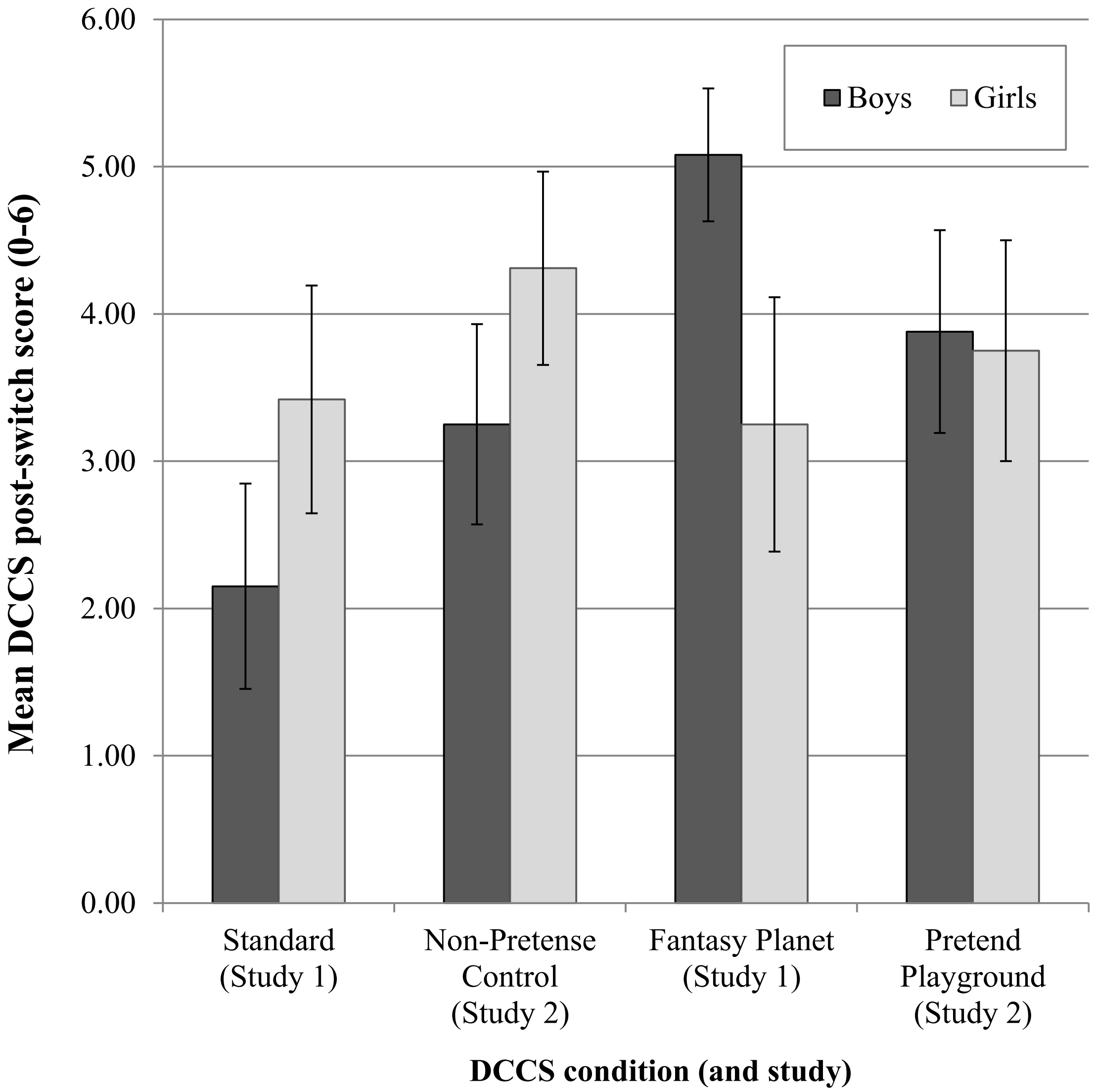



*Figure 2.*
Mean DCCS post-switch score by condition and sex (Studies 1 & 2; 43- to 46-month-olds only). Error bars represent standard errors.

Table 10.
*DCCS Post-switch Performance by Sex and Condition (Studies 1 & 2, 43- to 46-month-olds only)*

| | | DCCS post-switch (0-6) | |
|---|---|---|---|
| **Condition** | ***n*** | ***M (SD)*** | ***95% CI*** |
| | **Boys** | | |
| Standard (Study 1) | 13 | 2.15 (2.51) | [0.64, 3.67] |
| Non-Pretense Control (Study 2) | 16 | 3.25 (2.72) | [1.80, 4.70] |
| Fantasy Planet (Study 1) | 12 | 5.08 (1.56) | [4.09, 6.08] |
| Pretend Playground (Study 2) | 16 | 3.88 (2.75) | [2.41, 5.34] |
| | **Girls** | | |
| Standard (Study 1) | 12 | 3.42 (2.68) | [1.71, 5.12] |
| Non-Pretense Control (Study 2) | 16 | 4.31 (2.63) | [2.91, 5.71] |
| Fantasy Planet (Study 1) | 12 | 3.25 (2.99) | [1.35, 5.15] |
| Pretend Playground (Study 2) | 16 | 3.75 (3.00) | [2.15, 5.35] |

*Note.* CI = confidence interval.

Another set of cross-study comparisons arose because of the combination of the presence of a condition effect for boys in Study 1 and the absence of a condition effect for boys in Study 2. This combination of results raised the question of how the samples of boys in the two studies compared to each other. If the groups of boys were somehow different, it is possible that those differences influenced the presence versus absence of an effect of the experimental manipulations. Therefore, exploratory analyses were conducted to test for significant differences between the studies' samples of boys in terms of their age or CBQ subscale scores or their performance on the behavioral tasks included in both studies (PPVT, FDS, and Bear-Dragon). A series of ANOVAs indicated no significant differences, all *p*s > .05, although some trends approached significance. For the purposes of these analyses, only children within the Study 2 age range (43-46 months) were included; however, the boys in this subsample of Study 1 ($n$ = 25) had a somewhat higher average age ($M$ = 44.72, $SD$ = 1.10, 95% CI [44.27, 45.17] than the boys in Study 2 had ($M$ = 44.19, $SD$ = 1.12, 95% CI [43.78, 44.59]), $F$(1, 55) = 3.22, $p$ = .08, $\eta_p^2$ = .06. These boys in Study 1 also tended to have a higher Negative Affect score on the CBQ ($M$ = 4.31, $SD$ = 1.09, 95% CI [3.86, 4.76]) than did boys in Study 2 ($M$ = 3.88, $SD$ = 0.75, 95% CI [3.61, 4.15]), $F$ (1, 55) = 3.04, $p$ = .09, $\eta_p^2$ = .05. Overall, however, there were no significant differences between boys in the two studies' samples that could explain the presence of a condition effect (for boys) in Study 1 but not in Study 2.

**Discussion**

As expected, the random assignment process in Study 2 was successful in that the condition-based groups were equivalent in terms of age, receptive vocabulary (PPVT), secondary measures of EF (Bear-Dragon and FDS), and temperament (CBQ). Those initial results helped to verify the validity of comparing condition-based groups on DCCS performance to determine the effect of the experimental manipulation in Study 2. Also, in some analyses, preschoolers' DCCS performance (particularly regarding the highest level passed) was related to their PPVT scores, reflecting the established relation between EF and receptive vocabulary (e.g., Carlson et al., 2005; Carlson & Moses, 2001; Hongwanishkul et al., 2005). With the restricted age range in this study, it was only necessary to control for PPVT and not age. If pretense facilitates better EF performance, then the condition-based groups should have performed differently on the DCCS even after accounting for children's receptive vocabulary scores. However, in this study, children in the pretend-based experimental condition did not perform any differently on the DCCS than did children in the condition that did not involve any element of pretense.

The results from Study 2 must be considered within the context of the findings from Study 1 because Study 2 was specifically designed to address questions raised by the first experiment: (1) What was it about the experimental manipulation in Study 1 that facilitated better DCCS performance (for boys), and (2) Why was the effect apparent for boys but not for girls? I will address the second question (and my associated hypotheses) periodically within the context of the following discussion of the first question, which is of primary importance.

As for the first question, I hypothesized that if the incorporation of a pretense context facilitates better performance on the DCCS then there should be a significant difference between conditions in Study 2, with children in the Pretend Playground condition scoring higher. Because children in the pretend condition in Study 2 did not score any differently than children in that study's control condition scored, there is no support for this hypothesis or the associated explanation that pretense was responsible for the effect in Study 1. One interpretation of this null finding is that there is no facilitative effect of a pretense-based context on children's EF abilities (at least as measured by the DCCS). A more conservative interpretation is that there could be a facilitative effect of pretense but that it was not captured in this experiment. Perhaps the pretense manipulation was not strong enough, for example. Through the experimenter's language and intonation, she casually instructed children to pretend, but there were no physical activities, props, elaborate pretend scenarios, or assurances that children truly engaged in any pretense themselves. It is possible that stronger incorporation of pretense elements and children's involvement in the pretense would facilitate better performance on the DCCS.

In a study that was conducted at the same time as this research, White (2012) examined the effect of social distancing techniques on 3-year-olds' and 5-year-olds' performance on the EF Scale. Although the focus of the research was not on pretense per se, the most distal condition asked children to select a fictional character (from a given set) and pretend to be that character while they completed the card-sorting task. Children in that condition were given a sticker depicting the character to put on their hand as a reminder and a costume piece (e.g., a cape for Batman) to wear during the testing. Periodically, the experimenter reminded children to sort as the character (e.g., asking them, "Where should [character's name] put it?"). White found that 5-year-olds in this experimental condition passed higher levels of the EF Scale than did their same-age peers in a control condition that did not include any pretense or other psychological distancing. The author argued that 5-year-olds' performance on the EF Scale was facilitated by graduated degrees of social distancing, with the role-playing condition as the most distal version she tested. Although the author did not address such alternative interpretations, it is possible that children in the role-playing condition outperformed children in other conditions as a result of increased engagement in the activity due to the attractive sticker and costume pieces. An interpretation that merges the explanation of psychological distancing with that of engagement suggests that the stickers and costume pieces might have helped to heighten the pretense component. The pretense was therefore stronger in White's work than what was included in the present research, in which only simple verbal instructions were employed.

Thus, it is possible that the experimental manipulation in Study 2 did not demonstrate a facilitative effect of pretense on EF because its pretense component was too subtle. The simple inclusion of a pretense context is insufficient as a facilitative component for DCCS performance. While the basic idea and structure of the Pretend Playground condition were parallel to that of the Fantasy Planet condition from Study 1, the latter design might have been a more effective pretense manipulation due to some of the content differences. For example, the visual aid component or the particular theme of the pretense might have increased the pretense experience.

That reasoning relates back to the primary research question driving Study 2, which still remains unanswered: What was it about the Fantasy Planet condition that facilitated better performance (for boys) in Study 1? Regardless of a potential role for a pretense-based context, something else about the Fantasy Planet condition was at least partially responsible for the effect (for boys) in Study 1. This "something else" was present in the Fantasy Planet condition and absent in the modified Standard condition. In terms of Study 2, this mystery element was arguably either absent in both conditions or present in both conditions; either one of those arrangements would result in no significant difference between those two Study 2 conditions.

**What did the Fantasy Planet condition have that the Study 2 conditions both lacked?** By design, the Pretend Playground condition lacked the visual aid component and the particular pretense theme (and associated gender bias) that the Fantasy Planet condition had contained. Also by design, the Non-Pretense Control condition similarly lacked a visual aid and the pretense theme of outer space. First, I will address the visual aid component, which I mentioned as a possibly influential variable in the discussion of Study 1. To review: This visual aid might have had a facilitative effect by serving as an additional and visual marker of the switch in rules, providing a kind of pre-training exposure to the two rules, or illustrating the dimensions of color and shape separately. I hypothesized that if the visual aid were responsible for the effect in Study 1 then there should be no effect in Study 2, which was the case. However, it should also be noted that if the visual aid were the key component in Study 1, then children who are shown a visual aid in the context of a pretense manipulation should perform better on the DCCS than do children who are presented with a pretense manipulation without any visual aid. While neither study individually tested this hypothesis, it is instructive to compare the same-aged children from the relevant conditions across studies. In fact, children in the Fantasy Planet condition (Study 1) did not score higher on the DCCS than did children in the Pretend Playground condition (Study 2). This finding was the same whether data were collapsed across sex or considered separately for boys and for girls. This tentatively suggests that the visual aid alone cannot account for the collection of findings from both studies.

The second component that was present in the Fantasy Planet condition but absent in both Study 2 conditions was the theme of outer space, which also involved an unintentional masculine bias. As was mentioned in the discussion of Study 1, the outer space theme might have enhanced engagement, motivation, or mood for boys, which might have led to their better performance. Because neither condition in Study 2 incorporated this particular theme or a gender bias, the lack of a condition effect in this study is consistent with the possibility that the outer space theme (its content and/or its gender bias) was the facilitative component of the Fantasy Planet condition in Study 1. Again, however, the same cross-study comparison is relevant: Whether the analysis focuses on boys, girls, or both, children in the Fantasy Planet condition did not score higher on the DCCS than did same-aged children in the Pretend Playground condition. This tentatively suggests that the outer space theme (and its gender biases) also cannot fully account for the results from these two experiments.

