DEVELOPMENTAL NEUROPSYCHOLOGY, *28*(2), 617–644

Assessment of Hot and Cool Executive Function in Young Children: Age-Related Changes and Individual Differences

Donaya Hongwanishkul
Institute of Child Development
University of Minnesota

Keith R. Happaney
Department of Psychology
Lehman College
City University of New York

Wendy S. C. Lee and Philip David Zelazo
Department of Psychology
University of Toronto

Although executive function (EF) is often considered a domain-general cognitive function, a distinction has been made between the "cool" cognitive aspects of EF more associated with dorsolateral regions of prefrontal cortex and the "hot" affective aspects more associated with ventral and medial regions (Zelazo & Müller, 2002). Assessments of EF in children have focused almost exclusively on cool EF. In this study, EF was assessed in 3- to 5-year-old children using 2 putative measures of cool EF (Self-Ordered Pointing and Dimensional Change Card Sort) and 2 putative measures of hot EF (Children's Gambling Task and Delay of Gratification). Findings confirmed that performance on both types of task develops during the preschool period. However, the measures of hot and cool EF showed different patterns of relations with each other and with measures of general intellectual function and temperament. These differences provide preliminary evidence that hot and cool EF are indeed distinct, and they encourage further research on the development of hot EF.

Correspondence should be addressed to Philip David Zelazo, Department of Psychology, University of Toronto, Canada M5S 3G3. E-mail: zelazo@psych.utoronto.ca

Executive function (EF), which refers to the psychological processes involved in cognitive control, is usually considered a domain-general cognitive function (e.g., Denckla & Reiss, 1997; Zelazo, Carter, Reznick, & Frye, 1997). Although it is recognized that EF encompasses a variety of subfunctions (e.g., working memory, attentional flexibility) that work together in the service of goal-directed problem solving, it is generally assumed that these subfunctions, as well as the higher order function of EF, operate in a consistent fashion across content domains ranging from theory of mind (e.g., Frye, Zelazo, & Palfai, 1995; Hughes, 1998; Perner & Lang, 1999) to understanding symbols and word meaning (e.g., Bialystok & Martin, 2003; Deák , 2000; O'Sullivan, Mitchell, & Daehler, 2001) to reasoning about physical causality (e.g., Frye, Zelazo, Brooks, & Samuels, 1996).

In contrast to a domain general view of EF, however, recent research on the functions of ventral and medial regions of prefrontal cortex (VM–PFC) suggests that EF may operate differently in different contexts (e.g., Bechara, 2004; Clark, Cools, & Robbins, 2004; Damasio, 1994; Dias, Robbins, & Roberts, 1996; Hauser, 1999; Miller & Cohen, 2001; Rolls, 2004). In light of this research, Zelazo and Müller (2002) distinguished between two aspects of EF—the relatively "hot" affective aspects associated with VM–PFC and the more purely cognitive, "cool" aspects associated with dorsolateral prefrontal cortex (DL–PFC; cf. Metcalfe & Mischel, 1999). According to this heuristic framework, cool EF is more likely to be elicited by abstract, decontextualized problems (e.g., sorting by color, number, or shape in the Wisconsin Card Sorting Test [WCST]; Grant & Berg, 1948), whereas hot EF is more likely to be elicited by problems that involve the regulation of affect and motivation (i.e., the regulation of basic limbic system functions), as when one is required to reappraise the motivational significance of a stimulus (e.g., learning to choose advantageously in the Iowa Gambling Task; Bechara, Damasio, Damasio, & Anderson, 1994). Although it seems likely that measures of EF always require a combination of hot and cool EF (e.g., Manes et al., 2002) and, hence, that the difference between hot and cool EF is always a matter of degree, the hot–cool distinction encourages researchers to adopt a broader conception of EF that captures its more affective aspects.

