Individual Differences in Flexibility of Delay and Saving Behavior: Relations to Executive Function

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Wendy Seong-Cheng Lee

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

Dr. Stephanie M. Carlson, Dr. Ann S. Masten

May 2014

Acknowledgements

I would like to thank my adviser Dr. Stephanie M. Carlson for her guidance and support, for challenging me to think critically and experimentally, and for her unfailing excitement and faith in this project from start to end. I would also like to thank my co-adviser Dr. Ann S. Masten for her support and encouragement, especially during the early years, and for helping me see the importance of individual differences and developmental contexts, which influenced and informed my ideas for this dissertation work.

I am grateful to my other committee members Dr. Philip D. Zelazo and Dr. Gerald August for their feedback and support throughout the dissertation process.

I thank members of the lab, in particular the research staff, Catherine Schaefer and Joshua Harrod, and the various research assistants who have worked on this project, especially Hannah Saunders who assisted in data collection.

I thank the hundreds of families who took the time to come participate in my research and all the children who came to happily play games with me.

I am grateful to all my friends both near and far who have supported me unwaveringly throughout the years. In particular, I thank Donaya Hongwanishkul, who helped me in my journey to graduate school, and I am forever grateful to Sally Kuo, my friend, my co-instructor, and my partner-in-crime, who helped me every step of the way up to this moment.

Finally I thank my family, my parents who have sacrificed so much in their lives for their children and continue to support me every day, and my brother, who takes care of and indulges his silly sister who refuses to grow up!

Abstract

Delay of gratification is a complex decision-making behavior that is influenced by many contextual variables, such as cultural values, prior experience, and social trust. Research has shown that children's delay behaviors are sensitive to these variables, which can explain seemingly irrational behavior. It also suggests that immediate consumption does not always indicate poor impulse control. However, previous studies have examined only group differences, which neglect important individual differences in the ability to modify delay behavior. In two studies, we examined children's ability to adapt their delay and saving behavior according to the context and recent experience. It was predicted that children's ability to switch their behavior would be related to greater executive function. In Study 1 ($N = 140$), 3.5- and 4.5-year-old children were categorized as delayers or non-delayers based on a baseline delay choice task. In a second administration of the task, a risk of losing treats was associated with children's preferred choice (i.e., delaying or not delaying), encouraging children to switch their behavior. Children were again categorized as delayers or non-delayers. In Study 2 ($N = 142$), 3.5- and 4.5-year-old children were categorized as savers or spenders based on a baseline saving task where children could save marbles from a small marbles game for a later big marbles game. Children were then unable to play with the big game either because they had no marbles or because the big game was unexpectedly broken. In a second administration, children were again categorized as savers or spenders. In both studies, children who changed categorization across the two administrations, indicating a switch in behavior (e.g., delayer to non-delayer, spender to saver) scored higher on a measure of set-shifting. The results demonstrate that children can use the context and past experience to successfully adapt their delay and saving behaviors, and that executive function skills may be important in facilitating flexibility in these behaviors.

Table of Contents

List of Tables

List of Figures

Introduction

The ability to resist the temptation of immediate distractions and forgo immediate pleasure in pursuit of long-term goals has relevance for many domains of functioning from health (e.g., addiction, nutrition, exercise), finances (e.g., spending, saving, investing), relationships (e.g., marriage, parenting), to education and career (e.g., studying, working). Delay of gratification and the self-control processes that underlie it has roots in early childhood. In their classic laboratory paradigm, Mischel and colleagues measured how long preschool-aged children were able to wait when given the choice between having one marshmallow now and waiting for two marshmallows later (e.g., Mischel, 1974; Mischel et al., 2011; Mischel, Shoda, & Rodriguez, 1989). Younger children had difficulty waiting for delayed rewards, and often preferred strategies that made it more difficult to wait (Mischel & Mischel, 1983). Children's ability to delay improved with age (e.g., Mischel & Metzner, 1962; Mischel & Patternson, 1976) as they began using more effective strategies such as self- distraction (e.g., Mischel, 1974; Mischel, Ebbesen, & Raskoff Zeiss, 1972) and reappraisal or reframing the reward or situation (e.g., Mischel & Baker, 1975).

Mischel and colleagues discovered that individual differences in wait times and delay behavior during early childhood predicted a range of developmental outcomes into adolescence and adulthood, including SAT scores and academic competence, self-regulation, effective coping with stress and frustration, and positive peer relations (Ayduk et al., 2000; Ayduk et al., 2008; Eigsti et al., 2006; Mischel, Shoda & Peake 1988; Shoda, Mischel, Peake, 1990). Remarkably, Casey et al. (2011) found that children's ability to resist temptation at 4 years of age predicted their impulse control to positive emotional stimuli 40 years later, suggesting that there are long-term stable individual differences in delay of gratification.

Other researchers have reported similar outcomes associated with delay of gratification using various delay tasks, such as greater peer competence and peer relations (e.g., Olson, 1989; Olson & Hoza, 1993; Raver, Blackburn, Bancroft, & Torp, 1999), decreased antisocial and externalizing behaviors (e.g., Krueger, Caspi, Moffitt, & White, 1996; Trentacosta & Shaw, 2009), higher IQ (e.g., Funder & Block, 1989), and greater academic achievement and work habits (e.g., Duckworth & Seligman, 2005; Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009).

Delay of gratification has clear significant social and economic benefits for both individuals and society; yet it appears to be a consistent struggle for not only children but many adults. Instead of saving and investing our money for the future, we spend it on frivolous purchases. In the new year, we resolve to eat well and exercise, but give up by the third week. If rational behavior and action should maximize utility (e.g., Herrnstein, 1990), how do we explain failures to delay gratification?

The underlying assumption is that the rational choice is always to delay or wait for the larger reward, and therefore, failures are interpreted as either irrational behavior or the consequence of deficient self-control and willpower. However, this is a misguided assumption. Whether a smaller immediate reward or a later larger reward is more adaptive is determined by the intersection of proximal and distal factors that contextualizes the problem in important ways. For example, in a situation where we do not trust we will receive the later reward, or we have greater immediate needs, the wise choice would be to take the immediate reward. Failure to delay does not always indicate poor self-control, but may reflect a rational response within a given context (e.g., Fawcett, McNamara, & Houston, 2012; McGuire & Kable, 2012, 2013). Moreover, while being undercontrolled is often viewed negatively, being overcontrolled in these types of situations can also be maladaptive (Funder & Block, 1989).

Contextual Influences on Delay Behavior

There has been recent renewed interest in understanding the rationality behind delay behavior, or lack of delay behavior. In particular, there has been a focus on factors such as personal and cultural values (Carlson & Zelazo, 2011), social mistrust (Kidd, Palmeri, & Aslin, 2012; Michaelson, de la Vega, Chatham, & Munakata, 2013), and prior experience of reinforcement (McGuire & Kable, 2012, 2013), which may explain seemingly irrational impulsive behavior. However, the complexity of delay behavior has long been acknowledged and confirmed by early experimental findings. Delay of gratification can be influenced by many person and situation variables, which themselves can interact to determine behavior (e.g., Herzberger & Dweck, 1978; Metcalfe & Mischel, 1999; Mischel, 1966; 1974; Mischel & Shoda, 1995).

For example, the dimensions that define the task itself directly affect motivation to delay. According to the expectancy-value model, the likelihood of delaying depends on the combination of the subjective value of the rewards presented and the expected probability of receiving the later rewards (e.g., Mischel, 1966). With greater delays, the expectancy associated with later rewards decreases, as does the subjective value of later rewards, which are "discounted" (Mischel & Masters, 1966; Mischel & Metzner 1962). By varying the delay time and reward magnitude, one can calculate the rate at which a future reward is discounted, called the temporal discount function. Steeper discount functions indicate that an individual prefers an immediate reward over a later reward at lower values of the immediate reward and at shorter delay periods. There are both individual and age differences in discount functions across childhood, adolescence, and early adulthood (e.g., Scheres et al., 2006; Steinberg et al., 2009). Even young children have been shown to be sensitive to delay intervals and its effects on perceived attractiveness of rewards (e.g., Mischel, Grusec, & Master, 1969; Mischel & Metzner, 1962; Schwarz, Schrager, & Lyons, 1983). Children also respond differently to the types of rewards used (e.g., Bonato & Boland; 1983; Johnson, Parry, & Drabman, 1978; Koriat & Nisan, 1978), and to different reward magnitudes (e.g., Herzberger & Dweck, 1978; Koriat & Nisan, 1978; Lemmon & Moore, 2007).

More distal factors such as sociocultural norms and values may exert a less obvious influence on delay of gratification choices (e.g., Carlson & Zelazo, 2011). However, beliefs about the importance of saving or not saving, the value of self-control, or general trustworthiness of people, can have important implications. Mischel (1958) described research in Trinidad, showing that the delay behavior of children from different sub-cultural groups was consistent with sub-cultural attitudes towards delay. Gallimore, Weiss, and Finney (1974) also reviewed findings on Hawaiian-American children, whose seemingly poor delay behavior could be interpreted as adaptive when understood within the cultural importance of immediate consumption during social gatherings for building relationships, affiliation, and social credit.

Goals or values about delaying can also be primed experimentally. In a study by Kesek, Cunningham, Packer, and Zelazo (2011), children listened to a story that emphasized either maximizing rewards or receiving rewards immediately, and subsequently children in the maximizing rewards condition delayed significantly more than children in the immediate rewards condition. Use of more explicit modeling and verbal persuasion and reasoning can similarly influence delay behavior and attitudes (e.g., Bandura & Mischel, 1965; Nisan & Koriat, 1984; Staub, 1972). For example, Bandura and Mischel (1965) exposed delayers and non-delayers to live adult models who exhibited the opposite behavior, and subsequent retesting showed changes in delay behavior in the direction that was modeled.

Several recent studies have focused on expectations for receiving rewards. In Kidd et al. (2012), children completed an art activity during which some promised desirable art supplies were either delivered or not delivered by the experimenter. Children who experienced unreliability with the experimenter waited significantly less time during a subsequent delay of gratification task than children who experienced reliability. This study replicated earlier findings from Mahrer (1956), when an experimenter promised children a balloon the following day on 4 consecutive days, and either consistently, inconsistently or never delivered the balloons. When children were then given a choice between an immediate or delayed reward by the same experimenter, those who received the balloons consistently delayed more often than those who did not. When a different experimenter offered the choice, there was no difference in the delay choices between the groups.

Mistrust can also arise independent of direct recent experience with agents in a delay context. Strickland (1972) found that African-American children were less likely to delay with a White experimenter than an African-American experimenter, possibly from general mistrust of White individuals. More recently, Michaelson et al. (2013) reported a similar effect in adults, who were less likely to delay when delay choices were simply paired with faces and character vignettes that were rated low in trustworthiness.

Expectancy about rewards can be inferred not only from beliefs about the reliability or trustworthiness of the agent delivering the rewards, but also from non-social cues, such as statistical information in the environment which give cues to the probabilistic distribution of delay durations (McGuire & Kable, 2012, 2013). In a situation where the probability of delay duration is characterized by a Gaussian or normal distribution (see McGuire & Kable, 2013), the longer you wait, the closer you expect to be to the reward. For example, after sitting for a long period watching a play, you expect that the end should be close. In other situations characterized by heavy-tailed distributions, where rewards either come very quickly or are indefinitely long, the longer you wait, the longer you expect to wait. For example, when on hold for a call with no answer in the first 5 minutes, you expect that the wait could be very long. When there is uncertainty about when a future reward will arrive, as is the case in Mischel's classic paradigm, continued persistence in waiting may not always be appropriate (e.g., Dasgupta & Maskin, 2005; Fawcett et al., 2012; Rachlin, 2000). In theses situations, we should be guided by our temporal beliefs about the environment. In an experimental demonstration by McGuire and Kable (2012), adults completed a computer task consisting of trials offering choices between taking an immediate small reward and waiting for an unknown amount of time for a larger reward. Those exposed to conditions with delays following a heavy-tailed distribution showed less persistence over time than those exposed to delays following a Gaussian distribution, suggesting that people calibrate their persistence in an adaptive manner to the environment and past experience.

Taken together, these research studies provide collective evidence that delay of gratification and its associated paradigms are sensitive to contextual factors and are likely measuring more than self-control. The end goal is not always to delay. Sometimes not delaying is the most adaptive choice and is guided by rational considerations. Nonetheless, many of the above studies have relied primarily on showing group differences rather than individual change. Specifically, they do not address individual differences in the ability to respond adaptively to different contexts.

Without information about both children's baseline behavior and later posttest behavior, we cannot clearly identify which individuals did or did not modify their behavior in the new context (e.g., social mistrust), nor differentiate between those who were induced to shift their behavior (e.g., delayers choosing to not delay) and those who had no need to shift their behavior (e.g., non-delayers). Furthermore, if we know children's baseline responses, we can tailor manipulations to the individual, and assess change in delay behavior in both directions. That is, when given convincing reasons to do so, can delayers take immediate rewards and can non-delayers wait for later rewards?

With the exception of a few (e.g., Bandura & Mischel, 1965; Nisan & Koriat, 1984), previous studies have mostly ignored baseline preferences.

The Role of Executive Function

When a situation calls for behavior that conflicts with one's desires, tendencies, or predispositions, whether delaying when one typically prefers immediate gratification, or acting on impulse when one typically would refrain, self-control processes may play an important role. Self-control draws on executive functions (EF), neurocognitive processes involved in the conscious control of emotions, thoughts, and behaviors, including working memory, inhibitory control and set-shifting (Carlson, Zelazo, & Faja, 2013; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000; Zelazo & Müller, 2002). Emerging in the first year of life, EF undergoes significant development over the preschool years (Zelazo & Müller, 2002) as reflected by age-related improvements on a wide range of EF tasks administered to young children (Carlson, 2005). Neurological evidence regarding the prefrontal cortex, which is associated with EF, suggests a protracted development throughout childhood and into adolescence (e.g., Casey, Giedd, & Thomas, 2000; Giedd et al., 1999; Gogtay et al., 2004; Huttenlocher & Dabholkar, 1997) that parallels developmental changes at the behavioral and cognitive level.

In particular, set-shifting (i.e., switching between different mental sets, functions, or rules) may be associated with one's ability to adjust behavior across situations. In preschoolers, a commonly used measure of EF is the Dimensional Change Card Sort (DCCS; Frye, Zelazo, & Palfai, 1995; Zelazo, 2006), a task that involves sorting bivalent cards (e.g., red rabbit, blue boat) with target cards (e.g., blue rabbit, red boat), by one set of rules (e.g., color), and then switching and sorting by an incompatible set of rules (e.g., shape). Younger children often perseverate and continue sorting by the first set of rules, despite knowing the new rules. Flexibly switching to the second set of rules improves with age and may involve an increased ability to reflect on and hierarchically integrate higher order rules (e.g., if it is the color game, follow these rules; but if it is the shape game, follow these rules; Bunge & Zelazo, 2006; Zelazo & Frye, 1988). As EF matures, it can also facilitate greater attention to relevant contextual information, leading to reappraisal of stimuli and events, and reevaluating different response options (Cunningham & Zelazo, 2007; Zelazo & Cunningham, 2007), which are important processes in adaptive behavior. The same processes that are involved in reappraisal and switching between using incompatible rules might also enable children to flexibly switch behaviors across environmental contexts (e.g., if context A, then delay; if context B, then do not delay).

The Role of Emotion and Motivation: Hot Tasks

Although our primary focus is on delay of gratification, there are other behaviors that are characterized by the same context-dependency as delay of gratification, whereby the adaptiveness of a behavioral response is determined by the context or environment. For example, taking big risks in gambling is generally maladaptive, but it may be a wise strategy if substantial needs must be met in very little time. Similarly, although persistence is generally admired, persistence in an activity that no longer yields any reward or pleasure can be seen as stubborn or perseverative. As with delay of gratification, the same difficulties in interpretation arise when individuals exhibit risky behaviors or lack of persistence. They may be showing poor decision-making and poor control, or alternatively, they may be demonstrating rational adaptive responses.