**What did the Fantasy Planet condition and the Study 2 conditions all have in common?** The Fantasy Planet manipulation inherently emphasized change through the scripted language and the association of each game with a different entity (i.e., a different planet). This was pointed out earlier as a possible explanation for the (boys') better performance in this condition compared to (boys') performance in the modified Standard condition, which in some ways inadvertently emphasized sameness. By design, both conditions in Study 2 emphasized change in the same ways that the Fantasy Planet manipulation had done. Therefore, the lack of a difference between the Study 2 conditions is consistent with the argument that this emphasis on change was the component of the Fantasy Planet condition that facilitated (boys') better performance. Yet, again, a cross-study comparison reveals results that are inconsistent with this conclusion: If the emphasis on change (through language and the association between each game and a different entity) were the key component of the manipulation in Study 1 then children in the Non-Pretense Control condition should have scored higher than same-aged children in the modified Standard condition scored. However, there was no significant difference (for boys, girls, or both). Therefore, the argument that an emphasis on change facilitates better performance on the DCCS is not particularly compelling.

**Revisiting pretense.** Overall, the results discussed thus far show no compelling evidence that any of the components of the DCCS design that have been considered is responsible for the pattern of differences (and lack thereof) found in Studies 1 and 2. A pretense context, a visual aid, an outer space theme (and its associated gender biases), and a verbal emphasis on change were each systematically addressed. One by one, the arguments regarding their roles have been weakened by the results from condition comparisons both within and between studies. With no single variable clearly responsible for the boost in scores that boys in the Fantasy Planet condition showed in Study 1, it is worth considering more complex interactions among the variables.

For example, if any facilitative relation exists between pretense and EF then it is a complex relation applicable under some circumstances but not others. One complicated proposal is that a facilitative effect of pretense appears as long as the pretense theme is not associated with the opposite gender. A simpler conditional scenario might be that pretense is facilitative only when participants are boys. In Study 2, as indicated above, there were no differences between scores in the gender-neutral Pretend Playground condition and scores in the Non-Pretense Control condition (for boys or for girls). In addition, a more exploratory cross-study comparison showed that girls in the Pretend Playground condition performed similarly to girls in the Fantasy Planet condition. Given those two findings, there was no support for the argument that a pretense context facilitates better EF as long as the theme of the pretense is not biased towards the opposite gender. Because there was no condition effect for boys in Study 2, there was also no support for the hypothesis that there is a facilitative effect of pretense on DCCS performance for boys only.

To further address the possibility of a complex interplay between pretense and other variables, there are two other cross-study comparisons that could be considered – again, in an informal and exploratory fashion. These comparisons involve the pretense-based condition from one study and the control condition from the other study. Besides their cross-study nature and the lack of statistical justification, another problem with these comparisons is that they do not isolate the pretense context (presence vs. absence) from all other possibly influential variables including: the visual aid, the pretense theme and any associated gender biases, and the emphasis on change. This research did not test all possible combinations of each level of those variables. Despite this, it can be informative to consider combinations of components that might affect DCCS performance.

The first remaining cross-study comparison involves the pretense condition from Study 1 (Fantasy Planet) and the control condition from Study 2 (Non-Pretense Control). If the combination of a pretense context and certain conditions (i.e., with an outer space theme and/or its gender-related biases and/or with a visual aid) facilitates better EF performance then children in the Fantasy Planet condition should have scored higher than did same-aged children in the Non-Pretense Control condition. Overall, this was not the case.

The second remaining cross-study comparison involves the pretense condition from Study 2 (Pretend Playground) and the control condition from Study 1 (modified Standard). If the combination of a pretense context and other certain conditions (i.e., an emphasis on change versus sameness) facilitates better EF performance then children in the Pretend Playground condition should have scored higher than did same-aged children in the modified Standard condition. Again, overall, this was not the case.

The results from these two tentative cross-study comparisons do not provide any clarification regarding how Study 1 and Study 2 results can be reconciled. Further research would be necessary to better understand the complexities of an effect involving pretense and other variables, if one exists.

# Chapter IV: General Discussion

Considered collectively, results from Study 1 and Study 2 are largely inconclusive. While Study 1 data suggest that a pretense context for the DCCS might facilitate better performance for boys, there was no evidence of such a relation in Study 2. In addition, each primary alternative explanation for the findings in Study 1 was systematically evaluated and its argument weakened with data from Study 2 and cross-study comparisons. It seems unlikely that the visual aid, outer space theme with a masculine bias, or the verbal emphasis on change was solely responsible, as indicated in the earlier discussions. Also raised earlier was the possibility that there was something different about the boys in Study 1 that influenced the results. As presented in the Results and Discussion of that first study, although exploratory analyses suggested that boys in the two conditions might have differed in their Bear-Dragon performance, it is difficult to interpret that result given the statistical weakness of the difference and the unconventional nature of the task administration in Study 1. Overall, there was no clear indication that the random assignment process failed to create two comparable groups of boys in the first study.

Similarly, there were no significant differences between the sub-sample of 43- to 46-month-old boys in Study 1 and the sample of boys in Study 2 (with the same age range). Tests regarding age, CBQ subscale scores, and performance on the behavioral tasks (PPVT, FDS, and Bear-Dragon) showed only a couple of non-significant trends. In addition, participants were recruited in the same manner from the same database of families in the broader Seattle area. Although the studies took place in different years (and thus with different cohorts of boys), there is no particular reason to believe that the timing would influence the results of the DCCS manipulations, especially because the data collection for Study 1 itself occurred in three waves over the span of multiple years. So, there is also no reason to believe that the difference in results between the two studies was due to differences in the populations sampled.

Without further research it is unclear which aspect (or combination of aspects) of Study 1 was responsible for the significant difference in boys' performance in the Fantasy Planet condition compared to in the modified Standard condition. This is especially true given the null findings in Study 2 and in the tentative cross-study comparisons. If there had been no condition effect for anyone in Study 1 or Study 2, interpretation would be more straight-forward. Conclusions would also be clearer if the studies' results were more compatible in any other way (i.e., if the Sex × Condition interaction were replicated in Study 2, if it was followed by a main effect of Condition in Study 2 - once the pretense theme was more gender-neutral, or if boys in the Fantasy Planet condition outperformed boys in the Pretend Playground condition). It is the particular combination of results from the two studies that presents significant challenges for interpretation.

Prior to delving too far into further pursuit of a facilitative effect of pretense and an explanation for the collection of results found in these studies, it could be informative to attempt to replicate the effect found here in Study 1. Such replication would strengthen our conviction that further research based on this particular finding is worthwhile. This is even more necessary in light of Lillard and colleagues' (2012) recent critique of the literature that claims an effect of pretend play on EF. Besides those authors' arguments regarding published inconsistencies and null findings, it is worth noting that, given the field's publication bias against null findings, there might be a number of undistributed, unknown results that have failed to show evidence of a pretense-EF relation. Therefore, replication of an effect here could be important before investing significantly in further pursuit of the current line of research.

In an ideal world without practical concerns about limited resources, if further pursuit is warranted then a more precise and controlled set of experimental conditions could be implemented to directly address the questions that were only tentatively addressed through cross-study comparisons in the present research. For example, a study could include the addition of a Fantasy Planet condition without the visual aid, a control condition with a visual aid that parallels the original Fantasy Planet aid in Study 1, or a pretense condition with a more feminine theme that could parallel the masculine Fantasy Planet condition. More formal comparisons that truly isolate each variable would be necessary before the results from Study 1 and Study 2 can be understood fully.

At this point, however, there is one conclusion that can be drawn from this research with a reasonable level of certainty: A minor manipulation of the verbal DCCS instructions that includes references to pretense is insufficient to facilitate better performance on the task for this particular demographic. It is important to note four essential features of that stated conclusion: (1) It refers to a minor manipulation that includes pretense only through verbal instruction; (2) It indicates that such inclusion of pretense is not sufficient on its own; (3) It limits conclusions to the particular population sampled; and (4) It limits conclusions to the DCCS task. In this discussion, I will expand upon each of those features and discuss related limitations of this research, as well as suggestions for further study. I conclude with implications of this research.

**Minor Pretense Manipulation**

If the simple suggestion of pretending to travel between imaginary places during the DCCS were sufficiently helpful for children then children in the Pretend Playground condition should have performed better than their counterparts in the control condition. Similarly, the Fantasy Planet manipulation should have facilitated better performance for girls, as well as for boys.

Neither of those outcomes occurred. Therefore, if pretense were at all responsible for boys' better performance in Study 1 then it was most likely as one of a combination of elements that are only effective collectively. (This point will be elaborated upon in the next section.)

With pretense as the primary focus of this research, one important next question is whether a more substantive pretense manipulation would be more effective (and possibly sufficient). This issue was raised earlier, in the discussion of Study 2. In one of White's (2012) social distancing manipulations of the DCCS (EF Scale), participants were instructed to role-play while sorting cards. This manipulation incorporated props/costumes, stickers, and well-established fictional characters, and this manipulation was effective for one of the age groups studied. In contrast, the present studies explored pretense manipulations involving only verbal instructions to pretend (with and without a simple accompanying visual aid). Karniol and colleagues (2011) showed that children performed better on a DoG task when asked to pretend to be Superman even when there was no prop or costume (i.e., cape) involved. They argued that their work shows that children benefit from role-play (or, in their words, "cognitive self-transformation") strategies on the DoG task even when no props are provided; however, the authors did not directly address whether props make a role-play strategy comparatively more effective. Visual inspection of the means for duration of delay in their paper suggests that if there was a difference between having and not having a cape prop then children who were given a cape actually waited less time. Without proper analysis and controls, however, this trend is inconclusive. In addition, the role-play that was effective in their study involved a known fictional character, which might have inherently increased the strength of the pretense manipulation, independent of the cape. Finally, a finding regarding the DoG task might or might not even be applicable to the DCCS task. Overall, it seems quite possible that the absence of any prop, known character, or costume piece in the present research meant that the pretense manipulation was relatively weak.

Interestingly, anecdotal evidence suggests that at least some children enthusiastically agreed to the pretense premises (i.e., after the experimenter suggested that they pretend to be at an imaginary location these children said, “Yeah!” with some excitement); however, many children might not have actually engaged in pretense themselves, despite the experimenter’s suggestions. Neither the two current studies nor the other studies just cited (Karniol et al., 2011; White, 2012) incorporated any mechanism for measuring the degree of children’s participation in the pretense, although there is some mention of children’s Superman-like behaviors in Karniol and colleagues’ work. White and I each included questionnaires regarding children’s pretense-related proclivities more generally, but none of these researchers directly measured children’s pretense during the task. All of this considered, future research with the DCCS might involve both more intense pretense incorporation than what was included here and a formal measure of children’s pretense engagement during the task. For example, a pretense manipulation could be embellished with props, scenery, costumes, characters, and/or physical activity, and then video footage could be coded for children’s engagement using coding schemes based on previous observational work on children’s pretend play (see Elias & Berk, 2002; Haight, Parke, & Black, 1997; Smilansky & Shefatya, 1990; Youngblade & Dunn, 1995). Coding could include, for example, the degree to which children respond appropriately to their partner’s (the experimenter’s) pretense and behave in manners that are consistent with the established pretend stipulations (see Harris, 1994). It could be informative to have a single study that contrasts a more elaborate pretense manipulation with the type of minor manipulation that was used in the present research. It might be the case that pretense-based manipulations of the DCCS are more effective in facilitating better performance when they are more elaborate and/or when children themselves are more deeply involved in pretending.

This relates to other potentially influential variables that were mentioned in the earlier discussions: children's engagement, motivation, and mood. If pretense effectively enables better performance in certain circumstances, is it because children are engaged in the pretense or because children are engaged, period? The possibility remains that boys in the Study 1 Fantasy Planet condition had elevated engagement, motivation, or mood, which then led to better performance, although this argument is weakened by the fact that these boys did not score higher than did boys in the Pretend Playground or Non-Pretense Control conditions. It is assumed that the latter two conditions are likely less engaging (particularly for boys) than the Fantasy Planet condition; however, this was not measured in any manner. To add some perspective, it is worth noting that other studies investigating the use of pretense strategies for EF (or similar) tasks have neglected even to mention this issue as part of their discussion (Dias & Harris, 1988, 1990; Karniol et al., 2011; White, 2012).

That said, concerns about differences in engagement, motivation, or mood likely could be addressed through the incorporation of behavioral or psychophysiological measurements. Researchers in other areas have measured children's task engagement or motivation through observational coding systems involving, for example, on-task versus off-task behavior, attention, emotional expressions, persistence, secondary socializing, or a combination of such behaviors (e.g., Berhenke, Miller, Brown, Seifer, & Dickstein, 2011; de Kruif & McWilliam, 1999; Imhoff & Baker-Ward, 1999). In addition, there are indications that task engagement has been associated with a variety of psychophysiological measures, at least in adult samples. For example, in Fairclough and Venables's (2006) work, psychophysiological measures (including

EEG, heart rate, sinus arrhythmia, respiration rate, and eye-blink) predicted up to 53% of the variance in participants' self-reported task engagement. Also, fMRI research has implicated specific brain regions in adults' task engagement, suggesting that measurements of children's brain activity during tasks might also be informative. For example, measurement of the deactivation of a certain collection of brain regions--the default mode network (DMN)--has been related to task engagement (e.g., Greicius & Menon, 2004; Liddle et al., 2011). According to these researchers, the DMN is typically active during our default mode, such as rest or passive states that involve only intrinsic brain activity. In contrast, activity in these regions decreases during participation in demanding external tasks. If a task is insufficiently engaging, however, this deactivation will be attenuated (e.g., Greicius & Menon, 2004). Measurements of brain activity and its changes can thus help to quantify task engagement. This brief foray into the relevant literature suggests that researchers could consider adapting behavioral or psychophysiological approaches to measure children's DCCS task engagement (or motivation or mood). If such an approach were successful, these factors could be controlled for in future studies of the effects of experimental manipulations. This approach was beyond the scope of the current research project, but it could be worth exploration if the engagement/motivation/mood issue is a target of particular interest.