Prefrontal cortex as a whole undergoes considerable growth during the course of childhood, as indicated by age-related changes in volumes of gray and white matter (Giedd et al., 1999; Gogtay et al., 2004; Matsuzawa et al., 2001; Pfefferbaum et al., 1994), and interhemispheric connectivity (P. M. Thompson et al., 2000), among other measures. Until recently, however, research on the functional implications of these structural changes has focused almost exclusively on DL–PFC and cool EF. This research indicates that cool EF improves considerably between 3 and 4 years of age (e.g., Espy, Kaufmann, Glisky, & McDiarmid, 2001; Hughes, 1998; Zelazo, Müller, Frye, & Marcovitch, 2003; see also Zelazo & Müller, 2002, for review). Much less is known about the development of hot EF, although there is some indication that VM–PFC develops earlier than DL–PFC (e.g., Gogtay et al., 2004; Orzhekhovskaya, 1981), and interest in hot EF is growing (e.g., Blair, 2002; Carlson, Davis, & Leach, in press; Kerr & Zelazo, 2004; Pérez-Edgar & Fox, in press; Prencipe & Zelazo, 2005; see Happaney, Zelazo, & Stuss, 2004, and the June 2004 special issue of *Brain and Cognition*). In particular, there has been growing interest in the development of affective decision making, or decision making about events that have emotionally significant consequences (i.e., meaningful rewards and/or losses).

The distinction between hot and cool EF requires further clarification, but it has the potential to shed light on the role of EF in clinical disorders with childhood onset. Although deficits in EF have been implicated in a wide variety of disorders, including autism (e.g., Hughes, Russell, & Robbins, 1994), phenylketonuria (e.g., Diamond, Prevor, Callender, & Druin, 1997), and various externalizing disorders (e.g., Barkley, 1997; Hughes, Dunn, & White, 1998; Pennington, 1997), it seems unlikely that the same aspects of EF are involved in each disorder (Pennington & Ozonoff, 1996). Indeed, Zelazo and Müller (2002) suggested that whereas autism may be primarily a disorder of hot EF with secondary impairments in cool EF (cf. Dawson, Meltzoff, Osterling, & Rinaldi, 1998), ADHD may be mainly a disorder of cool EF (although differences may exist among subtypes and as a function of comorbidity; e.g., see Dinn, Robbins, & Harris, 2001).

The study presented here was designed to examine the early development of cool and hot EF in a single study. It was also designed to examine whether (and if so, how) these two aspects of EF may be related to one another and to measures of general intellectual functioning and temperament. Performance on two putative measures of cool EF—Self-Ordered Pointing task (Petrides & Milner, 1982) and a version of the Dimensional Change Card Sort (DCCS; Zelazo et al., 2003)—was compared to performance on two putative measures of hot EF—the Children's Gambling Task (Kerr & Zelazo, 2004) and a Delay of Gratification task (Prencipe & Zelazo, 2005). The rationale for selecting these measures is reviewed next.

MEASURES OF COOL EF

EF has long been considered synonymous with DL–PFC function, and two classic cases of DL–PFC function are working memory and flexible rule use. Working memory involves the simultaneous manipulation and maintenance of a representation so that this representation can guide responding (e.g., Baddeley, 1986; Daneman & Merikle, 1996). Neuropsychological and neuroimaging studies have consistently demonstrated the involvement of DL–PFC in working memory, in both adults (e.g., Curtis, Zald, & Pardo, 2000; Jacobson, 1936; Kane & Engle, 2002; Petrides & Milner, 1982) and children (Tsujimoto, Yamamoto, Kawaguchi, Koizumi, & Sawaguchi, 2004).

A common measure of working memory in adult patients is the Self-Ordered Pointing task (Petrides & Milner, 1982), in which participants are presented with a visual array of items, with the position of these items randomly changing from trial to trial. Participants begin by pointing to one item and are instructed not to point to the same item twice. Thus, to perform well, participants must keep the previously selected items in mind and use this information to inform subsequent responses. Developmental research has shown that performance on measures of working memory similar to the Self-Ordered Pointing task increases during the preschool period (see Gathercole, 1998, for a review). For example, Diamond et al. (1997) used a task in which rewards were placed in several (three or six) clearly distinct boxes that were scrambled from trial to trial. These authors found that children's performance improved from 15 months to about 7 years of age (the oldest age tested). Similar improvement on a spatial self-ordered working memory task from the Cambridge Neuropsychological Testing Automated Battery has been found during the early and middle childhood period (Luciana & Nelson, 1998, 2002). Archibald and Kerns (1999) used a more standard version of this task, in which children were presented with drawings of objects arranged in a matrix. The same objects were presented on each trial but in different locations, and children were instructed to point to a different picture on each trial. However, these authors examined older children—children between the ages of 7 and 12 years. To our knowledge, a standard picture version of the Self-Ordered Pointing task has never been used with preschoolers.