Tasks that involve affective and motivationally salient features, especially in the form of rewards and punishments, have been called “hot” tasks in the EF literature. They are distinguished from “cool” tasks characterized by decontextualized problems that are well-defined, abstract, and symbolic in nature, such as the DCCS. Cool tasks often have clear correct and incorrect answers that do not change with situational factors, and largely involve maximizing control. Some tasks considered cool include the DCCS, tapping a wooden peg once when the experimenter taps twice and vice versa, a measure of inhibitory control (Diamond & Taylor, 1996), and remembering and repeating a series of words or numbers in reverse order, a of measure working memory (Davis & Pratt, 1996). In contrast, hot tasks are often under the influence of multiple factors that affect how one should respond. Some tasks considered hot include delay of gratification, resisting temptation to peak at a gift (e.g., Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996; Murray & Kochanska, 2002), choosing between risky options with both gains and losses in gambling tasks (e.g., Carlson, Zayas, & Guthormsen, 2009; Crone & van der Molen, 2004; Kerr & Zelazo, 2004), responding to stimuli with changing reinforcement contingencies in object reversal tasks (e.g., Overman, Bachevalier, Schumann, & Ryan, 1996), and discontinuing a no-longer rewarded response in extinction tasks (e.g., Gladstone, 1969; Happaney & Zelazo, 2004).

In these hot tasks, rewarding or punishing stimuli and situations are often critical components of the task, and the motivation, emotion, and meaning elicited by these stimuli directly impact performance. There is likely individual variation in how children interpret and respond to these stimuli and events, which makes the task potentially non-uniform across children. These tasks are also characterized by choices and response options that are often not clearly defined as correct or incorrect. For example, in an extinction task, children may be told to continue responding to a stimulus and stop whenever they wish, but there is no clear expectation of when to do so (Happaney & Zelazo, 2004). In the Less is More task (Carlson, Davis, & Leach, 2005), children are shown a series of small and large piles of treats, and the pile they point to is given to a puppet and children receive the other pile; however, the experimenter does not suggest a right or wrong answer. This is in contrast to cool tasks that have clear goals, answers, and expectations. It is this openness in hot tasks that allows children to bring in other influences to bear, such as cultural beliefs and values. In an extinction task, children may continue to respond to a no-longer rewarded stimulus because they perceive it as important to the experimenter and value compliance. In the Less is More task, children may be socialized to value sharing and giving to others, and thus point to the larger pile in order to share more treats with the puppet, rather than receive more treats for themselves.

Perhaps a reflection of their sensitivity to context, hot tasks are more heterogeneous than cool tasks, and different paradigms do not consistently correlate or load on the same factor (e.g., Dalen, Sonuga-Barke, Hall, & Remington, 2004; Garon & Moore, 2007; Hongwanishkul, Happaney, Lee, & Zelazo, 2005; Olson, Schilling, & Bates, 1999; Prencipe et al., 2011; Sonuga-Barke, Dalen, & Remington, 2003). Despite this lack of coherence, hot tasks seem to differ from cool tasks in a number of ways. Hot and cool tasks often do not correlate and load on different factors (e.g., Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002; Murray & Kochanska, 2002; Olson, 1989; Olson et al., 1999; Smith-Donald, Raver, Hayes, & Richardson, 2007). Unlike cool tasks, hot tasks are often unrelated to IQ (e.g., Dalen et al., 2004; Hongwaniskul, et al., 2005; Sonuga-Barke et al., 2003; Toplak, Jain, & Tannock, 2005), and inconsistently related to age (e.g., Beck, Schaefer, Pang, & Carlson, 2011; Hongwanishkul et al., 2005; Schwarz et al., 1983; Thompson, Barresi, & Moore, 1997).

They also differentially predict developmental outcomes. Whereas cool tasks strongly predict academic outcomes (e.g., Blair & Razza, 2007; Brock et al., 2009; Raver et al., 2011; Willoughby, Kupersmidt, Voegler-Lee, & Bryant, 2011), there are mixed findings regarding the relation between delay of gratification and early academic skills (e.g., Duckworth & Seligman, 2006; Smith-Donald et al., 2007; Thorell, 2007). Delay of gratification often predicts peer functioning and competence, while sometimes, cool tasks in the same studies can show little or no relations (e.g., Olson, 1989; Olson & Hoza, 1993).

These differences between hot and cool tasks, as well as the heterogeneous nature of hot tasks, may partially be related to the complexity of hot tasks and their ability to be influenced by the context. It is important to note, however, that purely hot or cool tasks do not exist and differences are likely a matter degree, with some tasks primarily emphasizing cool aspects and some emphasizing hot aspects. Cool tasks can also be influenced by contextual factors, but to a lesser degree.

The Role of Temperament

Above we discussed contextual factors that may influence children's general preferences for delay, such as broad mistrust of others or cultural beliefs in the value of saving. Other factors may influence children's situational preferences for delay, such as the type of reward or experience with the agent. However, given the emotional and motivational aspects of delay behavior, as well as those measured in other hot tasks, another important determinant of children's behavior and preferences is their temperament. Without invoking rules of rationality and reason, children are predisposed behaviorally and physiologically to certain styles of responding, characterized by different levels of approach behavior toward rewards and different levels of fear and withdrawal. For example, a child who is impulsive with high approach tendencies will find it difficult to resist an appetitive immediate reward, but a child who is inhibited with low approach will find it less difficult. The children can be behaviorally indistinguishable in demonstrating delay behavior, but in the former case, exertion of self-control is required, whereas in the latter case, self-control is less relevant in the absence of a strong urge to be inhibited. Thus, children's temperamental style can physiologically bias their delay behavior, making it more or less difficult or desirable.

Temperament is understood broadly as the individual differences in emotional, motor, and attentional reactivity, as well as the individual differences in effortful control, which modulates that reactivity (Rothbart, 2007). These systems develop early in life, with regulation initially more reactive, and becoming more voluntary as effortful control develops with age. Individual differences have both physiological and neurological underpinnings, emerges early in infancy, and remains moderately stable across later childhood and adolescence (Derryberry & Rothbart, 1997).

The reactive motivational system, rooted in limbic circuits, is comprised of an appetitive-approach system, and a fear-avoidance system. The appetitive-approach system promotes approach behavior towards rewarding and positive stimuli and events, and is governed through circuits that include the basolateral amygdala, the ventral tegmental area, and the nucleus accumbens (Derryberry & Rothbart, 1997). The fear-avoidance system promotes inhibited motor activity, increased arousal, and heightened attention, and is governed through hippocampal and amygdala circuits that respond to novel signals, biologically prepared fear signals, and punishment (Derryberry & Rothbart, 1997). It also functions to inhibit and regulate the appetitive-approach system, although in a reflexive rather than voluntary manner. These reactive systems are in turn regulated by the anterior attentional system, which is involved in the control and flexibility of attention to spatial and semantic information, and has circuitry centered on the anterior cingulate region (Derryberry & Rothbart, 1997). Known as effortful control, the system is defined as the voluntary and flexible control of behavior through attentional and inhibitory mechanisms that serve to inhibit a dominant response and activate a subdominant response (Rothbart, Ahadi, & Evans, 2000; Rothbart & Bates, 1998). Although a dimension of temperament, it is considered closely related to EF and may capture stable individual differences in EF (e.g., Duckworth & Carlson, 2013; Zhou, Chen, & Main, 2012). In fact, many preschool effortful control tasks (e.g., Kochanska, Murray, & Coy, 1997; Kochanska et al., 1996) resemble those used to measure EF (e.g., Carlson, 2005), and relations between performance on EF tasks and parent ratings of effortful control have been reported (e.g., Chang & Burns, 2005; Rothbart, Ellis, Rueda, & Posner, 2003).

Effortful control modulates both the approach and fear systems as dictated by environmental demands. Having very high approach or very high fear/withdrawal can be problematic, especially when it cannot be adequately modulated according to the context. Several investigators such as Fox (1994) and Nigg (2000) suggest that undercontrol can lead to externalizing problems (e.g., attention deficit and hyperactivity disorder, oppositional behavior) and overcontrol can lead to internalizing problems (e.g., depression and anxiety). This is consistent with Kagan's work with behaviorally inhibited children who display fear and withdrawal in new unfamiliar situations, leading to restriction of exploration, avoidance of novelty, and high reactivity and distress (Kagan, Reznick, & Gibbons, 1989). Block and Block (1980) describe how overcontrol and inhibition, as well as undercontrol and impulsivity can be maladaptive; however, high ego-resilience can allow individuals to modulate their level of ego-control to flexibly adjust their behavior to varying situational demands.

Consistent with the idea that the extremes of the spectrum at either end can be maladaptive, in one study (Murray & Kochanska, 2002), preschoolers who scored high on effortful control tasks had greater internalizing symptoms, whereas those who scored low on effortful control had greater externalizing symptom. Those in the middle showed the most adaptive functioning. Similarly, in 4- to 6-year-olds, emotional regulation was best for children who exhibited intermediate performance on a battery of mostly hot inhibitory tasks (Carlson & Wang, 2007). Although it may appear that children with very high or low effortful control are at greater risk for problems, both reactive control and effortful control are confounded in these tasks. It is more likely that high approach and high fear/withdrawal may be driving these findings, although poor effortful control may also contribute to problematic behaviors.

Temperamental differences related to approach and fear/withdrawal may influence delay behavior, and accordingly, effortful control may play an important role in children's ability to adjust their delay behavior. In particular, when the reactive systems of approach and fear/withdrawal are engaged and activated, effortful control may be necessary to regulate behavior towards adaptive responses that run counter to those promoted by these reactive systems.

The Current Studies

We designed a novel dependent variable to evaluate flexibility in delay of gratification behavior. Beyond simply measuring delay behavior, we were interested in whether preschool-aged children would flexibly adjust their delay behavior when placed in a context where the most adaptive response is the behavior opposite to their baseline preferences. Using a within-subjects design, we first assessed children's baseline behavior and then assessed children's behavior a second time under new conditions that should encourage them to change their behavior according to what is adaptive or rational. We expected variability in children's responses with regard to switching, that those who adaptively altered their behavior would show greater EF on an independent measure. Furthermore, we investigated whether results could be replicated across different but related behaviors, specifically delay discounting and saving, which are both considered to be hot. This would provide stronger support that findings related to flexibility are not specific to a particular paradigm, and can be applied to a family of decision-making behaviors.

In Study 1, children were offered a series of choices between less treats immediately and more treats later, thus measuring their delay decisions (e.g., Mischel, 1966; Schwarz et al., 1983; Thompson et al., 1997). This is representative of a typical delay choice paradigm used in the literature and also displays characteristics that suggest, as with other delay tasks, it may be sensitive to situational factors and individual differences. Performance is not strongly or consistently related to IQ and other EF measures (e.g., Beck et al., 2011; Hongwanishkul et al., 2005; Schwarz et al., 1983). Reports regarding age-related differences have also been mixed, with age differences found in older children (e.g., Melikian, 1959; Mischel, 1958; Mischel & Metzner, 1962; Shipe & Lazare, 1969), but often not observed during the preschool period (e.g., Beck et al., 2011; Lindstrom & Shipman, 1972; Montgomery, 1976; Schwarz et al., 1983; Toner, Holstein, & Hetherington, 1977). However, there have been a few exceptions from recent studies (e.g., Hongwanishkul et al., 2005; Thompson et al., 1997).

In Study 2, children were offered a chance to play with a small marbles race game followed by a big marbles race game, but were only given 3 marbles at the beginning, requiring them to save marbles from the smaller game to play with the bigger game, thus assessing their decision to save for the future. This saving paradigm was developed by Metcalf and Atance (2011), who interestingly reported no age differences between 3-, 4-, and 5-year-olds, which suggests that, like delay choice, individual differences may be a significant determinant of baseline saving behavior. Furthermore, in their study, children were administered the paradigm twice, and a subset of children who initially did not save any marbles, saved one or more marbles the second time. This indicates that children's saving behavior can be influenced by experience, which makes it particularly suitable for our manipulations and interest in observing changes in saving behavior.

Wärneryd (1999) describes saving as the postponing of some consumption for future enjoyment. In discussions of delay of gratification behavior and self-control, saving is often not clearly distinguished from delaying, and sometimes the ability to delay gratification is seen as a precursor to saving (e.g., Ainslie, 1975; Wärneryd, 1999). Delay of gratification and saving are similar in that they both require inhibiting impulses towards immediate rewards and both unfold across time. Nonetheless, some argue that while closely related, saving and delay of gratification retain some important differences, at least as measured by standard tasks (Sonuga-Barke & Webley, 1993). Delay of gratification tests children's waiting ability, particularly in Mischel's paradigm, and waiting is the cost to attain a larger reward. The time is wasted and is not necessarily productive towards the end goal. Waiting is not equal to saving. While there is an inherent passage of time during saving, it is not a cost, but a necessity to provide opportunities to accumulate rewards across time. Unlike delay of gratification, saving involves the integration of multiple events during which immediate consumption must be given up repeatedly in service of a future goal.

The delay choice paradigm also involves multiple opportunities to delay, as in saving. Metcalf and Atance (2011), however, argue that there are still important differences. In the delay choice task, children do not give up the immediate reward, but simply need to wait to gain the reward plus more. In contrast, during the marbles saving task, children who save give up an opportunity for current enjoyment that is not recoverable. Saving may also involve more future-oriented thinking and considerations (Metcalf & Atance, 2011). Children need to consider future needs, prepare for unexpected future circumstances, and understand that present spending behaviors have future consequences (Otto, Schots, Westerman, & Webley, 2006).

We expected that individual differences in delay and saving flexibility would be related to various factors. First, we predicted that switching would be related to age, with older children showing more flexibility than younger children due to increasing EF skills that are associated with age (Carlson et al., 2013). Second, we predicted asymmetry in the direction of switching. The less taxing response is to be impulsive and take immediate rewards, and thus it may be more difficult for non-delayers and spenders to wait and save than for delayers and savers to choose immediate rewards or spend their resources. Third, we predicted that the ability to switch would be related to performance on a set-shifting task, if the same EF processes that underlie DCCS performance also underlie switching. In contrast, baseline delaying and saving was expected to be unrelated to, or not strongly related to age, consistent with previous findings, and also unrelated to set-shifting (e.g., Hongwanishkul et al., 2005).

We were also interested in exploring how individual differences in temperament may affect children's delay and saving behavior, given the relevance of the reactive motivation system in these reward paradigms. A widely-used parent-report questionnaire of children's temperament, the Children's Behavior Questionnaire (e.g., Putnam & Rothbart, 2006; Rothbart, Ahadi, Hershey, & Fisher, 2001) was included, from which three factors could be derived that represent the appetitive-approach (Surgency), fear-avoidance (Negative Affect), and effortful control (Effortful Control) systems respectively. Given the affective nature of the tasks, we predicted that greater approach tendencies would be related to less delay and saving behavior. If the fear-avoidance system inhibits the approach system, higher fear and avoidance tendencies would be related to greater delay and saving. Given the role of effortful control in regulating these temperamental systems, and its similarity and overlap with EF, it was predicted that effortful control, like set-shifting, would be associated with children's ability to adjust their delay and saving behavior. This may be particularly relevant when behavior must be shifted in a direction that is contrary to children's temperamental predisposition (e.g., non-delayer with high approach switching to delaying).

Study 1

Children's ability to be flexible with their delay behavior was examined using the delay choice paradigm with food treats. Children's baseline preference for delaying was assessed and children were then administered the delay choice task a second time with modifications to the delivery of treats such that their preference at baseline (i.e., to delay or not delay) was associated with a risk of losing treats. This should create incentive for children to shift their choices away from baseline choices that have become risky, and towards less preferred choices that are not risky.

Method

Participants. Children were recruited from a large Midwestern city through a university database of families that had expressed interest in research. Screening criteria included full-term birth and absence of any serious medical conditions and developmental disorders. A total of 41 children were excluded because they showed no clear preference to delay or not delay (see below), a procedure also used by Bandura and Mischel (1965) when they tailored their intervention to delayers and non-delayers. Excluded were 13 3.5-year-olds (10 females) and 28 4.5-year-olds (18 females) respectively. Significantly more 4.5-year-olds, $\chi^2(1, N = 181) = 7.75$, $p = .004$, and more females, $\chi^2(1, N = 181) = 6.88$ $p = .007$), were excluded**.** Ten children failed to start or complete the tasks due to fatigue, lack of interest, or noncompliance, and were also excluded from analyses.