Another possible explanation for why the minor pretense manipulations in these studies were not consistently facilitative is if only certain types of pretense are helpful in an EF task. This could be related to children's engagement or it could be independent of that component, and it relates back to the idea of specificity I discussed in the Introduction in response to Lillard and colleagues' (2012) stance on the pretense-EF relation. Carlson and Beck (2009) acknowledged that some pretense strategies might be more or less effective on the DoG based on whether they decrease or increase the hotness (i.e., affective salience) of the task. The pretense adopted in the current studies involved instructions to behave as if one is at an imaginary place that is associated with shape or color (whichever dimension is presently relevant). This general approach was based on Dias and Harris's (1988, 1990) work with counterfactual syllogisms and imaginary planets. Thus, in the current research there was no explicit object substitution, attribution of unreal traits to oneself or others, or role-play, although an argument could be made that children were implicitly encouraged to role-play as astronauts in the Fantasy Planet condition. In contrast, White's (2012) study with the EF Scale involved explicit instructions to children that they pretend to be specific established fictional characters. Perhaps pretense that involves role-play is a more effective EF strategy than other types of pretense. This possibility is consistent with the theoretical argument mentioned in the Introduction that role-play, in particular, might facilitate the psychological distancing from reality that, in turn, facilitates more reflection-based actions (Carlson & Beck, 2009).

Narrowing in even further, Karniol and colleagues' (2011) work with the DoG task suggests that a facilitative effect of pretense might be limited to role-play that has certain characteristics. Their work showed that role-play in which children pretended to be a character known to have relevant strengths (i.e., patient Superman) was more beneficial than role-play involving a character known to have conflicting traits (e.g., impulsive Dash from the movie *The Incredibles*). In addition, pretending to be another child who had the same relevant strengths as Superman was not as beneficial as pretending to be Superman specifically. The authors argued that children cognitively transformed themselves during the pretense and adopted the relevant characteristics of Superman; however, the authors did not clearly explain why children could not similarly transform themselves into the unknown fictional child. The main point, however, is that the specific type of or focus of the pretense could be important. This idea echoes conclusions drawn by others (e.g., Berk et al., 2006; Haight et al., 2006; Lillard et al., 2012) that the thematic content of children's pretend scenes affects how beneficial the pretense is for their emotional health.

In sum, there are reasonable arguments that the manipulations of the DCCS instructions that were embraced in the current research program were too minor in their depth of pretense and/or incorporated the "wrong" kind of pretense to facilitate preschoolers' better performance on the task.

**Minor Pretense Manipulation is Insufficient**

This second feature of the main conclusion from these studies emphasizes that a minor pretense manipulation (such as the ones used in this research) is not universally facilitative of better DCCS performance. If that were the case then there would have been main effects of condition in both studies here. Instead, there must be other important circumstances or design elements that are a necessary part of any facilitative effect involving pretense that might exist. It is worth noting that tentative conclusions based on the within-study and cross-study comparisons indicated that no individual element addressed in this research (i.e., visual aid, verbal emphasis on change) is singularly effective as a facilitative manipulation. So, it is not just the minor use of pretense that is insufficient on its own. The main focus of this work, however, remains on the pretense component.

The issues already presented above suggest that a pretense manipulation that involves a minor verbal instruction is insufficient, but there's also the possibility that pretense itself, regardless of the intensity, is insufficient. There is, of course, the possibility that pretense does not facilitate EF performance at all – which is consistent with the skepticism expressed by

Lillard and colleagues (2012). The pretense-EF relation summarized in the Introduction might not be causal in that direction or in any direction. However, it is also possible that the combination of pretense with one or more other variables is effective as an EF strategy. For example, maybe pretense and a visual aid and an emphasis on change (i.e., Fantasy Planet condition) can facilitate better performance, but the single features or sub-combinations are not significantly effective on their own. It might even be the case that certain types of pretense manipulations, when combined with other specific circumstances, facilitate EF for boys but not for girls. Or, as indicated in the Discussion of Study 1, children's sex and their other characteristics might have been intertwined in this research such that the benefit of a pretense manipulation is not limited to boys but to children with lower EF, for example. If that were the case, this would be consistent with other instances in which children with the lowest EF abilities benefited the most from an intervention (Diamond & Lee, 2011). This would also be consistent with Elias and Berk's (2002) finding that sociodramatic play predicted later self-regulation in children who were highest in impulsivity but not in children who were lowest. The pretense-EF relation might be strongest for children with the weakest self-control abilities.

As was mentioned earlier, this research did not test all possible combinations of each level of each variable. Ideally, future research would test manipulations that involve various combinations of pretense, visual aid, verbal emphasis on change, and any other identified factors, and determine which combinations facilitate better performance for which children. Identification of the role of factors that have not yet been identified might solve the current mystery regarding the presence of a condition effect for boys in Study 1 and the absence of such an effect in Study 2.

**Limited to the Population Sampled**

As with all research, it is important to consider the population that was represented in these studies' samples: typically developing children (ages 40-47 or 43-46 months) from middle- to upper-middle class, primarily White, English-speaking families living in the broader Seattle area. Before drawing conclusions about whether findings generalize to children who are different from this population in any way, additional research with appropriate sampling would be necessary.

In the meantime, it is relevant to consider how the particular population that was sampled here might have affected the results. For example, perhaps the pretense manipulations here are not sufficiently helpful for children in this particular age range. As mentioned above, White's (2012) role-playing manipulation was effective for only one of the age ranges that she studied. Although she found a social distancing effect for 5-year-olds, there was no such effect for 3-year-olds. Similarly, Kirkham, Cruess, and Diamond (2003) suggested that having the experimenter label only the relevant dimension of each card on the DCCS was helpful for 4-year-olds in their study but not for the 3-year-olds. Also, much of Dias and Harris's (1988, 1990) work on the effect of fantasy on children's reasoning about counterfactual syllogisms was based on samples of 5- and 6-year-olds with some studies including 4-year-olds but not 3-year-olds. Although Carlson and Beck (2009) included 3-year-olds in their research on children's strategies on a DoG task, these children were combined with young 4-year-olds in the "younger" group (range = 39 to 51 months), so the results for 3-year-olds alone are not presented. That said, this younger group spontaneously used pretense strategies less often than did their older counterparts (ages 52-60 months), suggesting that the use of pretense strategies might be different for 3-year-olds versus children who are almost 5. Children of various ages were similarly grouped together in Karniol and colleagues' (2011) three studies of pretense instructions for a DoG task. In their research, children in as broad a range as 38 to 63 months old were analyzed collectively. Although results from the full samples suggested that cognitive self-transformation into Superman facilitated better delay, it is unclear whether the same pattern would hold for the subset of 3-year-old children.

All of this considered, it is possible that the effectiveness of the more subtle manipulations tested in the present research was impacted by the youth of the preschooler participants. Maybe pretense is not as facilitative of EF in the early preschool years as it is for slightly older children. The age range for the current research program was selected specifically with the DCCS in mind. Based on prior research with this task, the transition from failing to passing tends to occur between 3 and 4 years old (e.g., Carlson, 2005; Zelazo, 2006). However, perhaps children are too young at that age to benefit from a pretense manipulation. In that case, even minor pretense-based manipulations might be effective with slightly older children, assuming a more formalized adoption of the various levels of the EF Scale (to avoid a ceiling effect on the main DCCS, Level 4). This issue alone does not explain how an effect was found for similarly aged boys in Study 1 but not those in Study 2 (although there was a trend that approached significance for the boys in Study 1 to be older than the boys in Study 2, even within the limited 43-46 month age range); however, it is still an important consideration when reflecting on the current set of studies and their results.

**Specific to the DCCS Task**

Any conclusions drawn from the current research must also be limited to the DCCS. Due to the particular priorities in this research program, the EF Scale was not administered exactly as designed. Therefore, even conclusions regarding the EF Scale must be tentative, in addition to conclusions regarding EF tasks that do not involve this type of card sort. The broader relation between pretense and EF might be explored further through the manipulation of other EF tasks. If a pretense manipulation similar to the one in this research facilitates better performance on another EF task (e.g., Grass/Snow; Carlson & Moses, 2001) even though it was insufficient for the DCCS, then those results could inform our conceptualization of each task and the pretense-EF relation.

In addition, it could be useful to compare pretense-based manipulations of EF tasks that are categorized as “hot” versus “cool” (Hongwanishkul et al., 2005) or “conflict” versus “delay” (Carlson & Moses, 2001). The research on pretense and other symbolic distancing strategies has usually focused on hot EF, including the Less is More task (Carlson et al., 2005), or on delay tasks (Carlson & Beck, 2009; Karniol et al., 2011), with the recent exception of White (2012). Although the theoretical connections between pretense and ToM are controversial (see Harris, Lillard, & Perner, 1994; Leslie, 1987), an argument could be made that the same types of EF tasks that relate to ToM could relate more strongly to pretense, as well. For example, research has shown that conflict-based tasks are more strongly related to measures of ToM than are delay-based tasks (e.g., Carlson & Moses, 2001). Therefore, perhaps conflict-based tasks are also more strongly influenced by pretense manipulations. It is noteworthy that some conflict-based EF tasks include an element of pretense in their basic design, despite no stated purpose to examine its effect: Bear-Dragon uses puppet characters with nice and naughty traits, respectively (Reed et al., 1984), and Less is More involves a naughty monkey puppet (Carlson et al., 2005). Meanwhile, I am not aware of any delay-based tasks that incorporate pretense into their basic design. Perhaps that pattern is not entirely happenstance.

**Other Limitations and Suggestions for Future Research**

Thus far, suggestions for future research have included: replication of the interaction effect in Study 1, additional and direct consideration of other potentially influential variables from Study 1 (e.g., Fantasy Planet without the visual aid), investigation of pretense manipulations with varying levels of intensity and different types of pretense, measurement of children's engagement in the pretense and the task more broadly, methods of controlling for children's general task engagement, combinations of pretense with other potentially facilitative features (e.g., visual aid), use of a broader age range (with age groups similar to and older than the present research), and application of pretense strategies within different EF tasks. These suggestions are based on an acknowledgment of some of the primary limitations of the present research.

As is the case with most experiments, there are other methodological concerns that critics might raise about the present research. One issue is the potential that experimenter bias influenced the results. This is often a genuine concern; however, I will argue that it was not problematic in this research. In these two studies, the experimenter for all sessions was the primary investigator and author, who had hypotheses regarding this work. While aiming to study the true nature of the pretense-EF relation, I also admittedly hoped that there would be facilitative effects of my manipulations, as that would be the most promising contribution to the field. These expectations or hopes could have influenced children's performance during the testing, since it was impossible to keep the experimenter blind to children's condition assignments. In fact, not having fully blind experimenters is an under-recognized criticism of most, if not all, studies involving a manipulation of the DCCS (e.g., Diamond et al., 2005; Honomichl & Chen, 2010; Kirkham et al., 2003; Kloo & Perner, 2005; Kloo et al., 2010; Mack, 2007; Perner & Lang, 2002), since the manipulations tend to affect the process that is overseen by the experimenter. The issue of blind (or "masked") experimenters was raised by Lillard and colleagues (2012) as a substantial criticism of the literature on potential benefits of pretend play. Here and in other cases in the literature, it is possible that children responded to unintentional signals from the experimenter. That said, results in the current studies were not actually reflective of the experimenter's hypotheses or preferences. If data were impacted by experimenter biases then they would most likely have shown stronger facilitative effects of the pretense-based manipulations or, at the least, a more conclusive set of results from the two studies. Therefore, while an ideal design would include an experimenter who is blind to at least hypotheses or condition, if not both, it is highly unlikely that the results obtained in this set of studies were influenced by experimenter biases.

Another methodological criticism of this research could be that there were missing data for some children on individual tasks. Some data were unusable due to a procedural error or another issue unrelated to the child, and it is certainly best to minimize those cases. Most missing data, however, resulted from children's refusal to complete the task, which might have related to those children's cognitive skills or other traits, such as attention spans, temperament, and, importantly, self-control abilities[8]. Given the secondary status of the PPVT, FDS, BDS, and Bear-Dragon, missing data for these items were unlikely to have a large effect on the primary results regarding the DCCS; however, they might have informed the investigations of covariates and "baseline"

[8] Exploratory analyses were conducted to compare children with data on specific tasks to children without data on those tasks in terms of age, sex, condition, CBQ subscores, and DCCS post-switch scores. A series of MANOVAs revealed the following significant differences: (1) In Study 2, children who were missing PPVT scores had higher Surgency (CBQ) scores than did children who had valid PPVT scores; (2) In Study 2, children without FDS scores ($n = 2$) had lower DCCS post-switch scores than did children who had valid FDS scores, since both of the children missing FDS scored 0 on the DCCS. There was also a non-significant trend in Study 1 for children without Bear-Dragon scores to have lower DCCS post-switch scores than did children who had valid Bear-Dragon scores.

differences between groups. More significant would be missing data on the DCCS. While the final sample did not include anyone who failed to complete the main (Level 4) DCCS (or who failed the pre-switch phase), this was by design, not by natural outcome. There was a small handful of children in each sample who were excluded due to child refusal to complete the DCCS or to low DCCS pre-switch scores. These children might have differed from children who had usable DCCS data and were kept in the final sample[9]. This is another way in which the generalizability of the results might be limited. Children enter this research as voluntary participants whose rights must be respected, but that means that children who prefer not to complete various tasks are not represented in the data.