Flexible rule use has frequently been assessed by the WCST (Grant & Berg, 1948), which is often regarded as "the prototypical EF task in neuropsychology" (Pennington & Ozonoff, 1996, p. 55). In this task, participants are presented with target cards that differ on various dimensions (shape, color, and number) and then shown test cards that match different target cards on different dimensions. Participants must discover the rule according to which each card must be sorted. After a certain number of consecutive correct responses, the correct dimension is switched, and participants must discover this new rule. Patients with DL–PFC lesions typically perseverate on the WCST, continuing to sort by the initial dimension (Milner, 1964). Recent research has confirmed the role of DL–PFC in performance on the WCST and other measures of task switching (e.g., Lombardi et al., 1999; Nagahama et al., 2001; Rogers, Andrews, Grasby, Brooks, & Robbins, 2000; Sohn, Ursu, Anderson, Stenger, & Carter, 2000; Wang, Kakigi, & Hoshiyama, 2001; see Demakis, 2003, for a review).

Research with children reveals that flexible rule use improves considerably between 3 and 5 years of age. One frequently used measure is the DCCS (Frye, Zelazo, & Palfai, 1995; Zelazo et al., 2003), in which children are asked to sort a series of colored shapes, first by one dimension (e.g., color) and then by the other (e.g., shape). Whereas 4-year-olds switch flexibly, 3-year-olds systematically perseverate on the preswitch rules during the postswitch phase, despite being able to describe the rules they fail to use (e.g., Bialystok, 1999; Carlson & Moses, 2001; Jacques, Zelazo, Kirkham, & Semcesen, 1999; Kirkham, Cruess, & Diamond, 2003; Munakata & Yerys, 2001; Perner, Stummer, & Lang, 1999; Zelazo, Frye, & Rapus, 1996). That is, like DL–PFC patients on the WCST, 3-year-olds display what Teuber (see Discussion section in Milner, 1964) called "a curious dissociation between knowing and doing" (p. 333). The DCCS resembles the WCST in that cards must be sorted first by one dimension and then by another. In contrast to the WCST, however, children are explicitly told the rules by which to sort cards; they do not need to infer them. In our study, children who passed the postswitch phase of the standard version were given a new, more difficult border version (see the Method section).

MEASURES OF HOT EF

Interest in hot aspects of EF has been bolstered by findings that patients with VM–PFC damage exhibit substantial impairments in their social and emotional decision-making abilities despite good performance on traditional (i.e., cool) measures of EF, such as working memory and flexible rule use (e.g., Anderson, Bechara, Damasio, Tranel, & Damasio, 1999; Bechara, Damasio, Tranel, & Anderson, 1998). To study the basis of these impairments in the laboratory, researchers have developed a number of measures, including measures of gambling (e.g., Bechara, 2004), risky decision making (e.g., Rogers, Everitt, et al., 1999; Rogers, Owen, et al., 1999), guessing with feedback (e.g., Elliot, Frith, & Dolan, 1997), and delay discounting (Monterosso, Ehrman, Napier, O'Brien, & Childress, 2001). Despite their differences, these measures all assess flexible decision making about events that have emotionally significant consequences (i.e., meaningful rewards and/or losses).