The final sample consisted of 140 children, including 79 3.5-year-olds ($M = 42.42$ months, $SD = .68$, 35 females), and 61 4.5-year-olds ($M = 54.51$ months, $SD = .77$, 28 females). The majority of children were White and non-Hispanic (84.3%), with the remaining sample consisting of children of White and Hispanic, African-American, Asian, Native Hawaiian/Pacific Islander, and mixed ethnicity. Families were predominantly two caregiver households (96.4%) and the majority of primary caregivers, mostly mothers, reported attaining a bachelors degree or higher (84.3%). The median and mode household income was $75 000 to $99 999.

Procedure and Measures. Families were invited to the lab for a one-hour visit where children were assessed individually in a quiet child-friendly laboratory room during a videotaped 40–45 minute session. Parents completed a general demographics questionnaire (e.g., family living situation and structure, parent education and income, child's ethnicity). Children first completed the standard delay choice task (9 trials) to determine their baseline performance. They were categorized as delayers if they delayed 6 to 9 trials, and non-delayers if they delayed 0 to 3 trials. Children who delayed 4 or 5 trials, showing no clear preference, were excluded from analyses.

After completing a standardized receptive vocabulary test, children were randomly assigned to either an experimental or a control-risk condition and administered a modified version of the delay choice task that introduced a risk of losing treats. In the experimental condition, children were encouraged to choose their non-preferred choice (i.e., delayers to not delay, non-delayers to delay) by associating risk with their preferred choice. A total of 77 children were assigned to the experimental condition, including 44 3.5-year-olds (20 females) and 33 4.5-year-olds (17 females). In the control-risk condition, children were encouraged to choose their preferred choice (i.e., delayers to delay, non-delayers to not delay) by associating risk with their non-preferred choice. This control-risk condition assessed the extent of spontaneous switching when there was no adaptive reason to do so. In fact, switching could be considered maladaptive since the less preferred choice was associated with risk. Switching may be observed, nonetheless, if children find risk or the stimuli of the risky option attractive. A total of 33 children were assigned to this control-risk condition, including 17 3.5-year-olds (7 females) and 16 4.5-year-olds (6 females).

An additional control-no-risk group was later added to confirm and replicate previous findings (Beck et al., 2011) of test-retest reliability of the delay choice task within a session. Children in this second control-no-risk condition were administered the delay choice task a second time, but with no risk involved. The 30 children in this group included 17 3.5-years-olds (5 females) and 13 4.5-year-olds (3 females).

After the second delay choice task, children were again categorized as delayers (5 to 9 trials) or non-delayers (0 to 4 trials). Children who delayed 4 or 5 trials were not excluded here since the number of trials delayed did not reflect children's natural preference without intervention, but was influenced by our experimental manipulation. If children's classification changed across the two tasks (e.g., delayer to non-delayer), they were considered to be switchers. If children's classification did not change across the two tasks (e.g., remained a delayer), they were considered to be non-switchers. All children then completed an independent measure of set-shifting ability adapted from the DCCS.

Delay Choice Baseline. This task was adapted from Thompson and colleagues (1997) and Prencipe and Zelazo (2005). Children were given a choice of several types of treats to use in the game to ensure that the treat was desirable (e.g., small crackers, sugary cereal, raisins). Then they were offered a series of choices between having one treat now, which was placed in a bowl and could be eaten immediately, and more treats later, which were placed in an opaque envelope to be retrieved at the end of the session. The side of presentation for the treats was counterbalanced. There were 9 trials in total, balanced across three types of contingencies (i.e., 1 versus 2, 1 versus 4, 1 versus 6) presented in a fixed mixed order. The experimenter first demonstrated both choices for herself, choosing now and choosing later. For each test trial, children were presented with the treats and asked to make a choice (i.e., "Do you want to choose 1 now or 4 later?").

Delay Choice Modified. In the modified task, the same treats, number of trials, order, and contingencies were used. Two panels with clear plastic tubes were introduced as part of the game and placed side by side in front of the child (see Figure 1). Both panels had a single vertical tube where treats were deposited. Halfway down the panel, the single tube split into two separate tubes. Hidden from the children was a tab at the back that controlled which tube the treats went down. In one panel, two bowls were placed under the two lower tubes to catch the treats. In this case, there was a 100% chance of receiving the treats. In the other panel, a bowl was placed under only one tube and the other tube led to an opaque sealed box. In this case, there was a chance that the treats would get stuck in the box and not be delivered.

Each choice (i.e., now or later) was associated with a different panel. In the experimental condition, the risky panel was associated with the later choice for delayers and with the now choice for non-delayers. In the control-risk condition, the risky panel was associated with the now choice for delayers and the later choice for non-delayers. Children in the control-no-risk condition had two non-risky panels. The experimenter demonstrated how both panels worked. For example, the experimenter showed delayers in the experimental condition how treats for "now" that fell into either bowl under the non-risky panel could be eaten, and how treats for "later" that fell into the bowl under the risky panel would be placed in the envelope. However, treats for "later" that fell into the box under the risky panel were lost and could not be retrieved. The experimenter demonstrated both choices for herself, choosing now and choosing later. Rule checks were administered, assessing children's knowledge of which choices were associated with which panels, as well as what happened to the treats with each bowl and box. Children were questioned up to two times with incorrect responses corrected. During test trials, the risky panel was fixed such that treats always fell into the box and were always lost. To mirror the behavior of the risky panel, the non-risky panel was also fixed such that the treats consistently fell into only one bowl on the right. At the end of the task, children were asked to identify which panel they liked the best.

Peabody Picture Vocabulary Test (PPVT-4). In this standardized test of receptive vocabulary (Dunn & Dunn, 2007), children were shown a series of pages with pictures and asked to point to 1 of 4 pictures on each page, corresponding to a word spoken by the experimenter. Standard administration was used.

Executive Function Scale (EF Scale™). This task developed by Carlson (2012) and was adapted and expanded from the DCCS (Zelazo, 2006) for a wider age range with multiple levels of complexity. Children were asked to sort cards into two opaque boxes affixed with target cards. There were seven levels of complexity with the start level dependent on the child's age (e.g., 3.5-year-olds started at level 3, 4.5-year-olds started at level 4). Level 4 was comparable to the standard DCCS task, where children were asked to sort red stars and blue trucks into boxes with target cards of a blue star and a red truck, first by color (pre-switch), and then by shape (post-switch). Level 3 involved different sorting and target cards where the dimension of color and shape were separated (i.e., a black star on red background, a black truck on blue background). If children passed the initial level, passing both pre-switch and post-switch, more difficult levels were administered until children failed. If children failed the initial level, lower levels were administered in a backward fashion until children passed. In this manner, the highest level at which children could perform was determined. In the now standardized protocol, there are 5 trials in both pre-switch and post-switch, and in order to pass a level, children must correctly sort at least 4 out of 5 trials in both pre-switch and post-switch. In the current study, a slightly earlier version of the task was used, with 6 trials in both pre-switch and post-switch, and sorting 5 out of 6 trials was considered passing a level. For further details on each level, please refer to Carlson and Schaefer (2012).

Children's Behavior Questionnaire Very Short Form (CBQ-VSF). Parents were asked to complete the Very Short Form of the CBQ (Putnam & Rothbart, 2006), a parent-report measure of children's temperament. Parents were asked to rate 36 items, indicating how true each item was of their children in the last 6 months on a 7-point Likert scale. The items described different behaviors children may show in various situations (e.g., "Is afraid of the dark"). Three dimensions of temperament were derived from the questionnaire, including Surgency, Negative Affect, and Effortful Control. Surgency included items related to impulsivity, activity level, high intensity pleasure, and shyness (negatively loading). Negative Affect included items related to frustration, discomfort, fear, sadness, and soothability (negatively loading). Effortful control included items related to low intensity pleasure, smiling and laughter, inhibitory control, attentional focusing, and perceptual sensitivity.

Results

Descriptive. An alpha level of .05 was used for all statistical analyses unless otherwise noted. All children completed all tasks, except 5 children in the control group were included who did not complete PPVT and EF Scale. No statistical differences were found between the experimental, control-risk, and control-no-risk groups with respect to age, sex, baseline delaying, PPVT, and EF Scale, and therefore the groups were analyzed together for baseline analyses. A total of 51 children (36.4%) were classified as delayers and 89 children (63.6%) as non-delayers. The proportion of non-delayers was significantly higher than expected by chance as indicated by a binomial test, $p = .002$. As shown in Table 1, children's baseline status was marginally related to age, with 44.3% of 4.5-year-olds categorized as delayers compared to 30.5 % of 3.5-year-olds, $\chi^2(1, N = 140) = 2.86$, $p = .065$. More females, 49.2%, were categorized as delayers than males, 26.0%, $\chi^2(1, N = 140) = 8.08$, $p = .004$. No relation was seen with PPVT, $t(136) = -.044$, $p = .965$. The sex difference remained significant after applying a Bonferroni corrected alpha rate of .0167 (0.05/3).

Switching. In the control-risk group ($n = 33$), only 5 children switched (15.2%), and in the control-no-risk group ($n = 30$), only 3 children switched (10.0%). Since the rate of switching in both groups did not differ, $\chi^2(1, N = 63) = .38$ $p = .410$, nor did the groups differ on other variables described above, the two groups were collapsed into a single control group for further analyses ($n = 63$). In the experimental group ($n = 77$), 27 children switched (35.1%), which is significantly higher than the 8 children (12.7%) who spontaneously switched in the control group, $\chi^2(1, N = 140) = 9.24$ $p = .002$. To further confirm that switching was low in the absence of any intervention, the first half of baseline trials (1 to 4) was compared to the second half of baseline trials (6 to 9), with trial 5 omitted. No differences were found in the mean number of trials delayed between the first half ($M = 1.49$, $SD = 1.56$) and the second half ($M = 1.53$, $SD = 1.59$), $t(139) = -.43$, $p = .669$, suggesting that within the span of the task, overall response patterns remained consistent with no spontaneous switching.

Differences between switchers and non-switchers were evident by the first trial of the modified delay task. In the group of baseline delayers, switchers were less likely to delay on the first trial ($M = .44$, $SD = .51$) than non-switchers ($M = 0.88$, $SD = .33$), $t(31) = 2.99$, $p = .006$. In the group of non-delayers, switchers were more likely to delay on the first trial ($M = .36$, $SD = .505$) than non-switchers ($M = .03$, $SD = .17$), $t(42) = -3.31$, $p = .002$. Interestingly, however, the experimental group as a whole, regardless of switching, also showed differences in the number of trials delayed between the first and second half of trials on the modified delay task. Delayers delayed more trials in the first half ($M = 2.30$, $SD = 1.31$) and less trials in the second half ($M = 1.91$, $SD = 1.42$), $t(32) = 2.42$, $p = .021$. Non-delayers delayed less trials in the first half ($M = 0.89$, $SD = 1.18$) and more trials in the second half ($M = 1.27$, $SD = 1.44$), $t(43) = -1.95$, $p = .058$. When groups were further broken down into switchers and non-switchers, the same patterns of delaying remained. However, likely due to small cell sizes, these differences were not significant, except for the group of delayers who switched. They delayed more trials in the first half

(*M* = 1.31, *SD* = 1.01) and less trials in the second (*M* = .75, *SD* = .775), $t(15) = 2.33$, p = .034.

In the experimental group, contrary to predictions, switching was not related to age, as shown in Table 2, $\chi^2(1, N = 77) = 3.50$, $p = .365$. It was also unrelated to sex, $\chi^2(1, N = 77) = .15$, $p = .114$ and PPVT, $t(75) = 1.32$, $p = .189$. As predicted, it was related to baseline behavior, with 48.0% of delayers switching and only 25.0% of non-delayers switching, $\chi^2(1, N = 77) = 4.57$, $p = .029$ (see Table 2), although after applying a Bonferroni corrected alpha rate of .0125 (.05/4), it was no longer significant.

To examine whether children recognized which panel was the better choice, which may have facilitated switching, children's self-reported preferences regarding the two panels were analyzed. Preference data were available for 72 children in the experimental group. Of the 25 children who switched, all but 2 (92.0%) identified the non-risky panel as their favorite, consistent with their behavioral responses. In contrast, of the 47 children who did not switch, 26 (55.3%) favored the risky panel and 21 (44.7%) favored the non-risky panel, significantly different from the pattern seen with switchers, $\chi^2(1, N = 72) = 15.38$ $p = .000$. Interestingly, in the group of non-switchers, almost half of the children favored the non-risky panel, although it was inconsistent with their behavior.

EF Scale. One of the main hypotheses was that children's ability to switch would be positively related to set-shifting, as indexed by the highest level passed on the EF Scale (0 to 7). An analysis of covariance was conducted with the highest level passed as the dependent variable, and switching and baseline delaying as independent variables, controlling for age and verbal ability. Both age and verbal ability have been shown to relate positively to the EF Scale (Beck et al., 2011). Results revealed that age was significantly related to EF Scale performance, $F(1, 71) = 39.88$, $p = .000$, partial $\eta^2 = .36$ as was PPVT standard scores, $F(1, 71) = 9.66$, $p = .003$, partial $\eta^2 = .12$. Controlling for both age and PPVT, switching was a significant predictor of EF Scale performance, $F(1, 71) = 5.09$, $p = .027$, partial $\eta^2 = .07$. Adjusted means controlling for age and PPVT showed that switchers had a mean score of 3.78 (*SE* = .20) and non-switchers had a mean score of 3.21 (*SE* = .15). As hypothesized, baseline delaying was not a significant predictor, $F(1, 71) = .84$, $p = .362$, and the interaction between switching and baseline delaying was also not significant, $F(1, 71) = 1.14$, $p = .290$.

CBQ. Two-tailed Pearson correlation analyses were conducted and are presented in Table 3. Contrary to predictions, baseline delaying was unrelated to Surgency, and also unrelated to Negative Affect, and Effortful Control. Switching was marginally correlated with Negative Affect, with those scoring higher on Negative Affect less likely to switch, $r = -.21$, $p = .06$. Switching was not correlated with Effortful Control, although Effortful control was marginally related to EF Scale $r = .15$, $p = .08$. Given results indicating that it was easier for delayers to switch compared to non-delayers, and that switching may be slightly different in these groups, separate correlational analyses were conducted for delayers and non-delayers. In delayers, switching was again marginally related to Negative Affect, $r = -.30$, $p = .08$, but unrelated to Effortful Control. In non-delayers, switching was unrelated to Negative Affect, but marginally related to Effortful Control, with those showing higher effortful control more likely to switch, $r = -.28$, $p = .07$. In this group, however, Effortful Control was correlated with EF Scale performance, $r = .33$, $p = .03$, and if this relation was controlled, then Effortful Control was no longer related to switching.

Discussion

We were interested in whether children would be able to modify their behavior in a delay of gratification choice task when the context changed and their preferred response became risky and no longer adaptive. Indeed results showed that children were able to adapt their behaviors, but children varied in their responses, with some children switching and others not switching. Less than half of the children changed their behavior, but a greater percentage of switching was observed in the direction of delaying to non-delaying. As mentioned earlier, it may indicate that shifting towards the consumption of rewards is easier and requires less control than shifting towards waiting. Furthermore, switchers could be differentiated from non-switchers by the first trial, suggesting that presenting new information and contingencies can be enough to cue a shift in children's behavior. However, as results also indicated that for all children, both switchers and non-switchers, the latter half of trials showed a greater shift towards the non-risky choice than the first half of trials, it suggests that negative feedback from losing their treats may facilitate switching behavior over time. Nonetheless, despite being aware of the new risk context, and being somewhat responsive to it, some children may not fully adapt their behavior. Interestingly, a subset of children who were unable to demonstrate full switching behavior still stated a preference for the non-risky option. Some children may have understood or recognized the better choice, and may have even marginally shifted their behavior, but did not fully follow through.