**Implications Regarding the DCCS**

Despite the inclusion of detailed Method sections in research publications and a published protocol for the DCCS (Zelazo, 2006), there are many variations in DCCS task administration (e.g., regarding sorting container, the practice phase, shapes/colors used, experimenter's phraseology) as reported in the literature. Some of the variations are purposeful manipulations designed to address a theoretical issue regarding preschoolers' traditionally poor task performance. For example, Diamond, Carlson, and Beck (2005) manipulated their test cards by contrasting traditional integrated stimuli with stimuli that separated the shapes from the colors to test their hypothesized role of inhibition. Other variations, however, are not a given study's focus, are not driven by theory, and/or are unintentional side effects of other efforts. Which details of DCCS administration are important to maintain, and which adjustments can be made

---

[9] Exploratory analyses were also conducted to compare the CBQ subscores of children who were excluded from each study due to their DCCS performance (i.e., refusal to complete the DCCS or low DCCS pre-switch scores) to the subscores of children who were part of the final samples. (Children who were excluded from the final samples for other reasons were not part of these analyses.) For each study, a series of MANOVAs showed no significant differences between these two sets of children.

without having a significant impact on research findings? These questions are important both for methodological authenticity and for our understanding of what construct the DCCS actually measures.

While the present research was not designed to address these methodological questions, the results are informative in this regard. In the course of Study 1 and Study 2, a number of adaptations were made to the standard DCCS protocol put forth by Zelazo (2006). Obviously, the addition of the pretense component for two conditions involved changes in the script and associated presentation of the task. In order to mirror those experimental conditions, the control conditions also involved adaptations to the experimenter's script. There were additional aspects of the DCCS presentation that were not analyzed in the present studies but nonetheless differed from other researchers' presentations of the task. In Study 1, there was the addition of a card stand, which the experimenter used to display each test card before handing it to the child to sort. In addition, the role of the context of the DCCS within the series of tasks comprising the testing session has not been evaluated and yet certainly differed between these two studies, as well as among the vast number of studies in the broader literature.

As indicated multiple times above, there was not enough evidence in the present research to suggest that any single analyzed change in the DCCS administration affected preschoolers' performance, since most comparisons of conditions with versus without particular elements showed no significant differences. Aside from the boys in Study 1 performing differently across conditions, children tended to perform similarly throughout the current studies, and even those boys' performances did not differ significantly from the scores seen in Study 2. It would be interesting to compare DCCS post-switch performance in the various conditions in the current studies to same-aged children's performance in other studies that have involved the more traditional DCCS. At the basic level, the pattern that often appears in the literature (a bimodal distribution of scores with a large percentage of 3-year-old children failing the task) was replicated in each condition of the present research. However, a more detailed comparison is made difficult by the variation in the number of test trials, ages studied, and particular data reported in the DCCS literature. More formal comparison, such as a meta-analysis with access to the necessary data, would strengthen the currently tentative argument that the present studies' changes in the experimenter's basic script and phraseology were not disruptive to capturing children's general struggles on the DCCS. If the multiple variations in the DCCS administration do not influence research findings, then this adds to the robustness of the DCCS as an EF measure, and researchers can rest assured that the maintenance of minor details of its administration is not important.

Whether or not a particular change to the DCCS has an effect on children's performance also informs our understanding of what makes the traditional DCCS difficult for preschoolers. Diamond, Carlson, and Beck's (2005) manipulation referenced above, for example, showed that children do not struggle as much with the DCCS when the dimensions are visually separated on the test cards. The authors argued that this manipulation identifies the key challenge in the traditional DCCS by showing children's success upon its removal. According to these authors, the key challenge in the DCCS is for children to inhibit their tendency to think of the pictured image in one way (e.g., as a truck) in order to think of it in another way (e.g., as a blue thing). Since other theorized impediments still remained a part of the DCCS in this manipulation, the authors argued that they cannot be the elements of the task that are truly responsible for children's frequent failure.

In contrast, the results from the present research are largely compatible with the wide variety of theories regarding what makes the DCCS so challenging. While the results here do not clearly support any particular theory, the argument could be made in each case that the manipulation simply did not minimize the theorized impediment to the extent necessary to facilitate significantly better performance. For example, the attentional inertia explanation (e.g., Kirkham et al., 2003) suggests that children have difficulty overcoming their tendency to pay attention to the dimension that was relevant in the pre-switch phase in order to sort by the other, newly relevant dimension in the post-switch phase. The pretense manipulation, as well as the visual aid and the verbal emphasis on change, had the potential to highlight the necessity to focus on a new dimension and draw children's attention towards it. However, the results here suggest that whatever highlighting these changes might have provided was not sufficient to assist children with shifting their attention, and that requirement remained a challenge.

The same reasoning applies to the consideration of the idea that the DCCS is challenging because it requires children to inhibit prepotent response tendencies (or action schemas) (e.g., Carlson & Moses, 2001). This theory indicates that children struggle with inhibition not of attention but of the behavioral response of sorting based on the first rule. So, it could be argued that the pretense manipulations in the current research, as well as the visual aid and verbal emphasis on change, failed to make that behavioral inhibition significantly easier for participants. Therefore, there were no consistent significant differences in children's performance with these particular manipulations.

From the perspective of the re-description hypothesis (e.g., Kloo & Perner, 2005), children struggle with the concept of thinking of the same entity in two different ways. Given the argument that pretending often involves the management of two conflicting representations regarding the same thing (e.g., object substitution involves a real and a pretend identity for the same object), contextualizing the task with pretense especially has the potential to facilitate children's ability to think of the pictured items in two different ways. The fact that referring to pretense during the task instructions was not sufficient to facilitate better performance (since it did not help both sexes in both studies) indicates that this minor reference was not enough. As indicated above, maybe a more intense immersion into a pretense mindset would sufficiently prime children's abilities to consider multiple perspectives of the same entity.

According to the negative priming argument (e.g., Perner & Lang, 2002), children struggle with the DCCS because they are challenged in the post-switch phase to reactivate precisely the same response that was necessarily inhibited in the pre-switch phase. It could be argued that the pretense manipulation, visual aid, and/or verbal emphasis on change might assist children with this reactivation by emphasizing and drawing attention to the change. As with the other arguments, however, it could be the case that these manipulations were not intense enough to sufficiently make easier the reactivation of the previously inhibited dimension.

As the final example, Zelazo and colleagues' CCC and CCC-r theories (Frye et al., 1995; Zelazo et al., 1996; Zelazo et al., 2003) indicate that children struggle with the DCCS because they cannot manage the hierarchical rule structure, which organizes the two dimension-based rules under a higher-order rule about which game is to be played. The pretense manipulations, visual aid, and verbal emphasis on change did not simplify this rule structure, so the CCC theories might (correctly) predict that there would be no change in children's performance. However, Zelazo and colleagues (2003) also indicated that children can manage two conflicting rule sets (without having a higher-order rule in which they are embedded) when the rule sets are applied in different contexts. It could be argued that the pretense manipulations in the present research provided explicitly different (imaginary) contexts for each rule set, with Planet Shape versus Planet Color and the Shape Playground versus the Color Playground. If those suggested contexts sufficiently produced different contexts for the two rule sets, then the CCC-r theory would predict that children would be better able to switch rules. Again, the lack of significantly improved performance in those pretense conditions is still compatible with the CCC-r theory if it is argued that the manipulation was unsuccessful in truly creating different contexts for children.

Thus the same basic argument is applicable to many of the popular theories regarding the challenges of the DCCS: In general, while one might think that the pretense manipulation (and/or the visual aid and/or the verbal emphasis on change) would help children to abide by the change in the rules, the results indicate that the particular changes studied in the present research were not helpful enough to enable significantly better performance. With no explicit attempt to differentiate among theories regarding the DCCS (e.g., through the definitive isolation of theorized impediments), the present research does not provide substantive support for or against any of the popular theories; rather, these data are arguably compatible with each of them.

**Conclusion**

With the broader goal of examining possible developmental benefits of preschoolers' pretending, this research focused specifically on the pretense-EF relation. This particular association has been implicated in prior theoretical and empirical literature, as summarized in the Introduction. In addition, there is a contemporary recognition of the importance of understanding EF and facilitating EF development. This recognition is apparent in academic communities, as exemplified by the vast associated literature (including the recent "Special Section on Self-Regulation, Effortful Control, and Executive Functions in Child Development" in *Child Development Perspectives*, Volume 6, Issue 2, June 2012), and in applied settings, as implied by the creation and adoption of a variety of associated intervention programs (see Diamond & Lee, 2011).

The pretense-EF relation might involve causality, and insight into any facilitative effects of pretense on EF can provide important clarifications for both basic research and applied purposes. In that vein, the two studies in this research program centered on the addition of a pretense context to a laboratory-based task that measures conflict-related EF abilities: the DCCS. The preliminary conclusion from this research is that the type of minor pretense manipulation adopted here is not sufficient to facilitate better DCCS performance in the population studied; however, I have presented multiple recommendations regarding further investigation into how pretense might be better used as a DCCS (or EF) strategy. Once formal laboratory-based experiments have clarified effective ways of using pretense in controlled EF settings, it will also be important to determine how findings translate to real-world applications. How can parents, teachers, and other community members who interact with children in EF-related situations help those children to take advantage of whatever assistance pretense can offer them? Building upon the extant body of work on pretense and EF, the present research program has contributed to our understanding of what we know and what we should still be pursuing in future research. With the addition of each clarification, we will better understand what we are witnessing and its implications when we observe pretend play within a classroom of preschoolers. Pretending to be someone else, that a doll is a real baby, or that a block of wood is a car might be good strategy besides being good fun.

**References**

Baddeley, A. (1981). The concept of working memory: A view of its current state and probable future development. *Cognition, 10*(1-3), 17-23. doi: 10.1016/0010-0277(81)90020-2

Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly, 23*(3), 299-313. doi: 10.1016/j.ecresq.2008.03.001

Beck, D. M., Carlson, S. M., & Rothbart, M. K. (2007). *Executive function, effortful control, and parent report of children's temperament*. Unpublished manuscript. University of Washington, Seattle.

Bergin, D. A. (1999). Influences on classroom interest. *Educational Psychologist, 34*(2), 87-98. doi: 10.1207/s15326985ep3402_2

Berhenke, A., Miller, A. L., Brown, E., Seifer, R., & Dickstein, S. (2011). Observed emotional and behavioral indicators of motivation predict school readiness in Head Start graduates. *Early Childhood Research Quarterly, 26*(4), 430-441. doi: 10.1016/j.ecresq.2011.04.001

Berk, L. E., Mann, T. D., & Ogan, A. T. (2006). Make-believe play: Wellspring for development of self-regulation. In D. Singer, R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Play=Learning: How play motivates and enhances children's cognitive and social-emotional growth* (pp. 209-330). New York: Oxford University Press.

Black, B. (1989). Interactive pretense: Social and symbolic skills in preschool play groups. *Merrill-Palmer Quarterly: Journal of Developmental Psychology, 35*(4), 379-397.

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

Boyatzis, C. J., & Eades, J. (1999). Gender differences in preschoolers' and kindergartners' artistic production and preference. *Sex Roles, 41*(7-8), 627-638. Retrieved from http://link.springer.com/journal/11199

Bretherton, I. (1989). Pretense: The form and function of make-believe play. *Developmental Review, 9*(4), 383-401. doi: 10.1016/0273-2297(89)90036-1

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*(2), 595-616. doi: 10.1207/s15326942dn2802_3

Carlson, S. M. (2012). *Executive Function Scale for Preschoolers*. Test Manual. Institute of Child Development, University of Minnesota.

Carlson, S. M., & Beck, D. M. (2009). Symbols as tools in the development of executive function. In A. Winsler, C. Fernyhough & I. Montero (Eds.), *Private speech, executive functioning, and the development of verbal self-regulation.* (pp. 163-175). New York, NY US: Cambridge University Press.

Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more: Executive function and symbolic representation in preschool children. *Psychological Science, 16*(8), 609-616. doi: 10.1111/j.1467-9280.2005.01583.x

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

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

Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development, 22*(4), 489-510. doi: 10.1016/j.cogdev.2007.08.002

Carlson, S. M., White, R. E., & Davis-Unger, A. (2012). *Evidence for a relation between executive function and pretense representation in preschool children*. Manuscript under review.

Colwell, M., & Lindsey, E. (2005). Preschool children's pretend and physical play and sex of play partner: Connections to peer competence. *Sex Roles, 52*(7/8), 497-509. doi: 10.1007/s11199-005-3716-8

Connolly, J. A., & Doyle, A. B. (1984). Relation of social fantasy play to social competence in preschoolers. *Developmental Psychology, 20*(5), 797-806. doi: 10.1037//0012-1649.20.5.797

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*(1), 25-31. doi: 10.1080/00049539508258765

de Kruif, R. E. L., & McWilliam, R. A. (1999). Multivariate relationships among developmental age, global engagement, and observed child engagement. *Early Childhood Research Quarterly, 14*(4), 515-536. doi: 10.1016/s0885-2006(99)00028-9

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*(5855), 1387-1388. doi: 10.1126/science.1151148

Diamond, A., Carlson, S. M., & Beck, D. M. (2005). Preschool children's performance in task switching on the Dimensional Change Card Sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology, 28*(2), 689-729. doi: 10.1207/s15326942dn2802_7

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

Dias, M., & Harris, P. L. (1988). The effect of make-believe play on deductive reasoning. *British Journal of Developmental Psychology, 6*(3), 207-221.