One of the most widely used measures of VM–PFC function is the Iowa Gambling Task (Bechara et al., 1994). In this task, participants are presented with four decks of cards that, when turned, reveal a combination of gains and losses (measured in play money). Participants are initially given a stake of $2,000 facsimile dollars and asked to win as much money as possible by choosing cards from any of the four decks (one card per trial). Consistently selecting from two of the decks (the advantageous decks) results in a net gain, whereas selecting from the other two (the disadvantageous decks) results in a net loss. Each card from the disadvantageous decks provides a higher reward than each card from the advantageous decks ($100 vs. $50), but the variable (and unpredictable) losses associated with cards from disadvantageous decks are much larger on average than the losses associated with the advantageous decks. Bechara and colleagues have demonstrated that, whereas healthy controls learn to avoid the disadvantageous decks, patients with damage to VM–PFC continue to choose from these disadvantageous decks (e.g., Bechara et al., 1994; Bechara, Tranel, & Damasio, 2000; Bechara, Tranel, Damasio, & Damasio, 1996; but see Manes et al., 2002, for some qualifications regarding lesion focus and extent).

Kerr and Zelazo (2004) created the Children's Gambling Task (a simplified version of the Iowa Gambling Task), in which children had to choose between (a) cards that offered more rewards (candies) per trial but were disadvantageous across trials due to occasional large losses and (b) cards that offered fewer rewards per trial but were advantageous overall. On later trials, 4-year-olds made more advantageous choices than would be expected by chance, whereas 3-year-olds made fewer. Other adaptations of the Iowa Gambling Task have revealed age-related improvements in performance in somewhat older children (Crone & van der Molen, 2004; Garon & Moore, 2004). However, 4-year-olds in the study by Kerr and Zelazo were far from ceiling, and we expected that this task would be developmentally sensitive until at least 5 years of age.

Another measure of affective decision making is delay discounting (e.g., Green, Myerson, & Ostaszewski, 1999; Monterosso et al., 2001), a paradigm originally developed for use with nonhuman animals. In this task, individuals are asked to choose between smaller, immediate rewards and larger, delayed rewards (e.g., \$9 immediately vs. \$10 in 1 week). By varying the delay and the amount offered immediately, it is possible to calculate the rate at which a reward is discounted over time. Rats with lesions to orbitofrontal cortex, a part of VM–PFC, are more likely than controls to opt for small, immediate rewards (e.g., Mobini et al., 2002; Kheramin et al., 2003), and evidence suggests that similar regions of PFC play an important role in this type of decision making in human participants as well (e.g., Krawczyk, 2002; Rahman, Sahakian, Cardinal, Rogers, & Robbins, 2001; Rogers, Owen, et al., 1999). In addition, Monterosso et al. (2001) found that performance on the Iowa Gambling Task was significantly correlated with performance on the delay discounting task in a group of cocaine-dependent individuals, and these authors concluded that the two tasks tap similar affective decision-making processes.

Versions of this delay paradigm (referred to as Delay of Gratification) have been used extensively with children (Mischel, Shoda, & Rodriguez, 1989, for review). In one version (the choice version), children choose between an immediate reward of lower value and a delayed reward of higher value. Although early studies found few age differences within the preschool range (e.g., Schwarz, Schrager, & Lyons, 1983; Toner, Holstein, & Hetherington, 1977), C. Thompson, Barresi, and Moore (1997) used a modified choice paradigm and found a significant increase between 3 and 4 years of age in children's tendency to choose delayed rewards. Prencipe and Zelazo (2005) adapted the procedure used by C. Thompson et al. to include additional trials. These authors used nine trial types, created by crossing three types of reward (stickers, pennies, candies) and three types of choice (one now vs. two later, one now vs. four later, one now vs. six later), and found that 4-year-olds were more likely to choose delayed rewards than were 3-year-olds. In addition, whereas 3-year-olds were less likely to choose delayed rewards than would be expected based on chance responding, 4-year-olds were more likely to do so. These results were robust across all three types of reward.

THIS STUDY: OBJECTIVES AND HYPOTHESES

The study presented here was designed to investigate the development of both cool and hot EF in a single sample of children to assess whether these aspects of EF were related and whether they showed similar patterns of relations with measures of general intellectual functioning and temperament. Two putative measures of cool EF—a Self-Ordered Pointing task and a modified DCCS—were compared to two putative measures of hot EF—the Children's Gambling Task and Delay of Gratification. It was expected that the measures of cool EF would be more highly correlated with each other than with the measures of hot EF, and vice versa.