It was hypothesized that switching would be associated with age and set-shifting skills. As predicted, children who were able to adjust their responses on the delay paradigm scored higher on the EF Scale, controlling for age and verbal ability. This suggests that EF skills, as measured independently in a cool abstract task, may be recruited in situations of a hot nature, where children must adjust their behavior to the environment in an adaptive manner. As discussed previously, performance in hot tasks are not consistently related to behavior in cool tasks, and in particular, delay of gratification is typically not related to measures of set-shifting. Surprisingly, no age differences in switching were observed, although age itself was strongly related to set-shifting, and switching was independently related to set-shifting. It is possible that a wider age range may be needed to reveal developmental differences.

There were also no significant age differences in baseline delaying, although there was a trend towards older children delaying more often, consistent with mixed findings in the literature. Females were more likely to delay at baseline than males, which has been previously reported in the literature (see Silverman, 2003). Also consistent with past findings, the delay choice task was not related to verbal ability or set-shifting (e.g., Beck et al., 2011; Hongwanishkul et al., 2005; Schwarz et al., 1983).

With regards to children's temperament, it was predicted that baseline delaying would show relations to surgency and negative affect, but this was not observed. It is possible that given the complex nature of delay, these reactive aspects of temperament may not be as prominent in light of other influences on behavior. Interestingly, switching was negatively associated with negative affect. Children who perseverated on their baseline response during the modified task often displayed high frustration and negative affect at continually losing their treats. Those prone to negative emotionality may experience such reactions more strongly and find it more difficult to disengage from the negative emotion to flexibly shift their behavior.

Associations have also been reported between higher negative affect and lower effortful control (e.g., Eisenberg et al., 1997; Kochanska, Coy, Tjebkes, & Husarek, 1998). Lower effortful control may hinder switching behavior, although no relation between negative affect and effortful control was observed in the current study. Garon and Moore (2006) also reported that preschoolers who scored higher on negative affect did more poorly on a children's gambling task. The task involves two decks of cards, an advantageous deck with lower gains, but also lower losses, and a disadvantageous deck with higher gains, but also higher loses. Although initially the disadvantageous deck may seem attractive, as children continue to select cards, over time the deck results in overall losses. Thus this task also assesses flexibility, as children potentially need to switch their selections from the risky deck to the non-risky over the course of the task.

As predicted, switching was related to effortful control, but this was only observed in non-delayers, and not delayers. Our results suggest that it is more difficult for children to switch from not delaying to delaying, and in such situations, control processes may be more important in flexibly modulating reactive behavior, and show stronger relations. Effortful control was also related to set-shifting, and when controlled, the relation between switching and effortful control was no longer significant. This points to the potential overlap between effortful control and EF and their role in facilitating flexible behavior.

There are, however, several limitations of the current study. By using risk to manipulate the adaptiveness of the context, a potential confound was introduced, where children who continued to select the risky option may have been attracted to risk-taking rather than unable to adapt to the changed context. An attraction to the risky option would also make it more difficult for these children to switch their behavior, in contrast to other children who are more risk-aversive and can be easily induced to shift away from risk. Nonetheless, this concern is mitigated by the fact that the control group did not show any substantial shifting towards the risky option, which would be expected if attraction to risk-taking was a strong influence on behavior. In this experimental paradigm, it was also behaviorally difficult to disentangle persistence and extinction (Happaney & Zelazo, 2004). When children fail to switch, they may be perseverating at a no-longer rewarded response, or they may be demonstrating persistence in reaching a desired goal. Children were led to believe that there was a chance treats could bypass the box, and continuing to choose the risky option could reasonably be an example of persistence. With continued trials and consistent negative feedback, however, expectations of obtaining treats should be diminished. Furthermore, in the current task, children were given an alternative response option where rewards were guaranteed, albeit not in their preferred manner or amount, which reduces the need to persist at a risky response.

In Study 2, these issues were addressed by employing a savings paradigm that did not involve an explicit introduction of risk, but used past experience to signal relevant contingencies and information in the situation. This also provided a more ecological real-world problem that children may encounter, without reliance on a risky situation that is artificially designed. Furthermore, the outcomes of children's choices would not be characterized by uncertainty with regards to whether their responses may be rewarded or not, but would be clear from their past experiences.

Study 2

One of the goals was to examine children's flexibility in at least two different paradigms examining closely related but different decision-making behaviors. If similar findings regarding children's flexibility and their relation to EF can be demonstrated, it would strengthen our confidence that these relations are not specific to a particular paradigm or manipulation but may indicate some broader function. As discussed, saving behavior is similar to delay of gratification in some respects, such as the need to control impulses toward immediate consumption or rewards. Saving differs from delaying, however, in that it may be more future-oriented, and it involves giving up a present consumption, that cannot be recovered, for a future consumption. The marbles saving paradigm (Metcalf & Atance, 2011) was used to measure flexibility in children's saving behavior. It allows a replication of Study 1 using a more realistic situation with which children are familiar, and also eliminates confounds associated with risk-taking and with persistence. Saving behavior as measured by this task furthermore appears to be sensitive to past experience, as reported by Metcalf and Atance (2011), and children's saving behavior can be influenced by the context and experience (Otto, 2003). The delay choice tasks was also included in Study 2 to assess whether baseline saving and baseline delaying would be related. It was expected that given their similarities, but also considering the complexity of these tasks, saving and delaying would be weakly or moderately related.

The marbles saving task involves giving children a limited number of marbles during play with a small marbles game and measuring how many marbles are saved for use with a later big marbles game. Children's baseline preference for saving was assessed during the small marbles game. Children who did not save marbles were unable to play with the big marbles game. A modification to the task was introduced such that children who saved marbles were also unable to play with the big marbles game, which was revealed to be broken. In the former case, the presentation of the big marbles game as promised should emphasize the certainty of future need for marbles, and that no extra marbles are available. In the latter case, the presentation of the broken big marbles game should emphasize the certainty of no future need for marbles. This experience should encourage children to adjust their subsequent behavior when given the chance to play the game a second time.

Assuming similar results on children's flexibility could be replicated in Study 2, we were additionally interested in whether flexibility in these experimental contexts would be reflected in flexibility in children's everyday lives. During the preschool years, children's behavior becomes more ritualistic, repetitive, compulsive, and perseverative, and in fact, can be characterized by increased inflexibility (Evans et al., 1997; Evans, Lewis, & Iobst, 2004). During this time, parents often report that children insist on strict invariable routines, have strong likes and dislikes, demand repetition, display high sensory-perceptual awareness, and are often being bothered by small details and changes. Becoming prevalent by age 2 and decreasing by age 6, Evans et al. (1997) argue that while these behaviors resemble those seen in adults with obsessive-compulsive disorder, they may be a normative part of development, and may involve emotional regulation and anxiety reduction (Evans, Gray & Leckman 1999). To measure these ritualistic and compulsive behaviors in early childhood, Evans and colleagues (1997) developed a parent-report measure, the Childhood Routine Inventory (CRI). We predicted that children who show flexibility in saving behavior would have less reported compulsive behaviors, as measured by the CRI.

Considering our observations that flexibility as measured by our experimental paradigm was positively associated with set-shifting, we also predicted that greater set-shifting would be associated with less compulsive behaviors. Pietrefesa and Evans (2007) reported that CRI measures of compulsive behaviors were negatively associated with set-shifting and response inhibition in 4- to 6-year-olds, which mirror findings of obsessive-compulsive behaviors and poor set-shifting in adults (e.g., Lucey et al., 1997; Zohar, LaBuda, & Moschel-Ravid, 1995). Furthermore, effortful control may also be related to less compulsive behaviors, given its overlapping relationship with EF.

Method

Participants. A new sample of children was recruited from the same population using the same methods and screening criteria described in Study 1. Three 3.5-year-old boys failed to start or complete the tasks due to fatigue, lack of interest, or noncompliance, and were excluded. The final sample consisted of 142 children, including 73 3.5-year-olds (*M* = 42.13 months, *SD* = .48, 36 females), and 69 4.5-year-olds (*M* = 53.91 months, *SD* = .52, 36 females). Similar to Study 1, the majority of children were White and non-Hispanic (88.0%), with the remaining sample consisting of children of White and Hispanic, African-American, Asian, and mixed ethnicity. Families were predominantly two caregiver households (94.4%) and the majority of primary caregivers, mostly mothers, reported attaining a bachelors degree or higher (88.7%). The median and mode household income was $100 000 to $124 999.

Procedure and Measures. Children were tested in the same lab environment as in Study 1, and parents completed the same demographics questionnaire. Children were first administered the standard delay choice task, to replicate findings regarding baseline delaying, and examine its relation to saving, followed by the marbles saving task to determine their baseline preference for saving. Children were shown both a small marbles game and a big marbles game. Children were told they would play first with the small game, and then later with the big game. However, children were given only 3 marbles, and therefore, must save at least 1 of the 3 marbles while playing the small game in order to have marbles to play the big game. Given the reported low to moderate level of saving (Metcalf & Atance, 2011), children who saved at least one marble were categorized as savers, and those who saved no marbles were categorized as spenders. Children were randomly assigned to an experimental or control condition.

In the experimental condition, spenders were given the opportunity to play with the big marbles game, but without marbles, children were unable to play. Savers were given the opportunity to play with the big marbles game, but they discovered that it was broken and not playable. In the control condition, children did not spend time with the big marbles game, but were told the experimenter made a mistake and they would play it later. Thus the control group did not receive any intervening feedback experience between the first and second playing of the small marbles game. A total of 102 children were assigned to the experimental condition, including 53 3.5-year-olds (26 females), and 49 4.5-year-olds (25 females). A total of 40 children were assigned to the control condition, including 20 3.5-year-olds (10 females) and 20 4.5-year-olds (11 females).

All children then completed a receptive vocabulary test after which the experimenter discovered 3 additional marbles. Children were given the chance to play with the small marbles game again, followed by the big marbles game. Using the same criteria above, children were again categorized as savers or spenders based on their saving behavior. Children who changed classifications across the two administrations of the savings task were categorized as switchers (e.g., saver to spender), whereas those who did not change classifications were categorized as non-switchers (e.g., remained saver). At the end of the second saving task, children were allowed to play with a third even bigger marbles game with extra marbles, followed by an independent measure of set-shifting.

Marbles Saving Task. This task was adapted from a paradigm developed by Metcalf and Atance (2011). Children were introduced to a small marbles game with a single run/chute and a more appealing big marbles game with 3 runs/chutes (see Figure 2). The experimenter demonstrated with the small game how once a marble was dropped down the start hole, it rolled into a box where it could not be retrieved or used again ("Once you put a marble down the hole, it rolls all the way down and it goes into the box. It stays inside and you can't use it again"). It was explained that both games worked the same way. Children would first play with the small marbles game for 3 minutes, followed by the big marbles game for 3 minutes. They were shown a transparent bag with marbles and it was emphasized that there were only 3 marbles left ("Look, there are only 3 marbles left. You *only* get 3 marbles to use for all your marble games today. Remember once you put a marble down the hole, you can't use it again"). The big marbles game was removed from the room and then rule checks were administered. Children were asked which game they were going to play with now and which game they were going to play with later. Incorrect answers were corrected and rule checks were repeated a second time.

Before handing children the marbles, children's choices were explicitly stated ("You can use some marbles now for the small game. You can also save some marbles for later for the big game."). Children were given the bag of marbles, and the experimenter started a timer for 3 minutes and retreated to a corner to work while the children played. To discourage interaction, the experimenter wore headphones and minimized conversation, replying as needed with neutral statements such as "Remember I need to work right now." When the timer rang, another rule check was administered. Children were asked if they remembered what they would do or play now, and incorrect answers were corrected. The small marbles game was removed, and the big marbles game was retrieved from the other room, and children were given 3 minutes with the big marbles game.

For savers, a switch was made such that, unbeknownst to the child, a different big marbles game was retrieved that was identical in outward appearance, but on closer inspection, had putty hidden in various tubes, blocking the marble pathways. The experimenter expressed surprise at discovering that the big game was broken, bringing children's attention to the putty ("Oh no, I think the big game might be broken! See the sticky stuff? There's some here and here. I think the marbles won't go down now and they might get stuck inside!"). Children were prompted to try a marble to confirm that it would get stuck. Any remaining marbles were collected to prevent children from entertaining themselves with the marbles.

In both cases, children spent 3 minutes with the big marbles game, but were unable to play with it, either because they lacked marbles, or because the game was broken. In the control group, the experimenter told children that a mistake was made and they had to play another game first before the big marbles game. Children did not spend time with the big marbles game and instead were directly administered the vocabulary test. After all children completed the vocabulary test, the experimenter discovered another 3 marbles and exclaimed that they would play the same game again, and the initial instructions regarding the small and big marbles game were repeated. Regardless of performance, all children were given extra marbles so that all children had 5 marbles to play with a surprise extra big marbles game.

Delay Choice Baseline. The delay choice baseline task (e.g., Prencipe & Zelazo, 2005; Thompson et al., 1997) used in Study 1 was administered with the same stimuli and instructions. There were 9 trials in total balanced across three types of contingencies (i.e., 1 versus 2, 1 versus 4, 1 versus 6) presented in a fixed mixed order, using a food treat chosen by the child. For each test trial, children were presented with a choice between one treat now and more treats later.

Peabody Picture Vocabulary Test (PPVT-4). The standardized test of receptive vocabulary (Dunn & Dunn, 2007) used in Study 1 was administered according to standard administration. Children were shown a series of pages with pictures, and for each page, children were asked to point to 1 of 4 pictures that corresponded to a word spoken by the experimenter.

Executive Function Scale (EF Scale™). The EF Scale (Carlson, 2012) used and described in Study 1 was administered as an independent measure of set-shifting and cognitive flexibility. The EF Scale is a card-sorting task with 7 levels of complexity. The starting level for 3.5-year-olds, and 4.5-year-olds was level 3 and level 4, respectively. When children passed a level, more difficult levels were administered until children failed. If children failed the initial level, easier levels were administered in a backward fashion until children passed. In this manner, the highest level at which children could pass was determined. In the current study, the standardized newer protocol was used, with 5 trials in both pre-switch and post-switch. Children were required to sort 4 out 5 cards correctly in both pre-switch and post-switch to pass a level. Children's performance across Study 1 and 2 were comparable as measured by the highest level passed, with similar mean performance for the two age groups, and similar relations seen with other variables such as age and verbal ability. This suggests that the number of trials likely did not affect the highest level passed, and the two versions of the task used across studies were comparable.

Children's Behavior Questionnaire Very Short Form (CBQ-VSF). Parents were asked to complete the Very Short Form of the CBQ (Putnam & Rothbart, 2006) as in Study 1. Parents rated how often certain behaviors were true of their children in the last 6 months on a 7-point Likert scale, which resulted in three factors or dimensions (i.e., Surgency, Negative Affect, Effortful Control).

Childhood Routines Inventory (CRI). Parents completed the CRI (Evans et al., 1997), which consisted of 19-items describing various compulsive, ritualistic, and repetitive behaviors children may engage in (e.g., "Strongly prefer to stick to one game or activity rather than change to a new one"). Parents were asked to rate how often each behavior occurred on a 5-point Likert scale. A total mean score for compulsive behaviors was calculated from averaging all items, and two factors were calculated based on specific items that had been shown to load together (see Evan et al., 1997). A Just Right factor was derived from averaging 5 items that captured children's sensory-perceptual need to have things be carried out or arranged in a particular way (e.g., "Lines up objects in straight lines or symmetrical patterns"). A Repetitive Behaviors factor was derived from averaging 4 items that captured children's repetitive behaviors and need for things to be the same (e.g., "Prefers the same household schedule or routine every day").

Results

Descriptive. An alpha level of .05 was used for all statistical analyses unless otherwise noted. All children completed all tasks, although 2 children in the experimental group did not complete the delay choice task. No statistical differences were found between the experimental and control groups with respect to age, sex, baseline delaying, baseline saving, PPVT, or EF Scale, and therefore, the two groups were combined for baseline analyses. For more direct comparison with Study 1, children were classified as delayers (6 to 9 trials) and non-delayers (0 to 3 trials), with those scoring in the mid-range omitted from delay analyses. Of the remaining 105 children, 47 children (44.8%) were delayers and 58 children (55.2%) were non-delayers. Unlike in Study 1, here the proportion of delayers and non-delayers did not differ from chance according to a binomial test, $p = .329$. There was a significant age difference, with 27 out of 44 children (61.4%) delaying in 4.5-year-olds, and 20 out of 61 children (32.8%) delaying in 3.5-year-olds, $\chi^2(1, N = 105) = 8.44$, $p = .003$. Delaying was also marginally related to sex, with 29 out 55 females (52.7%) delaying, compared to only 18 out of 50 males (36.0%) delaying, $\chi^2(1, N = 105) = 2.96$, $p = .063$. There was no relation with PPVT scores, $t(102) = -.46$, $p = .648$. Analyses were rerun categorizing delayers (5 to 9 trials) and non-delayers (0 to 4 trials) without excluding children, and results were comparable. After applying a Bonferroni corrected alpha rate of .0165 (.05/3), age differences remained significant.