Dias, M., & Harris, P. L. (1990). The influence of the imagination on reasoning by young children. *British Journal of Developmental Psychology, 8*(4), 305-318.

Dick, A. S., Overton, W. F., & Kovacs, S. L. (2005). The development of symbolic coordination: Representation of imagined objects executive function, and theory of mind. *Journal of Cognition and Development, 6*(1), 133-161. doi: 10.1207/s15327647jcd0601_8

Dunn, L. M., & Dunn, L.M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Elias, C. L., & Berk, L. E. (2002). Self-regulation in young children: Is there a role for sociodramatic play? *Early Childhood Research Quarterly, 17*(2), 216-238. doi: 10.1016/S0885-2006(02)00146-1

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin, 132*(1), 33-72. doi: 10.1037/0033-2909.132.1.33

Espy, K. A., & Bull, R. (2005). Inhibitory processes in young children and individual variation in short-term memory. *Developmental Neuropsychology, 28*(2), 669-688. doi: 10.1207/s15326942dn2802_6

Fairclough, S. H., & Venables, L. (2006). Prediction of subjective states from psychophysiology: A multivariate approach. *Biological Psychology, 71*(1), 100-110. doi: 10.1016/j.biopsycho.2005.03.007

Fein, G. G. (1981). Pretend play in childhood: An integrative review. *Child Development, 52*(4), 1095-1118. doi: 10.1111/j.1467-8624.1981.tb03157.x

Fein, G. G. (1987). Pretend play: Creativity and consciousness. In D. Görlitz & J. F. Wohlwill (Eds.), *Curiosity, imagination, and play: On the development of spontaneous cognitive motivational processes.* (pp. 281-304). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Field, T., Destefano, L., & Koewler, J. H. (1982). Fantasy play of toddlers and preschoolers. *Developmental Psychology, 18*(4), 503-508. doi: 10.1037/0012-1649.18.4.503

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483-527. doi: 10.1016/0885-2014(95)90024-1

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31-60. doi: 10.1037/0033-2909.134.1.31

Gleason, T. R. (2005). Mothers' and fathers' attitudes regarding pretend play in the context of imaginary companions and of child gender. *Merrill-Palmer Quarterly-Journal of Developmental Psychology, 51*(4), 412-436. doi: 10.1353/mpq.2005.0022

Goldstein, T. R., & Winner, E. (2010a). Engagement in roleplay, pretense, and acting classes predict advanced theory of mind skill in middle childhood. *Imagination, Cognition and Personality, 30*(3), 249-258. doi: 10.2190/IC.30.3.c

Goldstein, T. R., & Winner, E. (2010b). A new lens on the development of social cognition: The study of acting. In C. Milbrath & C. Lightfoot (Eds.), *Art and human development.* (pp. 221-247). New York, NY US: Psychology Press.

Greicius, M. D., & Menon, V. (2004). Default-mode activity during a passive sensory task: Uncoupled from deactivation but impacting activation. *Journal of Cognitive Neuroscience, 16*(9), 1484-1492. doi: 10.1162/0898929042568532

Haight, W. L., Black, J., Jacobsen, T., & , & Sheridan, K. (2006). Pretend play and emotion learning in traumatized mothers and children. In D. Singer, R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Play=Learning: How play motivates and enhances children's cognitive and social-emotional growth* (pp. 209-330). New York: Oxford University Press.

Haight, W. L., Parke, R. D., & Black, J. E. (1997). Mothers' and fathers' beliefs about and spontaneous participation in their toddlers' pretend play. *Merrill-Palmer Quarterly, 43*(2), 271-290.

Harris, P. L. (1994). Understanding pretence. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 235-259). Hillsdale, NJ: Erlbaum.

Harris, P. L. (2000). *The work of the imagination*. Oxford: Blackwell.

Harris, P. L., Lillard, A., & Perner, J. (1994). Commentary: Triangulating pretence and belief. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 287-293). Hillsdale, NJ: Erlbaum.

Hoffmann, J., & Russ, S. (2012). Pretend play, creativity, and emotion regulation in children. *Psychology of Aesthetics, Creativity, and the Arts, 6*(2), 175-184. doi: 10.1037/a0026299

Hongwanishkul, D., Happaney, K. R., Lee, W. S., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology, 28*(2), 617-644.

Honomichl, R. D., & Chen, Z. (2010). Relations as rules: The role of attention in the Dimensional Change Card Sort task. *Developmental Psychology*. doi: 10.1037/a0021025

Howe, N., Petrakos, H., Rinaldi, C. M., & LeFebvre, R. (2005). "This is a bad dog, you know...": Constructing shared meanings during sibling pretend play. *Child Development, 76*(4), 783-794. doi: 10.1111/j.1467-8624.2005.00877.x

Howes, C., & Matheson, C. C. (1992). Sequences in the development of competent play with peers: Social and social pretend play. *Developmental Psychology, 28*(5), 961-974. doi: 10.1037//0012-1649.28.5.961

Imhoff, M. C., & Baker-Ward, L. (1999). Preschoolers' suggestibility: Effects of developmentally appropriate language and interviewer supportiveness. *Journal of Applied Developmental Psychology, 20*(3), 407-429. doi: 10.1016/s0193-3973(99)00022-2

Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology, 30*(2), 190-200. doi: 10.1016/j.newideapsych.2011.11.001

Jacques, S., Zelazo, P. D., Kirkham, N. Z., & Semcesen, T. K. (1999). Rule selection versus rule execution in preschoolers: An error-detection approach. *Developmental Psychology, 35*(3), 770-780. doi: 10.1037/0012-1649.35.3.770

Karniol, R., Galili, L., Shtilerman, D., Naim, R., Stern, K., Manjoch, H., & Silverman, R. (2011). Why Superman can wait: Cognitive self-transformation in the delay of gratification paradigm. *Journal of Clinical Child and Adolescent Psychology, 40*(2), 307-317. doi: 10.1080/15374416.2011.546040

Kelly, R., & Hammond, S. (2011). The relationship between symbolic play and executive function in young children. *Australasian Journal of Early Childhood, 36*(2), 21-27.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*(5), 449-476. doi: 10.1111/1467-7687.00300

Kloo, D., & Perner, J. (2005). Disentangling dimensions in the Dimensional Change Card-Sorting task. *Developmental Science, 8*(1), 44-56. doi: 10.1111/j.1467-7687.2005.00392.x

Kloo, D., Perner, J., Aichhorn, M., & Schmidhuber, N. (2010). Perspective taking and cognitive flexibility in the Dimensional Change Card Sorting (DCCS) task. *Cognitive Development, 25*(3), 208-217. doi: 10.1016/j.cogdev.2010.06.001

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & et al. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*(2), 490-507. doi: 10.1111/j.1467-8624.1996.tb01747.x

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review, 94*(4), 412-426. doi: 10.1037/0033-295X.94.4.412

Liddle, E. B., Hollis, C., Batty, M. J., Groom, M. J., Totman, J. J., Liotti, M., . . . Liddle, P. F. (2011). Task-related default mode network modulation and inhibitory control in ADHD:

effects of motivation and methylphenidate. *Journal of Child Psychology and Psychiatry, 52*(7), 761-771. doi: 10.1111/j.1469-7610.2010.02333.x

Liew, J. (2012). Effortful control, executive functions, and education: Bringing self-regulatory and social-emotional competencies to the table. *Child Development Perspectives, 6*(2), 105-111. doi: 10.1111/j.1750-8606.2011.00196.x

Lillard, A. (1994). Making sense of pretence. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 211-234). Hillsdale, NJ: Erlbaum.

Lillard, A. (2001). Pretend play as Twin Earth: A social-cognitive analysis. *Developmental Review, 21*(4), 495-531. doi: 10.1006/drev.2001.0532

Lillard, A., Nishida, T., Massaro, D., Vaish, A., Ma, L., & McRoberts, G. (2007). Signs of pretense across age and scenario. *Infancy, 11*(1), 1-30. doi: 10.1207/s15327078in1101_1

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2012). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin*. doi: 10.1037/a0029321

Maccoby, E. E. (1998). *The two sexes: Growing up apart, coming together*. Cambridge, MA: Belknap Press of Harvard University Press.

Mack, W. (2007). Improving postswitch performance in the Dimensional Change Card-Sorting task: The importance of the switch and of pretraining by redescribing the test cards. *Journal of Experimental Child Psychology, 98*(4), 243-251. doi: 10.1016/j.jecp.2007.05.004

Marcovitch, S., Boseovski, J. J., & Knapp, R. J. (2007). Use it or lose it: Examining preschoolers' difficulty in maintaining and executing a goal. *Developmental Science, 10*(5), 559-564. doi: 10.1111/j.1467-7687.2007.00611.x

Matheson, D., & Spranger, K. (2001). Content analysis of the use of fantasy, challenge, and curiosity in school-based nutrition education programs. *Journal of Nutrition Education, 33*(1), 10-16. doi: 10.1016/s1499-4046(06)60004-3

Matte-Gagné, C., & Bernier, A. (2011). Prospective relations between maternal autonomy support and child executive functioning: Investigating the mediating role of child language ability. *Journal of Experimental Child Psychology, 110*(4), 611-625. doi: 10.1016/j.jecp.2011.06.006

McClelland, M. M., & Cameron, C. E. (2012). Self-regulation in early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Development Perspectives, 6*(2), 136-142. doi: 10.1111/j.1750-8606.2011.00191.x

McLoyd, V. C. (1980). Verbally expressed modes of transformation in the fantasy play of Black preschool children. *Child Development, 51*(4), 1133-1139. doi: 10.2307/1129554

McLoyd, V. C., Warren, D., & Thomas, E. A. C. (1984). Anticipatory and fantastic role enactment in preschool triads. *Developmental Psychology, 20*(5), 807-814. doi: 10.1037/0012-1649.20.5.807

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*(4907), 933-938. doi: 10.1126/science.2658056

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Muller, U., Zelazo, P. D., Lurye, L. E., & Liebermann, D. P. (2008). The effect of labeling on preschool children's performance in the Dimensional Change Card Sort task. *Cognitive Development, 23*(3), 395-408. doi: 10.1016/j.cogdev.2008.06.001

Munakata, Y., & Yerys, B. E. (2001). All together now: When dissociations between knowledge and action disappear. *Psychological Science, 12*(4), 335-337. doi: 10.1111/1467-9280.00361

Nader-Grosbois, N., & Vieillevoye, S. (2012). Variability of self-regulatory strategies in children with intellectual disability and typically developing children in pretend play situations. *Journal of Intellectual Disability Research, 56*(2), 140-156. doi: 10.1111/j.1365-2788.2011.01443.x

Nelson, L. J., Hart, C. H., & Evans, C. A. (2008). Solitary-functional play and solitary-pretend play: Another look at the construct of solitary-active behavior using playground observations. *Social Development, 17*(4), 812-831. doi: 10.1111/j.1467-9507.2008.00470.x

Nguyen, D., Kemp, N., & Want, S. C. (2011). The effects of funny and serious task content and expectations of fun versus importance on children's cognitive performance. *Australian Journal of Psychology, 63*(3), 154-162. doi: 10.1111/j.1742-9536.2011.00014.x

Oh, S., & Lewis, C. (2008). Korean preschoolers' advanced inhibitory control and its relation to other executive skills and mental state understanding. *Child Development, 79*(1), 80-99. doi: 10.1111/j.1467-8624.2007.01112.x

Overton, W. F., & Jackson, J. P. (1973). Representation of imagined objects in action sequences: A developmental study. *Child Development, 44*(2), 309-314. doi: 10.1111/j.1467-8624.1973.tb02156.x

Parker, L. E., & Lepper, M. R. (1992). Effects of fantasy contexts on children's learning and motivation: Making learning more fun. *Journal of Personality and Social Psychology, 62*(4), 625-633. doi: 10.1037/0022-3514.62.4.625

Perner, J., & Lang, B. (2002). What causes 3-year-olds' difficulty on the Dimensional Change Card Sorting task? *Infant and Child Development, 11*(2), 93-105. doi: 10.1002/icd.299

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

Pulaski, M. A. (1973). Toys and imaginative play. In J. L. Singer (Ed.), *The child's world of make-believe: Experimental studies of imaginative play.* (pp. 74-103). Oxford England: Academic Press.

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment, 87*(1), 102-112. doi: 10.1207/s15327752jpa8701_09

Qu, L., & Zelazo, P. D. (2007). The facilitative effect of positive stimuli on 3-year-olds' flexible rule use. *Cognitive Development, 22*(4), 456-473. doi: 10.1016/j.cogdev.2007.08.010

Rakoczy, H., Tomasello, M., & Striano, T. (2006). The role of experience and discourse in children's developing understanding of pretend play actions. *British Journal of Developmental Psychology, 24*, 305-335. doi: 10.1348/026151005X36001

Reed, M. A., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly-Journal of Developmental Psychology, 30*(2), 131-147.

Rubin, K. H., Fein, G. G., & Vandenberg, B. (1983). Play. In E. M. Hetherington (Ed.), *Handbook of child psychology: Socialization, personality, and social development* (4 ed., Vol. 4, pp. 693-774). New York: Wiley.

Smilansky, S., & Shefatya, L. (1990). *Facilitating play: A medium for promoting cognitive, socio-emotional, and academic development in young children*. Gaithersburg, MD: Psychosocial & Educational Publications.

Smith, P. K. (2002). Pretend play, metarepresentation, and theory of mind. In R. W. Mitchell (Ed.), *Pretending and imagination in animals and children* (pp. 129-141). United Kingdom: Cambridge University Press.

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

Taylor, M., & Carlson, S. M. (1997). The relation between individual differences in fantasy and theory of mind. *Child Development, 68*(3), 436-455. doi: 10.1111/j.1467-8624.1997.tb01950.x

Taylor, M., Carlson, S. M., Maring, B. L., Gerow, L., & Charley, C. M. (2004). The characteristics and correlates of fantasy in school-age children: Imaginary companions, impersonation, and social understanding. *Developmental Psychology, 40*(6), 1173-1187. doi: 10.1037/0012-1649.40.6.1173

Towse, J. N., Redbond, J., Houston-Price, C. M. T., & Cook, S. (2000). Understanding the Dimensional Change Card Sort: Perspectives from task success and failure. *Cognitive Development, 15*(3), 347-365. doi: 10.1016/S0885-2014(00)00021-6

Vygotsky, L. S. (1933/1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner & E. Souberman, Trans.). Cambridge, MA: Harvard University Press.

White, R. E. (2012). *Symbolic thought in the service of self-control: Effects of social psychological distancing on executive function in young children*. Unpublished doctoral dissertation. University of Minnesota, MN.

Yerys, B. E., & Munakata, Y. (2006). When labels hurt but novelty helps: Children's perseveration and flexibility in a card-sorting task. *Child Development, 77*(6), 1589-1607. doi: 10.1111/j.1467-8624.2006.00961.x

Youngblade, L. M., & Dunn, J. (1995). Individual differences in young children's pretend play with mother and sibling: Links to relationships and understanding of other people's feelings and beliefs. *Child Development, 66*(5), 1472-1492. doi: 10.2307/1131658

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*(1), 297-301. doi: 10.1038/nprot.2006.46

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*(1), 37-63. doi: 10.1016/S0885-2014(96)90027-1

Zelazo, P. D., & Müller, U. (2011). Executive function in typical and atypical development. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development (2nd ed.)* (pp. 574-603). Chichester: Wiley-Blackwell.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood: I. The development of executive function. *Monographs of the Society for Research in Child Development, 68*(3), 11-27. doi: 10.1111/j.0037-976X.2003.00261.x

Zhou, Q., Chen, S. H., & Main, A. (2012). Commonalities and differences in the research on children's effortful control and executive function: A call for an integrated model of self-regulation. *Child Development Perspectives, 6*(2), 112-121. doi: 10.1111/j.1750-8606.2011.00176.x

## Appendices

### Appendix A: Sample scripts for DCCS Conditions (Study 1)

DCCS Modified Standard Condition – Order 1 (Shape-Color)

**INTRODUCTION**

| | |
|---|---|
| *Set card stand.* | This is my stand, which I will use for the cards in this game.<br>It goes right here in front of me.<br>I'll put each card on this stand, where we can see it,<br>and then I'll hand the card to you when it's your turn, (ok?).<br>So, you stay seated in your chair,<br>and I'll hand you the card when it's time, ok? |
| *Set boxes.* | Now…here is a blue star, and here is a red truck.<br>Now, we're going to play a card game. |

**INITIAL RULES (Shape)**

| | |
|---|---|
| | This is called the shape game.<br>I learned this game from a friend of mine, and I thought that<br>you and I could play it together today!<br>Let's play right here on this table, since we're already set up<br>here with all the pieces.<br>We will play with two shapes: one of them is a star<br>and the other one is a truck.<br>So, in the shape game…. |
| *Gesture to boxes on "here" and "there" throughout.* | …all the stars go here,<br>and all the trucks go there.<br>See, here's a star. |
| *Place card in box.* | So, it goes here.<br>If it's a star, it goes here; if it's a truck, it goes there. |
| | Now, here's a truck.<br>Where does this one go? |

| | |
|---|---|
| **If correct:** | Very good. You know how to play the **SHAPE** game. |
| **If points only:** | Can you put it inside (the box) (for me)?<br>[or: Can you help me put this truck in the box?] |
| **If incorrect:** | No, this one's a truck,<br>so it has to go over here in the **SHAPE** game.<br>Can you help me put this truck down? |

Alright. Now, it's your turn.
So, remember…

*Say for cards 1, 2, & 3*

In this game, if it's a star, it goes here;
if it's a truck, it goes there.

Here's a ________. [*(1) truck, (2) star, (3) truck*]

*Interim Reminder for card 4*

Remember, we're playing the **SHAPE** game.
If it's a star, it goes here; if it's a truck, it goes there.
Here's a ________. [*(4) truck*]

*Say for cards 5 & 6*

In this game, if it's a star, it goes here;
if it's a truck, it goes there.

Here's a ________. [*(5) star, (6) star*]

**RULE SWITCH (Color)**

Now, we're going to play a new game.
We're not going to play the **SHAPE** game anymore.
We're going to play the **COLOR** game.
I learned this game from my friend, too!
Let's keep playing right here on this table,
where we're already set up with all the pieces.
We will play with two colors: one of them is blue,

and the other one is red.
So, in the color game…

…all the blue ones go here,
and all the red ones go there.
Remember…

*Say for cards 1, 2, & 3*

In this game, if it's blue, it goes here;
if it's red, it goes there.

Here's a _____ one. [*(1) red, (2) blue, (3) blue*]

*Interim Reminder for card 4*

Remember, we're playing the **COLOR** game.
If it's blue, it goes here; if it's red, it goes there.
Here's a _____ one. [*(4) red*]

*Say for cards 5 & 6*

In this game, if it's blue, it goes here;
if it's red, it goes there.

Here's a ______ one. [*(5) red, (6) blue*]

We finished all the cards! You did such a great job playing that game with me!

*(END OF TASK!)*

DCCS: Fantasy Planet Condition – Order 1 (Shape-Color)

**INTRODUCTION**

*Set card stand.*

This is my stand, which I will use for the cards in this game.
It goes right here in front of me.
I'll put each card on this stand, where we can see it,
and then I'll hand the card to you when it's your turn, (ok?).
So, you stay seated in your chair,
and I'll hand you the card when it's time, ok?

*Set boxes.*

Now…here is a blue star, and here is a red truck.
Now, we're going to play a card game.

**INITIAL RULES (Shape)**

When we play the game,
let's pretend we're on a faraway planet!

*Use visual aid, pointing to planets and tracing an arc pathway.*

We're not on Planet Earth anymore,
Let's pretend that we have taken a rocket ship all the way over here – to Planet SHAPE!
On this planet, we see lots and lots of shapes everywhere,
like stars and trucks!

*Put visual aid away.*

So, in the Planet Shape game…

*Gesture to boxes on "here" and "there" throughout.*

…all the stars go here,
and all the trucks go there.
See, here's a star.

*Place card in box.*

So, it goes here.
If it's a star, it goes here; if it's a truck, it goes there.

Now, here's a truck.
Where does this one go?

| | |
|---|---|
| **If correct:** | Very good. You know how to play the **PLANET SHAPE** game. |
| **If points only:** | Can you put it inside (the box) (for me)?<br>[or: Can you help me put this truck in the box?] |
| **If incorrect:** | No, this one's a truck,<br>so it has to go over here in the **PLANET SHAPE** game.<br>Can you help me put this truck down? |

Alright. Now, it's your turn.
So, remember…

*Say for cards 1, 2, & 3*

On this planet, if it's a star, it goes here;
if it's a truck, it goes there.

Here's a ________. [*(1) truck, (2) star, (3) truck*]

*Interim Reminder for card 4*

Remember, we're playing the **PLANET SHAPE** game.
If it's a star, it goes here; if it's a truck, it goes there.
Here's a ________. [*(4) truck*]

*Say for cards 5 & 6*

On this planet, if it's a star, it goes here;
if it's a truck, it goes there.

Here's a ________. [*(5) star, (6) star*]

**RULE SWITCH (Color)**

*Use visual aid again when talking about planets.*

Now, we're going to play a new game.
We're not going to play the **Planet Shape** game anymore.
Now, let's pretend we're on a different faraway planet!
Let's pretend that we have taken our rocket ship all the way over here – to Planet COLOR!
On this planet, we see lots and lots of colors everywhere, like blue and red!

*Put visual aid away.*

So in the Planet Color game…

…all the blue ones go here,
and all the red ones go there.
Remember…

*Say for cards 1, 2, & 3*

On this planet, if it's blue, it goes here;
if it's red, it goes there.

Here's a _____ one. [*(1) red, (2) blue, (3) blue*]

*Interim Reminder for card 4*

Remember, we're playing the **PLANET COLOR** game.
If it's blue, it goes here; if it's red, it goes there.
Here's a _____ one. [*(4) red*]

*Say for cards 5 & 6*

On this planet, if it's blue, it goes here;
if it's red, it goes there.

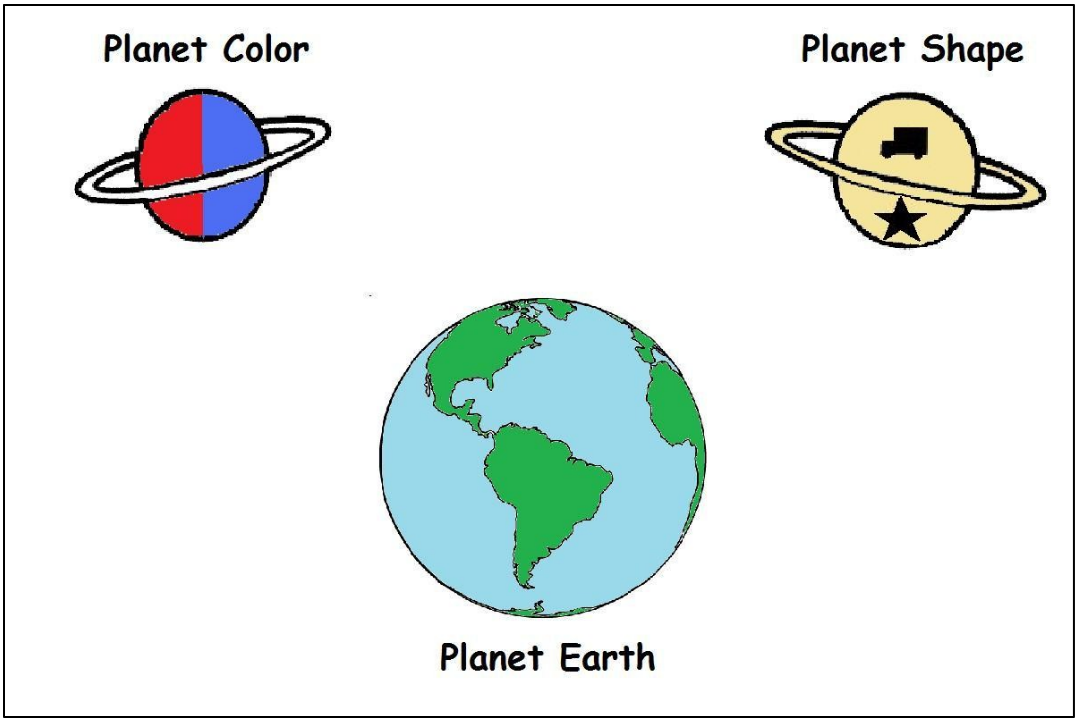
Here's a ______ one. [*(5) red, (6) blue*]

We finished all the cards! You did such a great job playing that game with me!

*(END OF TASK!)*

## Appendix B: DCCS Fantasy Planet Visual Aid (Study 1)

This image was presented in color as an 8.5” x 11” laminated visual aid in the Fantasy Planet condition:




Planet Color: red on the left half and blue on the right half with white ring
Planet Shape: pale yellowish-beige with black shapes
Planet Earth: blue water and green continents

## Appendix C: Scripts for Forward and Backward Digit Span (FDS/BDS)

### FORWARD DIGIT SPAN (ERNIE)

Let's play a game where you copy me, OK? (get Ernie)

Here's my friend Ernie. (says hi)

Ernie's going to copy me first, so whatever I say, he says it too!

Like this. If I say 1, 2, 3, Ernie says, 1, 2, 3 (diff voice). Like that!

Now you try. (put Ernie out of view)

I'll say some numbers and then you say those numbers, just exactly the same way I did.

Practice
Let's practice: 4, 2 **C's Final Response: __________** **# tries_____**
(correct as needed, up to 6 tries, then do test trials)

Great! Let's do some more like that.

Remember, whatever I say, you say it the same.

Test trials

Digits: (circle correct items; strike out failures; move up a level after a correct answer; discontinue after three consecutive failures)

| | 2-digits | | 3-digits | 4-digits | 5-digits |
|---|---|---|---|---|---|
| **Fail** | 1, 4 | **Pass→** | 6, 4, 2 | 1, 2, 4, 6 | 4, 2, 1, 3, 5 |
| ↓ | 5, 3 | | 3, 5, 1 | 3, 1, 5, 2 | 5, 6, 4, 3, 2 |
| | 2, 6 | | 2, 3, 6 | 6, 5, 3, 4 | 2, 3, 4, 1, 6 |

| 6-digits | 7-digits |
|---|---|
| 1, 2, 5, 9, 3, 7 | 9, 7, 2, 1, 5, 4, 6 |
| 7, 2, 5, 2, 1, 4 | 7, 4, 6, 2, 9, 3, 1 |
| 5, 4, 2, 8, 3, 6 | 8, 4, 6, 7, 2, 5, 9 |

**Highest Level on FDS: _____**
(Note to coders: If C could not do Level 2, Highest Level on FDS = 1.)

## BACKWARD DIGIT SPAN (ERNIE)

Now I have a new game where we say things backward. (get Ernie)

Here's Ernie again. (says hi)

Now Ernie's being silly, so whatever I say, he says it backward. He says it the wrong way.