During the baseline saving task, only 37 children (26.1%) were classified as savers and 105 children (73.9%) as spenders. This was significantly different from chance as indicated by a binomial test, $p = .000$. As seen in Table 4, children's baseline status as a saver or spender was unrelated to age, $\chi^2(1, N = 142) = 1.34$, $p = .167$, and sex, $\chi^2(1, N = 142) = 2.07$, $p = .106$. Saving was also unrelated to baseline delaying, $\chi^2(1, N = 105) = 1.65$, $p = .148$. When analyses were rerun without excluding children in our categorization of delayers and non-delayers, there was a marginally significant relationship. Of the 37 savers, 21 were delayers (56.8%), whereas of the 103 spenders, only 43 were delayers (41.7%), $\chi^2(1, N = 140) = 2.47$, $p = .084$. Savers also scored significantly higher on PPVT ($M = 127.28$, $SD = 10.82$) than spenders ($M = 117.95$, $SD = 13.28$), $t(138) = -3.80$, $p = .000$. PPVT differences remained significant after applying a Bonferroni corrected alpha rate of .0125 (0.05/4).

Switching. In the experimental group, 43 of 102 children (42.2%) switched, which was significantly higher than the 5 of 40 children (12.5%) who spontaneous switched in the control group, $\chi^2(1, N = 142) = 11.30$, $p = .000$. This suggests that the manipulation was effective. Switching was unrelated to baseline saving, although the proportion of switching was slightly higher when switching from saving to spending. As shown in Table 5, 50.0% of savers switched from saving to spending, whereas 39.7% of spenders switched from spending to saving, $\chi^2(1, N = 102) = .80$, $p = .256$. In the experimental group, switching was related to age, with 65.3% of 4.5-year-olds switching compared to only 20.8% of 3.5-year-olds switching, $\chi^2(1, N = 102) = 20.72$, $p = .000$ (see Table 5). More females than males switched, with 26 of 51 females (51.0%) switching, compared to 17 of 51 males (33.3%) switching, although this difference was only marginally significant, $\chi^2(1, N = 102) = 3.26$, $p = .054$. Switching was also related to PPVT, with switchers ($M = 123.37$, $SD = 13.39$) scoring higher standard scores than non-switchers (*M* = 117.68, *SD* = 12.99), $t(100) = -2.16$, $p = .033$. After applying a Bonferroni corrected alpha rate of .0125 (.05/4), only age differences remained significant.

To examine whether children recognized which was the adaptive choice, children's self-reported preferences for playing the small game or waiting for the big game were analyzed. Preference data were available for 98 children in the experimental group. Of the 12 savers who switched to spending, 9 children (75.0%) reported preferring to play with the small game, consistent with their behavior. Of the 11 savers who did not switch to spending, 6 children (54.5%) reported preferring to play with the small game, inconsistent with their behavior. Of the 31 spenders who switched to saving, 29 children (93.5%) reported preferring waiting for the big game, consistent with their behavior. Of the 44 spenders who did not switch to saving, 36 children (81.8%) reported preferring to wait for the big game, inconsistent with their lack of saving. Examining both savers and spenders together, of the 43 switchers, only 5 (11.6%) reported an inconsistent preference, whereas of the 55 non-switchers, 42 (76.4%) reported an inconsistent preference, $\chi^2(1, N = 98) = 40.52$, $p = .000$. These results are to those observed in Study 1, and suggest that some children may have recognized the better choice, despite not demonstrating it in behavior.

EF Scale. As in Study 1, one of the main hypotheses was that children's ability to switch would be positively related to set-shifting as assessed using the EF Scale. An analysis of covariance was conducted with the highest level passed as the dependent variable, and switching and baseline saving as independent variables, controlling for age and verbal ability. Results revealed that age was significantly related to EF Scale performance, $F(1, 96) = 47.05$, $p = .000$, partial $\eta^2 = .33$, as were PPVT standard scores, $F(1, 96) = 4.24$, $p = .042$, partial $\eta^2 = .04$. Controlling for both age and PPVT, switching on the marbles task was a significant predictor of EF Scale performance, $F(1, 96) = 6.43$, $p = .013$, partial $\eta^2 = .06$. Adjusted means controlling for age and PPVT showed that switchers had a mean score of 4.03 ($SE = .17$) and non-switchers had a mean score of 3.23 ($SE = .14$). Baseline saving was a marginally significant predictor, $F(1, 96) = 3.04$, $p = .084$, partial $\eta^2 = .03$, with an adjusted mean of 3.98 for savers ($SE = .216$) and an adjusted mean of 3.55 for spenders ($SE = .119$). The interaction between switching and baseline saving was also marginally significant, $F(1, 96) = 2.98$ $p = .088$, partial $\eta^2 = .03$.

Given the marginally significant results, separate analyses of covariance, controlling for age and PPVT, were conducted for savers and spenders. Switching was not a significant predictor of EF Scale performance in savers, $F(1, 20) = .19$, $p = .665$, with adjusted means of 4.30 ($SE = .29$) for switchers and 4.12 ($SE = .29$) for non-switchers. Switching was a significant predictor of EF Scale performance in spenders, $F(1, 74) = 9.34$, $p = .003$, partial $\eta^2 = .11$, with adjusted means of 3.95 ($SE = .22$) for switchers and 2.99 ($SE = .17$) for non-switchers.

CBQ. Two-tailed Pearson correlation analyses were conducted and are presented in Table 6. Contrary to predictions, baseline delaying was unrelated to Surgency, and it was also unrelated to Effortful Control. Interestingly, delaying was positively related to Negative Affect, with delayers more likely to have higher reported negative affect, $r = .25$, $p = .01$. Baseline saving was unrelated to all three CBQ dimensions. Switching was significantly correlated with Effortful Control, $r = .21$, $p = .01$, but unrelated to Negative Affect, as was observed in Study 1. Unlike Study 1, Effortful Control was unrelated to the EF Scale, but was correlated with PPVT, $r = .20$, $p = .02$. Controlling for PPVT, since it was also related to switching, the relation between switching and Effortful Control remained significant, $r = .22$, $p = .03$. Given results indicating that it was easier for savers to switch compared to spenders, and that switching may be slightly different in these groups, separate correlational analyses were conducted for savers and spenders. In savers, switching was significantly related to Effortful Control, $r = .43$, $p = .04$, which remained significant after controlling for PPVT, $r = .43$, $p = .04$. In spenders, switching was marginally related to Effortful Control, $r = .21$ $p = .06$, although it became non-significant after controlling for PPVT, $r = .16$, $p = .17$.

CRI. Two-tailed Pearson correlation analyses were conducted and are presented in Table 6. Neither baseline delaying or saving were related to compulsive behaviors. Interestingly, sex was associated with Just Right behaviors, $r = -.22$, $p = .01$, and Total behaviors, $r = -.17$ $p = .04$ with males showing more compulsive behaviors than females. Switching was unrelated to compulsive behaviors. However, EF Scale was marginally related to Repetitive behaviors with greater EF associated with less repetitive behaviors, $r = -.15$, $p = .08$. Effortful Control was positively related to Just Right behaviors, $r = .21$, $p = .01$ and marginally to Total behaviors, $r = .16$, $p = .06$, with those scoring higher on effortful control also scoring higher on compulsive behaviors. Negative Affect was positively related to Just Right behaviors, $r = .34$, $p = .00$ and Total behaviors, $r = .33$, p

= .00, and marginally to Repetitive behaviors, $r = .16$, $p = .06$. Surgency was negatively related to Just Right behaviors, $r = -.34$, $p = .00$ and Total behaviors, $r = -.24$, $p = .00$.

Discussion

In the present study, we attempted to replicate the findings of Study 1 using a different paradigm that assessed saving behavior. More children were categorized as spenders than savers, similar to Study 1, where more children were categorized as non-delayers than delayers. However, the rate of saving was particularly low, lower than that of delaying, at only 26.1%. This was lower than the 39% rate of saving reported by Metcalf and Atance (2011).

In particular, we were interested in whether children would be able to modify their saving behavior based on past experiences that provided information about the environmental context with respect to contingencies and rewards. In general, the results were replicated those from the first study. A comparable proportion of children switched their behavior, 35.1% in Study 1 and 42.2% in Study 2. The proportion of switchers in the saving paradigm was similar to that reported by Metcalf and Atance (2011), who reported 39% of spenders switched to saving during a second trial, which is the same proportion in spenders, 39.7%, observed in Study 2. A greater proportion of children switched from saving to spending, which mirrors findings from Study 1 that showed more switching from delaying to not delaying. Unlike the delay choice task, the savings task involved one-trial learning, where children only had the opportunity for feedback on one trial. However, children were still able to adjust their responses, consistent with our finding that switchers and non-switchers in the delay choice paradigm could be differentiated by their first delay trial before having a chance to learn through subsequent trials. Replicating the findings from Study 1, many of the children who were unable to switch stated a preference for the adaptive choice, which was not reflected in their behavior. It suggests that while they could not behaviorally switch their response, they appeared to recognize the better choice.

Switchers, as before, scored higher on set-shifting, controlling for age and verbal ability. When savers and spenders were analyzed separately, results showed that this relation may be stronger in spenders who need to shift towards saving. In spenders, the adjusted mean difference between switchers and non-switchers was equivalent to a level on the EF Scale. Unlike the first study, age differences were clearly observed. Older children were more likely to switch than younger children. Compared to the delay choice task, the savings task may require more complex understanding of casual connections between various elements in the task, such as how the small and big marbles game events were related, and how children's actions had consequences for the games. There was also no explicit instruction about the consequences (i.e., if you use all your marbles, you will have none left for the big game) as there was during the delay choice modified task (i.e., treats may be lost in the box). Children's choices of saving or not saving were, however, clearly presented. Consistent with this view is that verbal ability, which may help in understanding more complex tasks, was related to baseline saving and marginally related to switching.

Future-oriented thinking might also have played a bigger role in this task, and prior research suggests developmental change during the preschool years (see Atance &

O'Neill, 2005). An important aspect of adaptive and flexible behavior concerns the ability to consider our future selves and use such considerations to guide our current behaviors. Children's ability to understand the concept of future time and to consider their future selves develops during the preschool years (see Atance & O'Neill, 2005 for review). Some studies have suggested that young children have difficulty planning for future novel events, which improves between the ages of 3 and 5 years (e.g., Russell, Alexis, & Clayton, 2010; Suddendorf & Busby, 2005; Suddendorf, Nielsen, & von Gehlen, 2011). However, others have found that under less demands, such as the absence of strong semantically related distractors, even 3-year-olds can select items needed for a future scenario (Atance & Meltzoff, 2006).

Beyond verbal ability, baseline saving was not strongly related to other variables, which was similar to baseline delaying. While sex differences were seen for baseline delaying, with females delaying more, this was not the case for baseline saving, although females were marginally more like to switch in a savings context. In past paradigms assessing children's saving behavior, others have not found reliable sex differences (e.g., Otto et al., 2006; Sonuga-Barke & Webley, 1993). There was also some indication that baseline saving may be positively related to baseline delaying, although results were only observed when no children were excluded.

In general, the main results across both studies showed parallel patterns of older children and females delaying more often at baseline, although no such relations were seen in baseline saving, a weak relationship between baseline behaviors and set-shifting, a trends towards more switching to immediate rewards, and significant relations between switching in a hot task context and set-shifting in a cool task context. Main findings of both studies are summarized in Table 7 for ease of comparison.

With regards to children's temperament, as in Study 1, baseline delaying was unrelated to surgency, although it was related to negative affect, with those showing greater negative affect more likely to delay. Greater fear and inhibited behavior, which may itself reduce approach tendencies, may make it easier to resist immediate rewards. Baseline saving, however, was unrelated to the three temperament dimensions. Switching was significantly related to effortful control, with switchers having higher reported effortful control. This was also observed in Study 1, but only in non-delayers. Here effortful control was unrelated to set-shifting, and thus the relation between switching and effortful control could not be explained by set-shifting. Interestingly, it suggests that perhaps, while there is some overlap between effortful control and EF, both may have independent contributions to switching behavior. The relation between switching and negative affect seen in Study 1 was not replicated here. It may be that 9 trials of repeatedly losing treats during the delay choice modified paradigm elicited greater frustration and negative emotion. In the marbles saving paradigm, there was only a one-time frustration event, although it lasted for 3 minutes.

In Study 2 we added a questionnaire assessing children's ritualistic and compulsive behaviors to assess whether flexibility as assessed in our experimental paradigms would be related to flexibility in children's everyday behaviors. Contrary to predictions, switching was unrelated to compulsive behaviors, although compulsive behaviors were related to other variables associated with switching. For example, males were reported to have more compulsive behaviors, and interestingly, males were also less likely to switch their saving behavior. Set-shifting was marginally related to repetitive behaviors, with greater scores on shifting predicting lower reported repetitive behaviors. An inclination towards repetitive behaviors may be associated also with a tendency toward perseverative behaviors, which is related to poor performance on set-shifting tasks (e.g., Pietrefesa & Evans, 2007). Compulsive behaviors, particularly "just right" behaviors, were also associated positively with effortful control. Although the positive association seems counterintuitive, it may be that effortful control is required to maintain "just right" aspects in children's environment. Moreover, the subscale of effortful control includes items related to perceptual sensitivity, which is a characteristic of compulsive behaviors as assessed by the CRI. Compulsive behaviors were also positively related to negative affect. This is consistent with previous findings indicating that compulsive behaviors are associated with fears and phobias (e.g., Evans et al., 1999; Pietrefesa & Evans, 2007). Items related to fear are included in the Negative Affect subscale. Surgency, which is typically negatively associated with negative affect, as was observed in the two present studies, was negatively related to compulsive behaviors.

General Discussion

The ability to adjust one's behavior flexibly in response to different situations is an important aspect of adaptive functioning. What constitutes adaptive or rational behavior changes with the environment and with current internal goals, states, and motivations. Sometimes it is better to exert control; other times, it is better to let go. Across two paradigms investigating two hot decision-making behaviors, children's responses were indeed sensitive to contextual changes and past experience, confirming previous findings in the delay of gratification literature. Children were able to flexibly adjust their behavior in both directions, towards delaying and saving, and towards immediate consumption. More interestingly, however, individual differences emerged with respect to this flexibility of behavior, with some children switching to new responses, and others perseverating on old responses. Children who adjusted their behavior were differentiated by greater EF, specifically set-shifting. This finding suggests that certain EF skills may underlie the adaptation of behavior to different environmental demands, whether it requires one to exert or to relinquish control.

However, the need and recruitment of EF may vary depending on the direction of change required and its difficulty. There was a trend towards greater switching in the direction of immediate consumption. In Kesek et al. (2011), their implicit priming procedure appeared more effective in shifting children towards immediate rewards, Bandura and Mischel (1965) found stronger and more reliable effects with delayers responding to a live model encouraging immediate gratification, and similarly, a large decline in waiting time was observed when children interacted with an unreliable experimenter (Kidd et al., 2013). It may be easier to bias individuals towards immediate gratification as most individuals are drawn to appetitive rewards and waiting is typically aversive. Accordingly, such a switch may be less taxing on executive processes, and potentially require less EF. In fact, when savers and spenders were analyzed separately in Study 2, results suggested that the relation between switching and EF may be stronger when switching from spending to saving than the reverse. Moreover, in Study 1, switching and effortful control was correlated in only non-delayers who needed to switch towards delaying.