Like this, if I say the numbers "1, 2" Ernie says "2, 1" (diff. voice) Like that!

Isn't that silly?

Now you try. (put Ernie out of view)

Whatever I say, you say it backward, OK?

Practice
Let's practice: 1, 2 **C's Final Response**: ___________ **# tries**_____
(correct as needed, up to 6 tries, then do test trials)

Great! Let's do some more like that.

Test Trials

Remember, whatever I say, you say it backward.

Digits: (circle correct items; strike out failures; move up a level after a correct answer; discontinue after three consecutive failures)

| | 2-digits | | 3-digits | 4-digits | 5-digits |
|---|---|---|---|---|---|
| **Fail** ↓ | 1, 4 | **Pass**⟶ | 6, 4, 2 | 1, 2, 4, 6 | 4, 2, 1, 3, 5 |
| | 5, 3 | | 3, 5, 1 | 3, 1, 5, 2 | 5, 6, 4, 3, 2 |
| | 2, 6 | | 2, 3, 6 | 6, 5, 3, 4 | 2, 3, 4, 1, 6 |

**Highest Level on BDS:** _____

(Note to coders: If C could not do Level 2, Highest Level on BDS = 1)

**Appendix D: Script for Bear-Dragon**

# BEAR/DRAGON TASK

Warm-up

(make sure **C**hild can follow all of the directions by having C do the following; **E**xperimenter models the movements during warm-up.)

Before we start playing our next game, I'm going to ask you to do some silly things with me, like:

Stick out your tongue. (E remember to model)

Touch your ears.

Touch your teeth.

Touch your eyes.

Clap your hands.

Touch your feet.

Touch your head.

Touch your tummy.

Touch your nose.

Wave your hand.

Good job! You are ready to play this game…

Instructions

(E takes out the bear and dragon puppets)

I have a game to play with these puppets.
This puppet (show the bear) is a nice bear. When he talks to us, we will do what he tells us to do. This puppet (show the dragon) is not very nice. This puppet is a dragon. When he talks to us, we won't listen to him. If he tells us to do something, we won't do it.

Practice Trials

Let’s try some/practice.  We’ll try it first with the good bear.

(showing Bear) Here’s the good bear.  He says, "Touch your nose." (in a mellow, nice voice)

**Touch (1)**　　　**No touch (0)**　　　**# tries_____**

(If C does <u>not</u> touch nose)

Remember, we listen to the nice bear and do what he says because that's how we play the game.

(E repeats the command and may model the action until C succeeds)

(If C touches nose)

Very good!  Now let's practice with the naughty dragon.  In this game, we won't do what the dragon (tells/)asks us to do because isn’t (very) nice.

DRAGON:  Touch your tummy.  (in a low, gruff voice)

**Touch (1)**　　　**No touch (0)**　　　**# tries_____**

(If C touches stomach)

Remember, we are not going to listen to the mean dragon.  We won't do what he says because that's how we play the game.

(Repeat until C gets it right, holding C's hands down if necessary on the 6th try)

> (Specific script for repetitions was implemented in Study 2:)
>
> Try #2:  Remember, we are not going to listen to the mean dragon.  We won't do what he says because that's how we play the game.
>
> Try #3:  In this game, we don’t do what the dragon tells us to do.  If he says to do something, we won’t do it.
>
> Try #4:  Remember how we play this game:  if the dragon tells us to do something, we do not listen to him.  That’s how we play this game.
>
> Try #5:  Can you put your hands on the table?  Now, when the dragon tells you to do something, don’t do what he says.  Just keep your hands on the table instead.
>
> Try #6:  Put your hands on the table again.  Now, I’ll just gently hold your hands down with my hands.  So, when the dragon tells us to do something, we won’t do it.  We’ll keep our hands down here instead.

(After 6 tries or when C succeeds at both practice trials, continue)

Very good! Alright.

Rule Check

So, in this game…
When the Bear tells you to do something, do you do it? **Yes** **No**
(correct if necessary) # tries _____

And when the Dragon tells you to do something, do you do it? **Yes** **No**
(correct if necessary) # tries _____

OK, here we go.

Test Trials

(E does not model or give feedback on test trials)

1. **BEAR:** Stick out your tongue.

2. **DRAGON:** Touch your ears.

3. **BEAR:** Touch your teeth.

4. **DRAGON:** Touch your eyes.

5. **BEAR:** Clap your hands.

(Reminder given regardless of performance)
Remember the way we play this game. We do what the bear tells us to do because he is nice but we won't do what the dragon tells us to do because he isn't very nice.

6. **DRAGON:** Touch your feet.

7. **BEAR:** Touch your head.

8. **DRAGON:** Touch your tummy.

9. **BEAR:** Touch your nose.

10. **DRAGON:** Wave your hand.

Very good!

**Appendix E: Sample scripts for DCCS Conditions, Level 4 (Study 2)**

DCCS: Pretend Playground Condition – Order 1 (Shape, Color)

**INTRODUCTION**

| | |
|---|---|
| *Set boxes.* | Here is a blue star, and here is a red truck.<br>Now, we're going to play a card game. |

**INITIAL RULES (Shape)**

| | |
|---|---|
| | When we play the game,<br>let's pretend we're at a faraway playground!<br>We're not in this playroom anymore,<br>Let's pretend that we have driven our car all the way over to the **SHAPE** Playground!<br>At this playground, we see lots and lots of **shapes** everywhere, like **stars** and **trucks**!<br>So, in the **Shape** Playground game… |
| *Gesture to boxes on "here" and "there" throughout.* | …all the stars go here,<br>and all the trucks go there.<br>See, here's a star. |
| *Place card in box.* | So, it goes here.<br>If it's a star, it goes here; if it's a truck, it goes there. |
| | Now, here's a truck.<br>Where does this one go? |
| **If correct:** | Very good. You know how to play the **SHAPE PLAYGROUND** game. |
| **If points only:** | Can you put it inside (the box) (for me)?<br>[or: Can you help me put this truck in the box?] |
| **If incorrect:** | No, this one's a truck, so it has to go over here in the **SHAPE PLAYGROUND** game.<br>Can you help me put this truck down? |

Alright. Now, it's your turn.
So, remember…

*Say for cards 1, 2, & 3*

At this playground, if it's a star, it goes here;
if it's a truck, it goes there.

Here's a ________. [*(1) truck, (2) star, (3) truck*]

*Interim Reminder for card 4*

Remember, we're playing the **SHAPE PLAYGROUND** game.
If it's a star, it goes here; if it's a truck, it goes there.
Here's a ________. [*(4) truck*]

*Say for cards 5 & 6*

At this playground, if it's a star, it goes here;
if it's a truck, it goes there.

Here's a ________. [*(5) star, (6) star*]

**RULE SWITCH (Color)**

Now, we're going to play a new game.
We're not going to play the **Shape Playground** game anymore.
Now, let's pretend we're at a different faraway playground!
Let's pretend that we have driven our car all the way over to the **COLOR** playground!
At this playground, we see lots and lots of **colors** everywhere, like **blue** and **red**!
So in the **Color** Playground game…

…all the blue ones go here,
and all the red ones go there.
Remember…

*Say for cards 1, 2, & 3*

At this playground, if it's blue, it goes here;
if it's red, it goes there.

Here's a _____ one. [*(1) red, (2) blue, (3) blue*]

*Interim Reminder for card 4*

Remember, we're playing the **COLOR PLAYGROUND** game.
If it's blue, it goes here; if it's red, it goes there.
Here's a _____ one. [*(4) red*]

*Say for cards 5 & 6*

At this playground, if it's blue, it goes here;
if it's red, it goes there.

Here's a ______ one. [*(5) red, (6) blue*]

We finished all the cards! You did such a great job playing that game with me!

*(END OF TASK!)*

DCCS: Non-Pretense Control Condition – Order 1 (Shape, Color) – for boys*

* Boys' names were changed to girls' names (Julia and Samantha) for female participants

**INTRODUCTION**

*Set boxes.*

Here is a blue star, and here is a red truck.
Now, we're going to play a card game.

**INITIAL RULES (Shape)**

This is called the **shape** game.
I learned this game from my friend **Jonathan**, and I thought that you and I could play it together today!
He taught me how to play this way, and now I can teach you!
We will play with two **shapes**: one is a **star** and the other one is a **truck**.
So, in **Jonathan's shape** game….

*Gesture to boxes on "here" and "there" throughout.*

…all the stars go here, and all the trucks go there.
See, here's a star.

*Place card in box.*

So, it goes here.
If it's a star, it goes here; if it's a truck, it goes there.

Now, here's a truck.
Where does this one go?

---

**If correct:** Very good. You know how to play **Jonathan's SHAPE** game.

---

**If points only:** Can you put it inside (the box) (for me)?
[or: Can you help me put this truck in the box?]

---

**If incorrect:** No, this one's a truck, so it has to go over here in **Jonathan's SHAPE** game.
Can you help me put this truck down?

---

Alright. Now, it's your turn.

So, remember…

| | |
|---|---|
| *Say for cards 1, 2, & 3* | In **Jonathan's game**, if it's a star, it goes here;<br>if it's a truck, it goes there.<br><br>Here's a ________. [*(1) truck, (2) star, (3) truck*] |
| *Interim Reminder for card 4* | Remember, we're playing **Jonathan's SHAPE** game.<br>If it's a star, it goes here; if it's a truck, it goes there.<br>Here's a ________. [*(4) truck*] |
| *Say for cards 5 & 6* | In **Jonathan's** game, if it's a star, it goes here;<br>if it's a truck, it goes there.<br><br>Here's a ________. [*(5) star, (6) star*] |

**RULE SWITCH (Color)**

Now, we're going to play a new game.
We're not going to play **Jonathan's SHAPE** game anymore.
We're going to play the **COLOR** game.
I learned this game from my **different** friend, **Samuel**!
He taught me how to play this way, and now I can teach you!
We will play with two **colors**: one is **blue**,
and the other one is **red**.
So, in **Samuel's color** game…

…all the blue ones go here,
and all the red ones go there.
Remember…

| | |
|---|---|
| *Say for cards 1, 2, & 3* | In **Samuel's game**, if it's blue, it goes here;<br>if it's red, it goes there.<br><br>Here's a _____ one. [*(1) red, (2) blue, (3) blue*] |

| | |
|---|---|
| *Interim Reminder for card 4* | Remember, we're playing **Samuel's COLOR** game.<br>If it's blue, it goes here; if it's red, it goes there.<br>Here's a _____ one. [*(4) red*] |
| *Say for cards 5 & 6* | In **Samuel's game**, if it's blue, it goes here;<br>if it's red, it goes there.<br>Here's a ______ one. [*(5) red, (6) blue*] |

We finished all the cards!  You did such a great job playing that game with me!

*(END OF TASK!)*

**Appendix F: Sample scripts for each condition - the EF Scale, Levels 3 & 5 (Study 2)**

Level 3 Pre-switch (following Pretend Playground, Color trials):

## Two Rules w/ Separated Stimuli, Color Training

Okay, let's play a new game with cards!

*Introduction of Stimuli:*

I have these boxes here. (*Box on E's left has a* ***pink card with a black heart on it****. Box on E's right has a* ***yellow card with a black flower on it****.)*

This one is pink *(point to top of box)*
and this one is yellow *(point)*.

Now, look what I have here! I have pink ones and yellow ones. *(show example cards one at a time above boxes in center and then return them to the stack in experimenter's hand so they are not showing)*

***Demonstration Trials:***

Let's pretend we're at the **Color Playground**. At the Color Playground, all the pink ones go here *(pointing)* and all the yellow ones go there *(pointing)*.

See, here's a pink one *(hold up in center)*. It goes in the pink box *(put it in)*.
And here's a yellow one *(hold up in center)*. It goes in the yellow box *(put it in)*.

Now it's going to be your turn to play the **Color Playground game**!

***<u>Rule Check:</u>***

Can you show me where the pink ones go in the **Color Playground game**?

> **If Correct:** Very good, that's right.
> **If Incorrect:** Uh oh. Remember, **at the Color Playground**, if it's pink, it goes here; if it's yellow, it goes there *(pointing to appropriate boxes)*. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response: Correct(1) Incorrect(0)**

**Number of Tries:___________**

Can you show me where the yellow ones go in the **Color Playground game**?

> **If Correct:** Very good, that's right.
> **If Incorrect:** Uh oh. Remember, **at the Color Playground**, if it's pink, it goes here; if it's yellow, it goes there *(pointing to appropriate boxes)*. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response: Correct(1) Incorrect(0)**

**Number of Tries:___________**

Okay, let's play!

## <u>Level 3 Pre-switch:</u>
## Two Rules w/ Separated Stimuli, Color Test Trials

*E says rule <u>before</u> presenting card on each trial.*

E: **At this playground**, if it's pink, it goes here; if it's yellow, it goes there *(while pointing to appropriate boxes).*

1) *(yellow/heart)* Here's a yellow one. Pink _____ **Yellow _____**

E: **At this playground**, pink ones go here, yellow ones go there *(while pointing to appropriate boxes).*

2) *(pink/flower)* Here's a pink one. **Pink_____** Yellow _____

E: **At this playground**, if it's pink, it goes here; if it's yellow, it goes there *(while pointing to appropriate boxes).*

3) *(pink/flower)* Here's a pink one. **Pink _____** Yellow_____

E: **At this playground**, pink ones go here, yellow ones go there *(while pointing to appropriate boxes).*

4) *(yellow/heart)* Here's a yellow one. Pink _____ **Yellow _____**

E: **At this playground**, pink ones go here, yellow ones go there *(while pointing to appropriate boxes).*

5) *(yellow/heart)* Here's a yellow one. Pink _____ **Yellow _____**

**Total Correct: _______**

***CONTINUE only if child gets at least 4/5 correct. Otherwise, STOP***

## Level 3 Post-switch: Two Rules w/ Separated Stimuli, Shape Training

Now we're going to play a different game. We're not going to play the Color Playground game anymore. Let's pretend that we have driven our car all the way over to the **Shape Playground**.

**At the Shape Playground**, all the hearts go here (*pointing*), and all the flowers go there (*pointing*).

***(No demonstration trials or rule checks.)***

Okay, let's play!

**Level 3 Post-switch:**
**Two Rules w/ Separated Stimuli, Shape Test Trials**

*E says rule before presenting card on each trial.*

E: **At this playground**, if it's a heart, it goes here; if it's a flower, it goes there *(while pointing to appropriate boxes).*

1) *(pink/flower)* Here's a flower. Heart____ **Flower** _____

E: **At this playground**, hearts go here, flowers go there (*while pointing to boxes*).

2) *(yellow/heart)* Here's a heart. **Heart** ____ Flower _____

E: **At this playground**, if it's a heart, it goes here; if it's a flower, it goes there *(while pointing to appropriate boxes).*

3) *(pink/flower)* Here's a flower. Heart ____ **Flower** _____

E: **At this playground**, hearts go here, flowers go there (*while pointing to boxes*).

4) *(pink/flower)* Here's a flower. Heart ____ **Flower** _____

E: **At this playground**, hearts go here, flowers go there (*while pointing to boxes*).

5) *(yellow/heart)* Here's a heart. **Heart** ____ Flower _____

**Total Correct: ________**

[END OF TASK]

Level 3 Pre-switch (following Non-Pretense Control, Color trials, Boy):

## Two Rules w/ Separated Stimuli, Color Training

Okay, let's play a new game with cards!

*Introduction of Stimuli:*

I have these boxes here. (*Box on E's left has a **pink card with a black heart on it**. Box on E's right has a **yellow card with a black flower on it**.)*

This one is pink *(point to top of box)*
and this one is yellow *(point)*.

Now, look what I have here! I have pink ones and yellow ones. *(show example cards one at a time above boxes in center and then return them to the stack in experimenter's hand so they are not showing)*

***Demonstration Trials:***

This is my friend **Samuel's Color game**. In **Samuel's Color game**, all the pink ones go here *(pointing)* and all the yellow ones go there *(pointing)*.

See, here's a pink one *(hold up in center)*. It goes in the pink box *(put it in)*.
And here's a yellow one *(hold up in center)*. It goes in the yellow box *(put it in)*.

Now it's going to be your turn to play **Samuel's Color game**!

***<u>Rule Check:</u>***

Can you show me where the pink ones go in **Samuel's Color game**?

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, in **Samuel's Color game**, if it's pink, it goes here; if it's yellow, it goes there *(pointing to appropriate boxes)*. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response:** **Correct(1)** **Incorrect(0)**

**Number of Tries:___________**

Can you show me where the yellow ones go in **Samuel's Color game**?

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, in **Samuel's Color game**, if it's pink, it goes here; if it's yellow, it goes there *(pointing to appropriate boxes)*. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response:** **Correct(1)** **Incorrect(0)**

**Number of Tries:___________**

Okay, let's play!

## <u>Level 3 Pre-switch</u>
## Two Rules w/ Separated Stimuli, Color Test Trials

*E says rule <u>before</u> presenting card on each trial.*

E: **In Samuel's game**, if it's pink, it goes here; if it's yellow, it goes there *(while pointing to appropriate boxes).*

1) *(yellow/heart)* Here's a yellow one. Pink _____ **Yellow _____**

E: **In Samuel's game**, pink ones go here, yellow ones go there *(while pointing to appropriate boxes).*

2) *(pink/flower)* Here's a pink one. **Pink_____** Yellow _____

E: **In Samuel's game**, if it's pink, it goes here; if it's yellow, it goes there *(while pointing to appropriate boxes).*

3) *(pink/flower)* Here's a pink one. **Pink _____** Yellow_____

E: **In Samuel's game**, pink ones go here, yellow ones go there *(while pointing to appropriate boxes).*

4) *(yellow/heart)* Here's a yellow one. Pink _____ **Yellow _____**

E: **In Samuel's game**, pink ones go here, yellow ones go there *(while pointing to appropriate boxes).*

5) *(yellow/heart)* Here's a yellow one. Pink _____ **Yellow _____**

**Total Correct: _______**

***CONTINUE only if child gets at least 4/5 correct. Otherwise, STOP***

## Level 3 Post-switch: Two Rules w/ Separated Stimuli, Shape Training

Now we're going to play a different game. We're not going to play **Samuel's Color** game anymore. We're going to play my friend **Jonathan's Shape game**.

**In Jonathan's Shape game**, all the hearts go here (*pointing*), and all the flowers go there (*pointing*).

***(No demonstration trials or rule checks.)***

Okay, let's play!

## Level 3 Post-switch:
## <u>Two Rules w/ Separated Stimuli, Shape Test Trials</u>

*E says rule before presenting card on each trial.*

E: **In Jonathan's game**, if it's a heart, it goes here; if it's a flower, it goes there *(while pointing to appropriate boxes).*

1) *(pink/flower)* Here's a flower. Heart____ **Flower** _____

E: **In Jonathan's game**, hearts go here, flowers go there (*while pointing to boxes*).

2) *(yellow/heart)* Here's a heart. **Heart** ____ Flower _____

E: **In Jonathan's game**, if it's a heart, it goes here; if it's a flower, it goes there *(while pointing to appropriate boxes).*

3) *(pink/flower)* Here's a flower. Heart ____ **Flower** _____

E: **In Jonathan's game**, hearts go here, flowers go there (*while pointing to boxes*).

4) *(pink/flower)* Here's a flower. Heart ____ **Flower** _____

E: **In Jonathan's game**, hearts go here, flowers go there (*while pointing to boxes*).

5) *(yellow/heart)* Here's a heart. **Heart** ____ Flower _____

**Total Correct: ________**

[END OF TASK]

Level 5 – Pretend Playground:

## Two Rules Mixed, Training

Okay, let's play a new game with cards!

Continuing from Level 4: *(slowly explain)*

In this game, we sometimes **pretend we're at the Shape Playground**, and we sometimes **pretend we're at the Color Playground**. I'm going to show you a card, and I'll say, "Pretend we're at the Shape Playground" or "Pretend we're at the Color Playground."

***Demonstration Trials:***

If I say, "Pretend we're at the Shape Playground," you pretend we're at the Shape Playground. At the Shape Playground, all the stars go here (*point*) and all the trucks go there (*point*).

See, here's a star *(hold up in center)*. It goes in the star box *(put it in)*.
And here's a truck *(hold up in center)*. It goes in the truck box *(put it in)*.

If I say, "Pretend we're at the Color Playground," you pretend we're at the Color Playground. At the Color Playground, all the blue ones go here (*point*) and all the red ones go there (*point*).

See, here's a blue one *(hold up in center)*. It goes in the blue box *(put it in)*.
And here's a red one *(hold up in center)*. It goes in the red box *(put it in)*.

Now it's going to be your turn to play this game!

***<u>Rule Check:</u>*** *(E is not showing a card during these checks*)

So what game do you play if I say, "Pretend we're at **the Shape Playground**?" *(shape game).*

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, if I say shape, play the Shape Playground game. If I say color, play the Color Playground game. *(Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response: Correct(1) Incorrect(0)**

**Number of Tries:___________**

And how do you play the **Shape Playgound game**? (stars go here, trucks go there)

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, **at the Shape Playground**, if it's a star, it goes here; if it's a truck, it goes there *(pointing to appropriate boxes)*. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response: Correct(1) Incorrect(0)**

**Number of Tries:___________**

So what game do you play if I say, "Pretend we're at **the Color Playground**?" *(color game).*

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, if I say shape, play the Shape Playground game. If I say color, play the Color Playground game. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response: Correct(1) Incorrect(0)**

**Number of Tries:___________**

And how do you play the **Color Playground game**? (blue ones go here, red ones go there)

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, **at the Color Playground**, if it's blue, it goes here; if it's red, it goes there *(pointing to appropriate boxes)*. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response:** **Correct(1)** **Incorrect(0)**

**Number of Tries:___________**

Okay let's play!

**Level 5: Pretend Playground**
**Two Rules Mixed, Test Trials**
*Shaded trials indicate color trials.*

*On each trial, hold up a card and say,* "Pretend we're at the (Shape/Color) Playground."

| | | |
|---|---|---|
| 1) *(red star)* Pretend we're at the Color Playground. | Blue_____ | **Red____** |
| 2) *(red star)* Pretend we're at the Shape Playground. | **Star_____** | Truck___ |
| 3) *(blue truck)* Pretend we're at the Color Playground. | **Blue_____** | Red_____ |
| 4) *(red star)* Pretend we're at the Color Playground. | Blue_____ | **Red____** |
| 5) *(red star)* Pretend we're at the Shape Playground. | **Star_____** | Truck___ |
| 6) *(blue truck)* Pretend we're at the Shape Playground. | Star_____ | **Truck___** |
| 7) *(blue truck)* Pretend we're at the Color Playground. | **Blue_____** | Red_____ |
| 8) *(blue truck)* Pretend we're at the Color Playground. | **Blue_____** | Red_____ |
| 9) *(blue truck)* Pretend we're at the Shape Playground. | Star_____ | **Truck___** |
| 10) *(red star)* Pretend we're at the Shape Playground. | **Star_____** | Truck___ |

**Total Shape Correct: _______**

**Total Color Correct: _______**

[END OF TASK]

Level 5 – Non-Pretense Control (Boy):

## <u>Two Rules Mixed, Training</u>

Okay, let's play a new game with cards!

Continuing from Level 4: *(slowly explain)*

In this game, we sometimes play my friend **Jonathan's Shape Game**, and we sometimes play my friend **Samuel's Color Game**. I'm going to show you a card, and I'll say, "Play Jonathan's Shape Game" or "Play Samuel's Color Game."

***<u>Demonstration Trials:</u>***

If I say, "Play Jonathan's Shape Game," you play Jonathan's Shape Game. In Jonathan's Shape Game, all the stars go here (*point*) and all the trucks go there (*point*).

See, here's a star *(hold up in center)*. It goes in the star box *(put it in)*.
And here's a truck *(hold up in center)*. It goes in the truck box *(put it in)*.

If I say, "Play Samuel's Color Game," you play Samuel's Color Game. In Samuel's Color Game, all the blue ones go here (*point*) and all the red ones go there (*point*).

See, here's a blue one *(hold up in center)*. It goes in the blue box *(put it in)*.
And here's a red one *(hold up in center)*. It goes in the red box *(put it in)*.

Now it's going to be your turn to play this game!

***<u>Rule Check:</u>*** *(E is not showing a card during these checks***)**

So what game do you play if I say, "Play **Jonathan's Shape Game**?" *(shape game).*

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, if I say shape, play Jonathan's Shape Game. If I say color, play Samuel's Color Game. *(Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response:** **Correct(1)** **Incorrect(0)**

**Number of Tries:___________**

And how do you play **Jonathan's Shape Game**? (stars go here, trucks go there)

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, in Jonathan's Shape game, if it's a star, it goes here; if it's a truck, it goes there *(pointing to appropriate boxes)*. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response:** **Correct(1)** **Incorrect(0)**

**Number of Tries:___________**

So what game do you play if I say, "Play **Samuel's Color Game**?" *(color game).*

**If Correct:** Very good, that's right.

**If Incorrect:** Uh oh. Remember, if I say shape, play Jonathan's Shape Game. If I say color, play Samuel's Color Game. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response:** **Correct(1)** **Incorrect(0)**

**Number of Tries:___________**

And how do you play **Samuel's Color Game**? (blue ones go here, red ones go there)

> **If Correct:** Very good, that's right.
> **If Incorrect:** Uh oh. Remember, in Samuel's Color Game, if it's blue, it goes here; if it's red, it goes there *(pointing to appropriate boxes)*. (*Repeat question. If again incorrect, give reminder 1 more time only, then mark as incorrect and continue.)*

**Response: Correct(1) Incorrect(0)**

**Number of Tries:___________**

Okay let's play!

**Level 5: Non-Pretense Control**
**Two Rules Mixed, Test Trials**
*Shaded trials indicate color trials.*

*On each trial, hold up a card and say,* "Play (Jonathan's Shape/Samuel's Color) game."

| | | |
|---|---|---|
| 1) *(red star)* Play Samuel's Color game. | Blue_____ | **Red_____** |
| 2) *(red star)* Play Jonathan's Shape game. | **Star_____** | Truck_____ |
| 3) *(blue truck)* Play Samuel's Color game. | **Blue_____** | Red_____ |
| 4) *(red star)* Play Samuel's Color game. | Blue_____ | **Red_____** |
| 5) *(red star)* Play Jonathan's Shape game. | **Star_____** | Truck_____ |
| 6) *(blue truck)* Play Jonathan's Shape game. | Star_____ | **Truck_____** |
| 7) *(blue truck)* Play Samuel's Color game. | **Blue_____** | Red_____ |
| 8) *(blue truck)* Play Samuel's Color game. | **Blue_____** | Red_____ |
| 9) *(blue truck)* Play Jonathan's Shape game. | Star_____ | **Truck_____** |
| 10) *(red star)* Play Jonathan's Shape game. | **Star_____** | Truck_____ |

**Total Shape Correct: _______**

**Total Color Correct: _______**

[END OF TASK]