Considering children's well-documented gains in EF during the preschool years (e.g., Carlson, 2005; Zelazo & Müller, 2002), age was surprisingly not consistently related to switching. Age differences were observed during the saving paradigm, but not during the delay choice paradigm. The saving paradigm was more complex, requiring children to make more connections regarding the task with less explicit guidance. Factors related to age and development may be most relevant in such problems that demand greater perception and understanding of social situations, complex causal connections, and various contingencies.

While there was a consistent relation observed between EF and switching behavior, no consistent relation was seen between EF and baseline behaviors. This is similar to previous findings in the literature regarding delay of gratification and other hot tasks. In light of the complex nature of these hot tasks and the multiple factors both internal and external that may contribute to the observed behavior, it is clear that these tasks are not a pure measure of control. Although EF processes may be recruited, their effect may be obscured by other influences on behavior, and relations between hot tasks and other self-control measures may not always be clearly observed or measured.

Nature of Children's Switching Behavior

Interestingly, children were capable of adapting their behavior based on one-trial learning, as seen in the savings paradigm, where a single feedback experience was enough to cue children to adjust their behavior. The finding that switchers and non-switchers could be differentiated by the first trial in the delay choice modified paradigm suggests that, even without the opportunity for actual feedback and direct experience, children were able to use new information about contingencies in the environment to guide their behavior. Nonetheless, there was a role for learning. Differences in children's choices between the first and second half of trials in the delay choice modified paradigm showed that by the end of the task, children were more likely to select the non-risky option than at the beginning of the task, presumably from repeated feedback across trials. This was pattern was even seen in non-switchers, although they did not fully shift their behavior. Children's switching behavior is likely not a matter of simply switching or not switching, but can be understood in terms of how quickly and how much children switch, and how much information and feedback is needed to facilitate change in behavior.

Interestingly, some children who did not show switching behavior seemed to understand which choice was more adaptive. In Study 1, approximately half of the children were unable to switch behaviorally, but stated a preference for the adaptive choice. In Study 2, over three quarters of the children who did not switch stated a preference that was inconsistent with their actions. Thus it appears that while some children were unable to behaviorally switch, they demonstrated some understanding about which choices or responses were more adaptive. This abulic dissociation between knowledge and using that knowledge is a common phenomenon seen in the DCCS (Zelazo & Frye, 1988). Children know the rules of the new game, but are unable to follow through in action. They may have trouble reflecting on the rules and integrating them in a high-order fashion to facilitate switching (Zelazo & Frye, 1988). Nisan and

Koriat (1977) reported similar findings, that over two thirds of children who did not delay answered that a “smart” child would chose to delay. In a more recent study (Prencipe & Zelazo, 2005), 3-year-olds tended to choose immediate rewards for themselves but later rewards for another person. Since these children can recognize the rational or adaptive choice, although they may not follow through behaviorally, they may be more mature than those who do not show adaptation of their behavior and cannot identify the adaptive choice. It is possible that these children may be able to benefit with a great number of trials and feedback, which is also supported by trends showing some behavior change across trials.

Temperament and Compulsive Behaviors

In hot tasks involving delaying or saving, children’s temperamental reactivity may impact how children respond to rewarding and punishing stimuli and events, depending on children’s approach tendencies and their predisposition to fear, withdrawal, and negative affect. Surprisingly, there was little association observed between these dimensions of temperament and children’s baseline delay and saving behavior, although delaying was positively related to negative affect in Study 2. Just as the complexity of standard hot tasks may obscure associations with EF control processes, it may also make it difficult to differentiate and identify associations with temperamental reactivity. That is, while these aspects of reactivity may have an effect on behavior, in the context of other potentially stronger influences, they may be more difficult to detect or less relevant.

Despite the few findings related to baseline behaviors, the regulatory component of effortful control was related to switching, but also showed some relations to EF. It suggests a close overlapping relationship between effortful control and EF, and in fact, some of the items related to effortful control on the CBQ include behaviors related to inhibitory control and attention, which are important processes of EF. However, they may also make independent contribution toward flexible behavior. In Study 2, effortful control was not correlated with EF, but was correlated with switching. Items related to perceptual sensitivity and noticing changes in the environment were part of the subscale measuring effortful control. It is possible that these behaviors may facilitate switching behavior by enhancing children's perception of context changes, which is necessary before children can act on those changes. Greater negative affect was also related to less switching, which highlights the potentially important role of negative emotions. Negative emotions may emerge when children are pushed towards behavior that is less preferred, less rewarding, or perhaps contrary to their natural temperament. Likewise, negative emotions may also arise when children fail to adjust their behavior to changes in the environment or past events, and experience negative or frustrating consequences.

In view of these various findings regarding children's flexibility as measured by our experimental paradigms, it was of interest to examine whether children's flexibility in the laboratory would be reflected in their everyday behaviors. During the preschool years, there is a developmental increase in the number of ritualistic and compulsive behaviors (Evans et al., 1997), and it was predicted that children who showed more flexibility would be less likely to engage in compulsive behaviors. Surprisingly, children's compulsive behaviors was unrelated to switching in our paradigms, although they were related to other skills and capacities that were correlated positively with switching, such as effortful control and EF. Whereas greater EF was related to less compulsive behaviors, greater effortful control was related to more compulsive behaviors, especially “just right” behaviors. The positive relationship with effortful control, as discussed above, may be related to perceptual sensitivity, and noticing minute changes in the environment, which is an aspect of “just right” behaviors.

Limitations and Future Directions

The parallel results of the two studies are largely encouraging, but several limitations should be considered. First, although our interpretation is that children’s set-shifting skills facilitated their ability to adapt their behavior in the delay and savings task, the EF Scale was administered only at the end of the session, and thus one can argue that the better performance of switchers was related to generalization effects. That is, switching during the delay and savings tasks led to switching on the EF Scale. Simple exposure to the opportunity to switch was likely not an influence, as indicated by the lack of spontaneous switching in the control group, although children who switched may have somehow increased their propensity to switch again during the EF Scale. It should be noted that children switched of their own accord without feedback from the experimenter and no explicit training or encouragement regarding switching was provided. To rule out the possibility with certainty, however, the order of tasks should be counterbalanced. It would also be informative if one could show that training or developmental improvements in EF leads to greater switching in the current paradigms.

Relatedly, another complexity not explored in the current studies is that what is adaptive in a given environment may vary across individuals. Recent research has shown that what is adaptive can be shaped by characteristics of one's early childhood environment, such as harshness and unpredictability. These types of characteristics may be found in lower socioeconomic status (SES) environments. In a series of experiments by Griskevicius, Delton, Robertson, and Tybur (2011), when adults were primed with mortality cues through a false news article, individuals of lower childhood SES reported attitudes and plans for having children at an earlier age than non-primed controls. In contrast, individuals of higher childhood SES reported plans for having children later than controls. Although initially there appeared to be no difference between the groups, when primed with mortality cues, a context change, one group shifted towards a faster life strategy (i.e., having children earlier), while the other shifted towards a slower life strategy (i.e., having children later). This is argued to be adaptive respectively for each group within an evolutionary framework (see Ellis, Figueredo, Brumbach, & Schlomer, 2009). The same phenomenon was replicated with risk preferences, delay discounting, and probabilistic discounting (Griskevicius et al., 2013; Griskevicius, Tybur, Delton, & Robertson, 2011), which all showed that when primed with mortality or resource scarcity, lower childhood SES individuals shifted towards risky behaviors and immediate rewards, whereas higher childhood SES individuals shifted in the opposite direction towards less risky behaviors and delaying. This illustrates that what is considered an adaptive response to a particular problem can differ across individuals. Further research is needed to better understand what are adaptive responses for different individuals and what are important variables where individuals begin to diverge, such as childhood SES.

Continued research is needed to answer questions regarding the domain specificity or domain general function of this flexibility in behavior. For example, flexibility in one type of behavior may or may not be related to flexibility in another type of behavior. In the current study, if the same children experienced both paradigms, would the switchers in the delay paradigm also be switchers in the saving paradigm? In light of the associations between flexibility, core EF processes, and effortful control, we might expect at least some correspondence between flexibility in various situations. Although delaying and saving behaviors retain some important differences, they are also quite similar, and it would be important to examine flexibility with other types of behaviors. Furthermore, the current paradigms have emphasized the adaptive goals of maximizing utility and pleasure, but other types of adaptive goals, such as those in the social, moral, and cultural domain may also motivate behavioral changes. Research examining the types of behaviors in which children can show flexibility, and the types of adaptive goals to which children are sensitive will help us gain a better understanding of when flexibility in behavior is important and the extent to which EF are recruited in these circumstances.

Conclusion

Overall, the current study findings show that at a basic level, children's delay of gratification and saving behavior are sensitive to context and past experience, and furthermore, children's ability to adjust their behavior according to what is presently adaptive is likely supported by EF. Importantly, these findings show that although some children may exhibit poor control by default, in the right context, the same children can be capable of demonstrating control, and in fact, may have more mature EF. When there is no consistent goal or correct way to behave, it is not surprising that children can appear more controlled in one situation and less controlled in another.

This highlights the methodological value of assessing behavior across multiple contexts, especially when tasks are sensitive to contextual factors, such as delay of gratification and other hot tasks. With multiple measurements, we can better understand how the behavior is affected by key factors, and gain a more accurate and broader assessment of a child's abilities and developmental profile, as performance in a single context may not be representative of a child's competence. The importance of evaluating behavior in multiple contexts has been noted in work with behaviorally inhibited children (Davidson, Jackson, & Kalin, 2000) and children with attention deficit hyperactivity disorder (Barkley, 1990), where symptomatic behavior is sometimes only elicited or detected in certain contexts. As children appear heterogeneous in their response to changing environmental demands, measurements of this variability itself may also impart valuable information.

Delay of gratification and saving are clearly complex decision-making behaviors that integrate multiple influences including sociocultural factors, situational factors, and individual characteristics. As is distinctive of many hot tasks, delay of gratification and saving problems are not clearly defined. They involve ambiguity, uncertainty, emotion, and motivation. Because there are no clear rules to guide behavior, we must look to the context to identify the most optimal or adaptive course of action given past, present, and future concerns. What is considered adaptive can differ across individuals and groups of individuals, and there are many reasons one may choose to delay or not delay. Delaying and saving is not always the most adaptive response, and accordingly, lack of such behavior should not be indicative of deficient control or willpower. It is important in our assessments of children's behavior in such tasks to acknowledge the important distinction between competence and performance, and to interpret behaviors carefully in light of contextual and individual factors.

Because complex and "hot" real-world problems are dependent on the context, the ability to flexibly adjust one's behavior to changing demands is a critical aspect of adaptive functioning. Adaptive goals can vary, from, maximizing rewards and pleasure to maintaining social relationships. In all cases, we must attend to both contextual information in the environment and internal goals, states, and motivations, and analyze the fit between the self and environment to respond in the most adaptive manner. EF processes likely play an important role in facilitating flexible behavior, especially when certain behaviors or decisions are difficult to enact because they conflict with one's desires or natural tendencies, or perhaps elicits fear or anxieties. Development itself has been described as movement towards increasing flexibility, differentiation, and organization, allowing the individual to adapt to the environment with a greater range of means, behavioral patterns, and functions (Sroufe & Rutter, 1984). As children's behavioral repertoire and EF skills develop over time, it confers an in increasing ability to navigate and engage with the world in a flexible, adaptive, and successful manner.

References

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*(4), 463-496. doi:10.1037/h0076860

Atance, C. M., & Meltzoff, A. N. (2006). Preschoolers' current desires warp their choices for the future. *Psychological Science, 17*(7), 583-587. doi:10.1016/j.cogdev.2005.05.001

Atance, C. M., & O'Neill, D. K. (2005). The emergence of episodic future thinking in humans. *Learning and Motivation, 36*(2), 126-144. doi:10.1016/j.lmot.2005.02.003

Ayduk, O., Mendoza-Denton, R., Mischel, W., Downey, G., Peake, P. K., & Rodriguez, M. (2000). Regulating the interpersonal self: Strategic self-regulation for coping with rejection sensitivity. *Journal of Personality and Social Psychology, 79*(5), 776-792. doi:10.1037/0022-3514.79.5.776

Ayduk, O., Zayas, V., Downey, G., Cole, A. B., Shoda, Y., & Mischel, W. (2008). Rejection sensitivity and executive control: Joint predictors of borderline personality features. *Journal of Research in Personality, 42*(1), 151-168. doi:10.1016/j.jrp.2007.04.002

Bandura, A., & Mischel, W. (1965). Modification of self-imposed delay of reward through exposure to live and symbolic models. *Journal of Personality and Social Psychology, 2*(5), 698-705. doi:10.1037/h0022655

Barkley, R. A. (1990). A critique of current diagnostic criteria for attention deficit hyperactivity disorder: Clinical and research implications. *Journal of*

Developmental and Behavioral Pediatrics, 11(6), *343-352.* doi:10.1097/00004703-199012000-00014

Beck, D. M., Schaefer, C., Pang, K., & Carlson, S. M. (2011). Executive function in preschool children: Test-retest reliability. *Journal of Cognition and Development, 12*(2), 169-193. doi:10.1080/15248372.2011.563485

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

Block, J. H., & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Development of cognition, affect, and social relations: Minnesota Symposia on Child Psychology* (Vol. 13, pp. 39-101). Hillsdale, NJ: Erlbaum.

Bonato, D. P., & Boland, F. J. (1983). Delay of gratification in obese children. *Addictive Behaviors, 8*(1), 71-74. doi:10.1016/0306-4603(83)90059-X

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

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science, 15*(3), 118-121. doi:10.1111/j.0963-7214.2006.00419.x

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, February). Scaling the development of executive function in preschool children. In S. M. Carlson (Organizer), *Methodological advances in the study of executive function development*. Paper symposium presented at the SRCD Themed Meeting: Developmental Methodology, Tampa, FL.

Carlson, S., Davis, A., & Leach, J. (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 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., & Schaefer, C. (2012). Executive Function Scale for Early Childhood. Test manual. University of Minnesota.

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., Zayas, V., & Guthormsen, A., (2009). Neural correlates of decision making on a gambling task. *Child Development, 80*(4), 1076-1096. doi:10.1111/j.1467-8624.2009.01318.x

Carlson, S. M., Zelazo, P. D., & Faja, S. (2013). Executive function. In P. D. Zelazo (Ed.), *Oxford handbook of developmental psychology* (pp. 706-724). New York: Oxford University Press.

Carlson, S. M., & Zelazo, P. D. (2011). The value of control and the influence of values. *Proceedings of the National Academy of Sciences, 108*(41), 16861-16862. doi:10.1073/pnas.1113235108

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology, 54*(1-3)*,* 241-257. doi:10.1016/S0301-0511(00)00058-2

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

Chang, F., & Burns, B. M. (2005). Attention in preschoolers: Associations with effortful control and motivation. *Child Development, 76*(1), 247-263. doi:10.1111/j.1467-8624.2005.00842.x

Crone, E. A., & van der Molen, M. W. (2004). Developmental changes in real life decision making: Performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Developmental Neuropsychology, 25i(3)*, 251-279. doi:10.1207/s15326942dn2503_2

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences, 11*(3), 97-104. doi:10.1016/j.tics.2006.12.005

Dalen, L., Sonuga-Barke, E. J. S., Hall, M., & Remington, B. (2004). Inhibitory deficits, delay aversion and preschool AD/HD: Implications for the dual pathway model. *Neural Plasticity, 11*(1-2), 1-11. doi:10.1155/NP.2004.1

Dasgupta, P., & Maskin, E. (2005). Uncertainty and hyperbolic discounting. *American Economic Review, 95*(4), 1290-1299. doi:10.1257/0002828054825637

Davidson, R. J., Jackson, D. C., & Kalin, N. H. (2000). Emotion, plasticity, context, and regulation: Perspectives from affective neuroscience. *Psychological Bulletin, 126*(6)*,* 890-909. doi:10.1037/0033-2909.126.6.890

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

Derryberry, D., & Rothbart, M. K. (1997). Reactive and effortful processes in the organization of temperament. *Development and Psychopathology, 9*(4)*,* 633-652. doi:10.1017/S0954579497001375

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do". *Developmental Psychobiology, 29*(4), 315-334. doi:10.1002/(SICI)1098-2302(199605)29:4<315::AID-DEV2>3.0.CO;2-T

Duckworth, A. L., & Carlson, S. M. (2013). Self-regulation and school success. In B. W. Sokol, F. M. E. Grouzet, & U. Müller (Eds.), *Self-regulation and autonomy: Exploring the social, developmental, educational, and neurological dimensions of human conduct* (pp. 208-230). New York: Cambridge.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*(12), 939-944. doi:10.1111/j.1467-9280.2005.01641.x

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4 Manual.* Bloomington, MN: NCS Pearson Inc.

Eigsti, I., Zayas, V., Mischel, W., Shodea, Y., Ayduk O., Dadlani, M. B., . . . Casey, B. J. (2006). Attentional control in preschool predicts cognitive control at age eighteen. *Psychological Science, 17*(6), 478-84. doi:10.1111/j.1467-9280.2006.01732.x

Eisenberg, N., Guthrie, I. K., Fabes, R. A., Resier, M., Murphy, B. C., Holgren, R., . . . Losoya, S. (1997). The relations of regulation and emotionality to resiliency and competent social functioning in elementary school children. *Child Development, 68*(2)*,* 295-311. doi:10.1111/j.1467-8624.1997.tb01941.x

Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk: The impact of harsh versus unpredictable environments on the evolution and development of life history strategies. *Human Nature, 20*(2)*,* 204-268. doi:10.1007/s12110-009-9063-7

Evans, D. W., Leckman, J. F., Carter, A., Reznick, J. S., Henshaw, D., King, R. A., & Pauls, D. (1997). Ritual, habit, and perfectionism: The prevalence and development of compulsive-like behavior in normal young children. *Child Development, 68*(1), 58-68. doi:10.1111/j.1467-8624.1997.tb01925.x

Evans, D. W., Gray, F. L., & Leckman, J. F. (1999). The rituals, fears and phobias of young children: Insights from development, psychopathology and neurobiology. *Child Psychiatry and Human Development, 29*(4), 261-276. doi:10.1023/A:1021392931450

Evans, D. W., Lewis, M. D., & Iobst, E. (2004). The role of the orbitofrontal cortex in normally developing compulsive-like behaviors and obsessive-compulsive disorder. *Brain and Cognition, 55*(1), 220-234. doi:10.1016/S0278-2626(03)00274-4

Fawcett, T. W., McNamara, J. M., & Houston, A. I. (2012). When is it adaptive to be patient? A general framework for evaluating delayed rewards. *Behavioural Processes, 89*(2), 128-136. doi:10.1016/j.beproc.2011.08.015

Fox, E. (1994). Attentional bias in anxiety: A defective inhibition hypothesis. *Cognition and Emotion, 8*(2)*,* 165-195. doi:10.1080/02699939408408934

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

Funder, D. C., & Block, J. (1989). The role of ego-control, ego-resiliency, and IQ in delay of gratification of adolescence. *Journal of Personality and Social Psychology, 57*(6), 1041-1050. doi:10.1037/0022-3514.57.6.1041

Gallimore, R., Weiss, L. B., & Finney, R. (1974). Cultural differences in delay of gratification: A problem of behavior classification. *Journal of Personality and Social Psychology, 30*(1), 72-80. doi:10.1037/h0036606

Garon, N., & Moore, C. (2006). Negative affectivity predicts individual differences in decision making in preschoolers. *Journal of Genetic Psychology: Research and Theory on Human Development, 167*(4), 443-462. doi:10.3200/GNTP.167.4.443-462

Garon, N., & Moore, C. (2007). Awareness and symbol use improves future-oriented decision making in preschoolers. *Developmental Neuropsychology, 31*(1), 39-59. doi:10.1080/87565640709336886

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., . . . Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2*(10), 861-863. doi:10.1038/13158

Gladstone, R. (1969). Age, cognitive control and extinction. *Journal of Experimental Child Psychology, 7*(1)*,* 31–35. doi:10.1016/0022-0965(69)90083-6

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., . . . Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, *101*(21), 8174-8179. doi:10.1073/pnas.0402680101

Griskevicius, V., Ackerman, J. M., Cantú, S. M., Delton, A. W., Robertson, T. E., Simpson, J. A., … Tybur, J. M. (2013). When the economy falters, do people spend or save? Responses to resource scarcity depend on childhood environments. *Psychological Science, 24*(2), 197-205. doi:10.1177/0956797612451471

Griskevicius, V., Delton, A. W., Robertson, T. E., & Tybur, J. M. (2011). Environmental contingency in life history strategies: The influence of mortality and socioeconomic status on reproductive timing. *Journal of Personality and Social Psychology, 100*(2), 241-254. doi:10.1177/0956797612451471

Griskevicius, V., Tybur, J. M., Delton, A. W., & Robertson, T. E. (2011). The influence of mortality and socioeconomic status on risk and delayed rewards: A life history theory approach. *Journal of Personality and Social Psychology, 100*(6), 1015-1026. doi:10.1037/a0022403

Happaney, K., & Zelazo, P. D. (2004). Resistance to extinction: A measure of orbitofrontal function suitable for children? *Brain and Cognition, 55*(1), 171-184. doi:10.1016/S0278-2626(03)00270-7

Herrnstein, R. J. (1990). Rational choice theory: Necessary but not sufficient. *American Psychologist, 45*(3), 356-367. doi:10.1037/0003-066X.45.3.356

Herzberger, S. D., & Dweck, C. S. (1978). Attraction and delay of gratification. *Journal of Personality, 46*(2), 215-227. doi:10.1111/j.1467-6494.1978.tb00176.x

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., & 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. doi:10.1207/s15326942dn2802_4

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology, 387*(2), 167-178. doi:10.1002/(SICI)1096-9861(19971020)387:2<167::AID-CNE1>3.0.CO;2-Z

Johnson, W. G., Parry, W., & Drabman, R. S. (1978). The performance of obese and normal size children on a delay of gratification task. *Addictive Behaviors, 3*(3-4), 205-208. doi:10.1016/0306-4603(78)90020-5

Kagan, J., Reznick, J. S., & Gibbons, J. (1989). Inhibited and uninhibited types of children. *Child Development, 60*(4), 838-845. doi:10.2307/1131025

Kerr, A., & Zelazo, P. D. (2004). Development of "hot" executive function: The children's gambling task. *Brain and Cognition, 55*(1), 148-157. doi:10.1016/S0278-2626(03)00275-6

Kesek, A., Cunningham, W. A., Packer, D. J., & Zelazo, P. D. (2011). Indirect goal priming is more powerful than explicit instruction in children. *Developmental Science, 14*(5), 944-948. doi:10.1111/j.1467-7687.2011.01043.x

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition, 126*(1), 109-114. doi:10.1016/j.cognition.2012.08.004

Kochanska, G., Coy, K. C., Tjebkes, T. L., & Husarek, S. J. (1998). Individual differences in emotionality in infancy. *Child Development, 69*(2), 375-390. doi:10.1111/j.1467-8624.1998.tb06196.x

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*(2), 263-277. doi:10.1111/j.1467-8624.1997.tb01939.x

Kochanska, G., Murray, K. T., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (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

Koriat, A., & Nisan, M. (1978). Delay of gratification as a function of exchange values and appetitive values of the rewards. *Motivation and Emotion, 2*(4), 375-390. doi:10.1007/BF00993332

Krueger, R. F., Caspi, A., Moffitt, T. E., White, J., & Stouthamer-Loeber, M. (1996). Delay of gratification, psychopathology, and personality: Is low self-control specific to externalizing problems? *Journal of Personality, 64*(1), 107-129. doi:10.1111/j.1467-6494.1996.tb00816.x

Lemmon, K., & Moore, C. (2007). The development of prudence in the face of varying future rewards. *Developmental Science, 10*(4), 502-511. doi:10.1111/j.1467-7687.2007.00603.x

Lindstrom, D. R., & Shipman, V. C. (1972). Technical report 12: Mischel technique. In V. C. Shipman (Ed.), *Disadvantaged children and their first school experiences: Technical report series.* Princeton, NJ: Educational Testing Service.

Lucey, J. V., Burness, C. E., Costa, D. C., Gacinovic, S., Pilowsky, L. S., Ell, P. J., … Kerwin, R. W. (1997). Wisconsin Card Sorting Task (WCST) errors and cerebral blood flow in obsessive-compulsive disorder (OCD). *Psychology and Psychotherapy, 70*(4), 403-411. doi:10.1111/j.2044-8341.1997.tb01916.x

Mahrer, A. R. (1956). The role of expectancy in delayed reinforcement. *Journal of Experimental Psychology, 52*(2), 101-106. doi:10.1037/h0040837

McGuire, J. T., & Kable, J. W. (2012). Decision makers calibrate behavioral persistence on the basis of time-interval experience. *Cognition, 124*(2), 216-266. doi:10.1016/j.cognition.2012.03.008

McGuire, J. T., & Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological Review, 120*(2), 395-410. doi:10.1037/a0031910

Melikian, L. (1959). Preference for delayed reinforcement: An experimental study among Palestinian Arab refugee children. *Journal of Social Psychology, 50*(1), 81-86. doi:10.1080/00224545.1959.9921980

Metcalf, J. L., & Atance, C. M. (2011). Do preschoolers save to benefit their future selves? *Cognitive Development, 26*(4), 371-382. doi:10.1016/j.cogdev.2011.09.003

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review, 106*(1), 3-19. doi:10.1037/0033-295X.106.1.3

Michaelson, L., de la Vega, A., Chatham, C. H., & Munakata, Y. (2013). Delaying gratification depends on social trust. *Frontiers in Psychology, 4*, 355. doi:10.3389/fpsyg.2013.00355

Mischel, W. (1958). Preference for delayed reinforcement: An experimental study of a cultural observation. *Journal of Abnormal and Social Psychology, 56*(1)*,* 57-61. doi:10.1037/h0041895

Mischel, W. (1966). Theory and research on the antecedents of self-imposed delay of reward. In B.A. Maher (Ed.), *Progress in experimental personality research* (Vol. 3, pp. 85-132). New York: Academic Press.

Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 249-292). San Diego, CA: Academic Press.

Mischel, W., Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., . . . Shoda, Y. (2011). 'Willpower' over the life span: Decomposing self-regulation. *Social Cognitive and Affective Neuroscience, 6*(2), 252-256. doi:10.1093/scan/nsq081

Mischel, W., & Baker, N. (1975). Cognitive appraisals and transformations in delay behavior. *Journal of Personality and Social Psychology, 31*(2)*,* 254-261. doi:10.1037/h0076272

Mischel, W., Ebbesen, E. B., Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology, 21*(2), 204-218. doi:10.1037/h0032198

Mischel, W., Grusec, J., & Masters, J. C. (1969). Effects of expected delay time on the subjective value of rewards and punishments. *Journal of Personality and Social Psychology, 11*(4), 363-373. doi:10.1037/h0027265

Mischel, W., & Masters, J. C. (1966). Effects of probability of reward attainment on responses to frustration. *Journal of Personality and Social Psychology, 3*(4), 390-396. doi:10.1037/h0022966

Mischel, W., & Metzner, R. (1962). Preference for delayed reward as a function of age, intelligence, and length of delay interval. *Journal of Abnormal and Social Psychology, 64*(6), 425-431. doi:10.1037/h0045046

Mischel, H. N., & Mischel, W. (1983). The development of children's knowledge of self-control strategies. *Child Development, 54*(3), 603-619.

Mischel, W., & Patternson, C. J. (1976). Substantive and structural elements of effective plans for self-control. *Journal of Personality and Social Psychology, 34*(5), 942-950. doi:10.1037/0022-3514.34.5.942

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*(2), 246-268. doi:10.1037/0033-295X.102.2.246

Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology, 54*(4)*,* 687-696. doi:10.1037/0022-3514.54.4.687

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

Montgomery, G. T. (1976). Delay of gratification in children: A function of magnitude of reward and the delay cue. *Journal of Experimental Child Psychology, 21*(3)*,* 549-555. doi:10.1016/0022-0965(76)90082-5

Murray, K. T., & Kochanska, G. (2002). Effortful control: Factor structure and relation to externalizing and internalizing behaviors. *Journal of Abnormal Child Psychology, 30*(5), 503-514. doi:10.1023/A:1019821031523

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220-246. doi:10.1037/0033-2909.126.2.220

Nisan, M., & Koriat, A. (1984). The effect of cognitive restructuring on delay of gratification. *Child Development, 55*(2), 492-503. doi:10.2307/1129960

Olson, S. L. (1989). Assessment of impulsivity in preschoolers: Cross-measure convergences, longitudinal stability, and relevance to social competence. *Journal of Clinical Child Psychology, 18*(2)*,* 176-183. doi:10.1207/s15374424jccp1802_9

Olson, S. L., & Hoza, B. (1993). Preschool developmental antecedents of conduct problems in children beginning school. *Journal of Clinical Child Psychology, 22*(1)*,* 60-67. doi:10.1207/s15374424jccp2201_6

Olson, S. L., Schilling, E. M., & Bates, J. E. (1999). Measurement of impulsivity: Construct coherence, longitudinal stability, and relationship with externalizing problems in middle childhood and adolescence. *Journal of Abnormal Child Psychology, 27*(2)*,* 151-165. doi:10.1023/A:1021915615677

Otto, A. (2003). Saving in childhood and adolescence: Insights from developmental psychology. *Economics of Education Review, 33*, 8-18. doi:10.1016/j.econedurev.2012.09.005

Otto, A. M. C., Schots, P. A. M., Westerman, J. A. J., & Webley, P. (2006). Children's use of saving strategies: An experimental approach. *Journal of Economic Psychology, 27*(1), 57-72. doi:10.1016/j.joep.2005.06.013

Overman, W. H., Bachevalier, J., Schuhmann, E., & Ryan, P. (1996). Cognitive gender differences in very young children parallel biologically-based cognitive gender differences in monkeys. *Behavioral Neuroscience, 110*(4), 673-684. doi:10.1037/0735-7044.110.4.673

Pietrefesa, A. S., & Evans, D. W. (2007). Affective and neuropsychological correlates of children's rituals and compulsive-like behaviors: Continuities and discontinuities with obsessive-compulsive disorder. *Brain and Cognition, 65*(1), 36-46. doi:10.1016/j.bandc.2006.02.007

Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., & Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of Experimental Child Psychology, 108*(3), 621-637. doi:10.1016/j.jecp.2010.09.008

Prencipe, A., & Zelazo, P. D. (2005). Development of affective decision-making for self and other: Evidence for the integration of first- and third-person perspectives. *Psychological Science, 16*(2), 501-505. doi:10.1111/j.0956-7976.2005.01564.x

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

Rachlin, H. (2000). *The science of self-control.* Cambridge, MA: Harvard University Press.

Raver, C. C., Blackburn, E. K., Bancroft, M., & Torp, N. (1999). Relations between effective emotional self-regulation, attentional control, and low-income preschoolers' social competence with peers. *Early Education and Development, 10*(3)*,* 333-350. doi:10.1207/s15566935eed1003_6

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' preacademic skills: Self-regulation as a mediating mechanism. *Child Development, 82*(1), 362-378. doi:10.1111/j.1467-8624.2010.01561.x

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

Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16*(4), 207-212. doi:10.1111/j.1467-8721.2007.00505.x

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology, 78,* 122-135. doi:10.1037/0022-3514.78.1.122

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development, 72*(5), 1394-1408. doi:10.1111/1467-8624.00355

Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social emotional and personality development* (5th ed., pp. 105–176). New York: Wiley.

Rothbart, M. K., Ellis, L. K., Rueda, M. R., & Posner, M. I. (2003). Developing mechanisms of temperamental effortful control. *Journal of Personality, 71*(6), 1113-1143. doi:10.1111/1467-6494.7106009

Russell, J., Alexis, D., & Clayton, N. (2010). Episodic future thinking in 3- to 5-year-old children: The ability to think of what will be needed from a different point of view. *Cognition, 114*(1)*,* 56-71. doi:10.1016/j.cognition.2009.08.013

Scheres, A., Dijkstra, M., Ainslie, E., Balkan, J., Reynolds, B., Sonuga-Barke, E., & Castellanos, F. X. (2006). Temporal and probabilistic discounting of rewards in children and adolescents: Effects of age and ADHD symptoms. *Neuropsychologia, 44*(11), 2092-2103. doi:10.1016/j.neuropsychologia.2005.10.012

Schwarz, J. C., & Schrager, J. B., & Lyons, A. E. (1983). Delay of gratification by preschoolers: Evidence for the validity of the choice paradigm. *Child Development, 54*(3), 620-625. doi:10.2307/1130048

Shiner, R. L., & DeYoung, C. G. (2013). The structure of temperament and personality traits: A developmental perspective. Executive function. In P. D. Zelazo (Ed.), *Oxford handbook of developmental psychology* (pp. 113-141). New York: Oxford University Press.

Shipe, D., & Lazare, S. (1969). Two measures of delay of gratification: Age and socioeconomic status in young children. *Proceedings of the Annual Convention of the American Psychological Association, 4,* 461-462.

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology, 26*(6), 978-986. doi:10.1037/0012-1649.26.6.978

Silverman, I. W. (2003). Gender differences in delay of gratification: A meta-analysis. *Sex Roles, 49*(9-10), 451-463. doi:10.1023/A:1025872421115

Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-Regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly, 22*(2), 173-187. doi:10.1016/j.ecresq.2007.01.002

Sonuga-Barke, E. J. S., Dalen, L., & Remington, B. (2003). Do executive deficits and delay aversion make independent contributions to preschool attention-deficit/hyperactivity disorder symptoms? *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(11), 1335-1342. doi:10.1097/01.chi.0000087564.34977.21

Sonuga-Barke, E. J. S., & Webley, P. (1993). *Children's saving: A study in the development of economic behavior.* Hillsdale, NJ: Lawrence Erlbuam.

Sroufe, L. A., & Rutter, M. (1984). The domain of developmental psychopathology. *Child Development, 55*(1), 17-29. doi:10.2307/1129832

Staub, E. (1972). Effects of persuasion and modeling on delay of gratification. *Developmental Psychology, 6*(1), 166-177. doi:10.1037/h0032212

Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child Development, 80*(1), 28-44. doi:10.1111/j.1467-8624.2008.01244.x

Strickland, B. R. (1972). Delay of gratification as a function of race of the experimenter. *Journal of Personality and Social Personality, 22*(1), 108-112. doi:10.1037/h0032385

Suddendorf, T., & Busby, J. (2005). Making decisions with the future in mind: Developmental and comparative identification of mental time travel. *Learning and Motivation, 36*(2), 110-125. doi:10.1016/j.lmot.2005.02.010

Suddendorf, T., Nielsen, M., & von Gehlen, R., (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science, 14*(1), 26-33. doi:10.1111/j.1467-7687.2010.00950.x

Thompson, C., Barresi, J., & Moore, C. (1997). The development of future-oriented prudence and altruism in preschoolers. *Cognitive Development, 12*(2), 199-212. doi:10.1016/S0885-2014(97)90013-7

Thorell, L. B. (2007). Do delay aversion and executive function deficits make distinct contributions to the functional impact of ADHD symptoms? A study of early academic skill deficits. *Journal of Child Psychology and Psychiatry, 48*(11), 1061-1070. doi:10.1111/j.1469-7610.2007.01777.x

Toner, I. J., Holstein, R. B., & Hetherington, E. M. (1977). Reflection-impulsivity and self-control in preschool children. *Child Development, 48*(1), 239-245. doi:10.2307/1128904

Toplak, M. E., Jain, U., & Tannock, R. (2005). Executive and motivational processes in adolescents with Attention-Deficit-Hyperactivity Disorder (ADHD). *Behavioral and Brain Functions, 1*(1), 8. doi:10.1186/1744-9081-1-8

Trentacosta, C. J., & Shaw, D. S. (2009). Emotional self-regulation, peer rejection, and antisocial behavior: Developmental associations from early childhood to early adolescence. *Journal of Applied Developmental Psychology, 30*(3), 356-365. doi:10.1016/j.appdev.2008.12.016

Wärneryd, K. E. (1999). *The psychology of saving: A study on economic psychology.* Cheltenham, UK: Edward Elgar Publishing.

Willoughby, M., Kupersmidt, J., Voegler-Lee, & M., Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Developmental Neuropsychology, 36*(2), 162-180. doi:10.1080/87565641.2010.549980

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., & Cunningham, W. A. (2007). Executive function: Mechanisms underlying emotion regulation. In J. Gross (Ed.), *Handbook of emotion regulation* (pp. 135-158). New York: Guilford Press.

Zelazo, P. D., & Frye, D. (1988). Cognitive complexity and control: II. The development of executive function in childhood. *Current Directions in Psychological Science, 7*(4), 121-126. doi:10.1111/1467-8721.ep10774761

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 445-469). Oxford: Blackwell.

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

Zohar, A. H., LaBuda, M., & Moschel-Ravid, O. (1995). Obsessive-compulsive behaviors and cognitive functioning: A study of compulsivity, frame shifting and Type A activity patterns in a normal population. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 8*(3), 163-167.

Appendix 1

Table 1

Number of Baseline Delayers and Non-delayers by Age and Sex in Study 1

	3.5 years			4.5 years			Total		
Sex	D	ND	TOT	D	ND	TOT	D	ND	TOT
Male	9 (20.4)	35 (79.5)	44 (100)	11 (33.3)	22 (66.7)	33 (100)	20 (26.0)	57 (74.0)	77 (100)
Female	15 (42.8)	20 (57.1)	35 (100)	16 (57.1)	12 (42.8)	28 (100)	31 (49.2)	32 (50.8)	63 (100)
Total	24 (30.4)	55 (69.6)	79 (100)	27 (44.3)	34 (55.7)	61 (100)	51 (36.4)	89 (63.6)	140 (100)

Note. D = Delayers, ND = Non-delayers. Percentages within age group and sex are presented in brackets.

Table 2

Number of Switchers and Non-switchers by Age and Baseline Delaying in Study 1

	Delayers			Non-delayers			Total		
Age	S	NS	TOT	S	NS	TOT	S	NS	TOT
3.5 years	8 (50.0)	8 (50.0)	16 (100)	9 (31.0)	20 (69.0)	29 (100)	17 (37.8)	28 (62.2)	45 (100)
4.5 years	8 (47.1)	9 (52.9)	17 (100)	2 (13.3)	13 (86.7)	15 (100)	10 (31.3)	22 (68.8)	32 (100)
Total	16 (48.5)	17 (51.1)	33 (100)	11 (25.0)	33 (75.0)	44 (100)	27 (35.1)	50 (64.9)	77 (100)

Note. S = Switcher, NS = Non-switchers. Percentages within baseline delaying and age group are presented in brackets.

Table 3

Correlations Between Child Characteristics, Task Performance, and CBQ Dimensions in Study 1

Variables	1	2	3	4	5	6	7	8	9
1. Age	—								
2. Sex	.02	—							
3. Delay Baseline	.14[†]	.24*	—						
4. PPVT	.04	.06	.00	—					
5. EF Scale	.56**	.10	.17*	.20*	—				
6. Switching	-.07	.16	.24*	-.15	.12	—			
7. Effortful Control	-.05	.34**	.10	.01	.15[†]	.13	—		
8. Negative Affect	.02	.11	.08	.03	.03	-.21[†]	-.04	—	
9. Surgency	.03	-.07	-.13	.20*	-.07	-.00	-.22**	-.20*	—

Note. Correlations are presented for the full sample ($N = 140$); however, correlations involving *Switching* are based on only the experimental group ($n = 77$) where the variable is relevant.

[†]$p < .10$. *$p < .05$. **$p < .01$.

Table 4

Number of Baseline Savers and Spenders by Age and Sex in Study 2

	3.5 years			4.5 years			Total		
Sex	Sa	Sp	TOT	Sa	Sp	TOT	Sa	Sp	TOT
Male	8 (21.6)	29 (78.4)	53 (100)	14 (42.4)	19 (57.6)	20 (100)	22 (31.4)	48 (68.6)	73 (100)
Female	8 (22.2)	28 (77.8)	49 (100)	7 (19.4)	29 (80.6)	20 (100)	15 (20.8)	57 (79.2)	69 (100)
Total	16 (21.9)	57 (78.1)	73 (100)	21 (30.4)	48 (69.6)	69 (100)	37 (26.1)	105 (73.9)	142 (100)

Note. Sa = Savers, Sp = Spenders. Percentages within age group and sex are presented in brackets.

Table 5

Number of Switchers and Non-switchers by Age and Baseline Saving in Study 2

	Savers			Spenders			Total		
Age	S	NS	TOT	S	NS	TOT	S	NS	TOT
3.5 years	6 (54.5)	5 (45.5)	11 (100)	5 (11.9)	37 (88.1)	42 (100)	11 (20.8)	42 (79.2)	53 (100)
4.5 years	6 (46.2)	7 (53.8)	13 (100)	26 (72.2)	10 (27.8)	36 (100)	32 (65.3)	17 (34.7)	49 (100)
Total	12 (50.0)	12 (50.0)	24 (100)	31 (39.7)	47 (60.3)	78 (100)	43 (42.2)	59 (57.8)	102 (100)

Note. S = Switchers, NS = Non-switchers. Percentages within baseline saving and age group are presented in brackets.

Table 6

Correlations Between Child Characteristics, Task Performance, CBQ Dimensions, and CRI Behaviors in Study 2

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	—												
2. Sex	.03	—											
3. Delay Baseline	.28**	.17	—										
4. Saving Baseline	.10	-.12	.13	—									
5. PPVT	.06	.03	.04	.31**	—								
6. EF Scale	.68**	.10	.15	.17*	.22**	—							
7. Switching	.45**	$.18^{\dagger}$	.25*	.09	.21*	.54**	—						
8. Effortful Control	.04	.32**	.08	-.06	.20*	.11	.25*	—					
9. Negative Affect	.20*	.08	.25*	-.05	.02	.01	-.05	.07	—				
10. Surgency	.04	.05	-.05	.03	-.00	.03	.08	$-.14^{\dagger}$	-.25**	—			
11. Just Right	.05	-.08	.07	-.11	-.06	.00	.11	.21*	.34**	-.34**	—		
12. Repetitive	-.15	-.22*	.00	-.11	-.12	$-.15^{\dagger}$	-.03	.02	$.16^{\dagger}$	-.13	.60**	—	
13. Total CRI	-.01	-.17*	.07	-.12	-.08	-.06	.02	$.16^{\dagger}$	.33**	-.24**	.87**	.84**	—

Note. Correlations are presented for the full sample (N = 142); however, correlations involving *Switching* are based on only the experimental group (n = 102) where the variable is relevant, and correlations involving *Delay Baseline* exclude children delaying 4 or 5 trials (n = 105) to maintain consistency with Study 1.

$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

Table 7

Comparison of Main Findings Related to Baseline and Switching Behaviors in Study 1 and Study 2

	Delay Baseline[a] (Study 1)	Delay Baseline[a] (Study 2)	Saving Baseline (Study 2)	Switching – Delay (Study 1)	Switching – Saving (Study 2)
Age group	Yes[†] 4.5 > 3.5	Yes** 4.5 > 3.5	No	No	Yes*** 4.5 > 3.5
Sex	Yes** F > M	Yes[†] F > M	No	No	Yes[†] F > M
Delay[a]	--	--	No	Yes*[c] D > ND	--
Saving	--	No	--	--	No
PPVT	No	No	Yes*** Sa > Sp	No	Yes*[c] S > NS
EF Scale[b]	No	No	No	Yes* S > NS	Yes* S > NS

Note. F = Females, M = Males; Sa = Savers, Sp = Spenders; D = Delayers, ND = Non-delayers; S = Switchers, NS = Non-switchers.
[a] Excluding children who delayed 4 or 5 trials.
[b] Controlled for age and PPVT.

[c] Not significant after Bonferroni correction for multiple tests

[†]$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Appendix 2

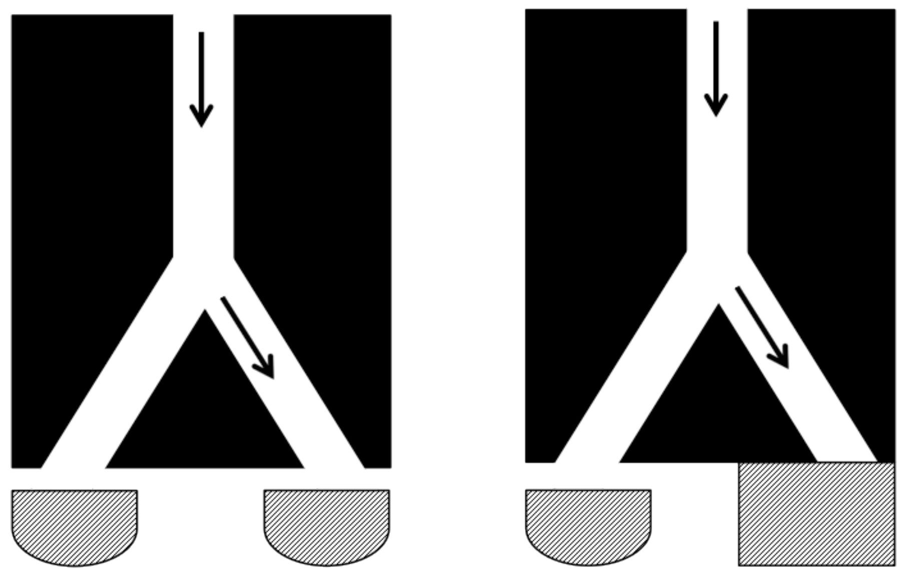

Figure 1. Panels used to deliver immediate and delayed treats to child during modified delay choice task. Each panel was associated with either the *now* choice or the *later* choice, with one conferring a risk of losing treats (i.e., box), and one providing certainty of treats (i.e., two bowls). Assignment of panels was based on children's baseline performance, with the risky panel assigned to children's preferred response in the experimental condition, and non-preferred response in the control condition.

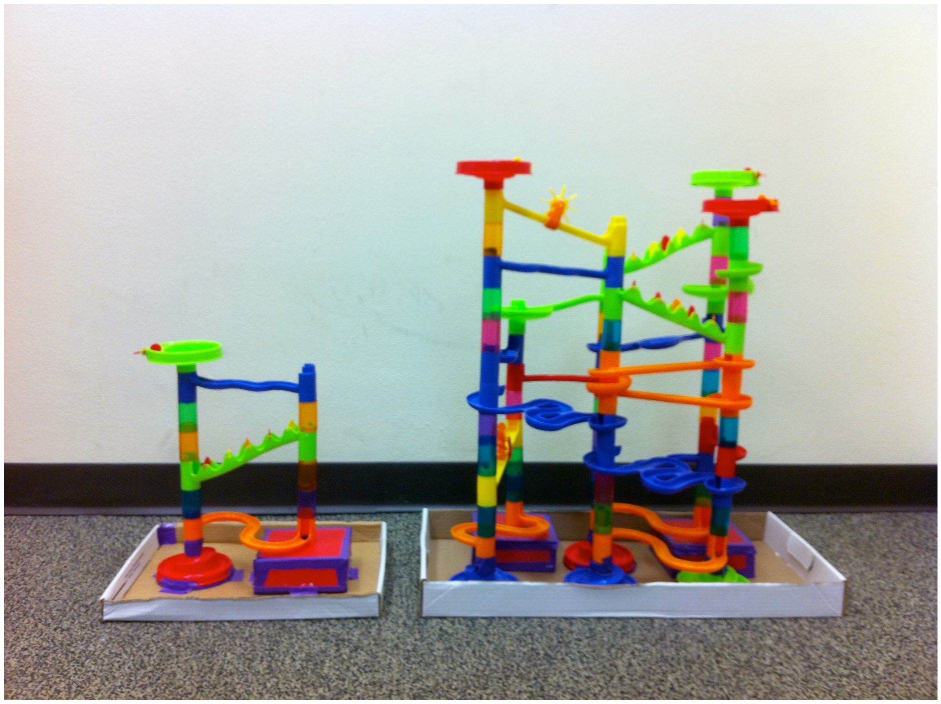

Figure 2. Small marbles game and big marbles game used in saving task.