It should be kept in mind, however, that like VM–PFC and DL–PFC, hot and cool EF are presumed to be parts of an interactive functional system—in the normal case, they should work together, even in a single situation. Thus, it is probably impossible to design a task that is a pure measure of hot or cool EF (although it should be possible to design tasks that emphasize one or the other). For this reason, and because the same genetic and environmental influences may affect all aspects of PFC, it was expected that hot and cool measures would be related to one another to some extent.

It was also expected that hot and cool measures might show different patterns of correlations with measures of general intellectual functioning and temperament. More specifically, it was expected that performance on cool EF tasks would be highly related to general intellectual functioning, given both the relatively abstract nature of measures of general intellectual functioning and previous research using measures of cool EF (e.g., Ardila, Pineda, & Rosselli, 2000; Arffa, Lovell, Podell, & Goldberg, 1998; Duncan et al., 2000; Gray, Chabris, & Braver, 2003). In contrast, given the well-documented failure of standardized measures of intelligence to capture differences in emotional and social competence (e.g., Gardner, 1983; Goleman, 1995), performance on measures of hot EF tasks was expected to be only weakly related to general intellectual functioning, if at all. Finally, hot and cool EF were expected to be related to different dimensions of temperament, as assessed by the Children's Behavior Questionnaire (CBQ; Rothbart, Ahadi, & Hershey, 1994). Whereas children's emotional reactivity to negative situations (i.e., the factor of Negative Affectivity) was expected to be positively related to hot EF (just as adults' levels of self-reported negative affect are related to VM–PFC activity; Zald, Mattson, & Pardo, 2002), children's attentional self-regulation (i.e., the factor of Effortful Control) was expected to be positively related to cool EF (e.g., Rothbart et al., 1994; Ruff & Rothbart, 1996; Wolfe & Bell, 2003).

METHODS

Participants

One hundred six children between 3.0 and 5.9 years of age were recruited from a database of parents who had expressed interest in having their children participate in research. The study consisted of two separate sessions, which occurred approximately 2 weeks apart. Six of the initial 106 children did not return for the second session, and 2 refused to play during both sessions. The remaining 98 children were divided into three age groups, with the final sample consisting of 33 children at 3 years (16 girls, 17 boys; $M = 41.02$, $SD = 3.84$, range = 35.9–46.8 months), 32 at 4 years (16 girls, 16 boys; $M = 54.06$, $SD = 3.61$, range = 48.0–58.8 months), and 33 at 5 years (16 girls, 17 boys; $M = 66.13$, $SD = 3.00$, range = 61.4–71.3 months).

Design and Overview of the Procedure

Children were tested individually. At the beginning of each session, the experimenter played with children for 10 min to help them become comfortable with both the experimenter and experimental setting. In the first session, children received three standardized tests of general intellectual function, administered in a counterbalanced order. In addition, parents were asked to fill out the CBQ, which took about 30 min to complete. The first session took approximately 1 hr to complete.

In the second session, children received four measures of EF in a counterbalanced order: the Self-Ordered Pointing task, the DCCS, the Children's Gambling Task, and Delay of Gratification. This session took approximately 1 hr 15 min, with a break (approximately 10 min) provided after completion of the first two tasks.

Measures of Cool EF

Self-Ordered Pointing task. Children were shown sets of pictures presented on laminated sheets of paper (21.6 × 27.9 cm) in a three-ring binder (see Figure 1). The pictures across sets were distinct. To demonstrate the task, children were first shown a sheet containing two pictures and instructed to select one, after which the page was turned and a new sheet was shown that contained the same two pictures in different locations. Children were told that they must point to a picture they had not yet chosen. If children responded correctly to these instructions, the test phase began. If children did not respond correctly, they were given another demonstration trial. All children successfully completed this practice phase within two demonstration trials.

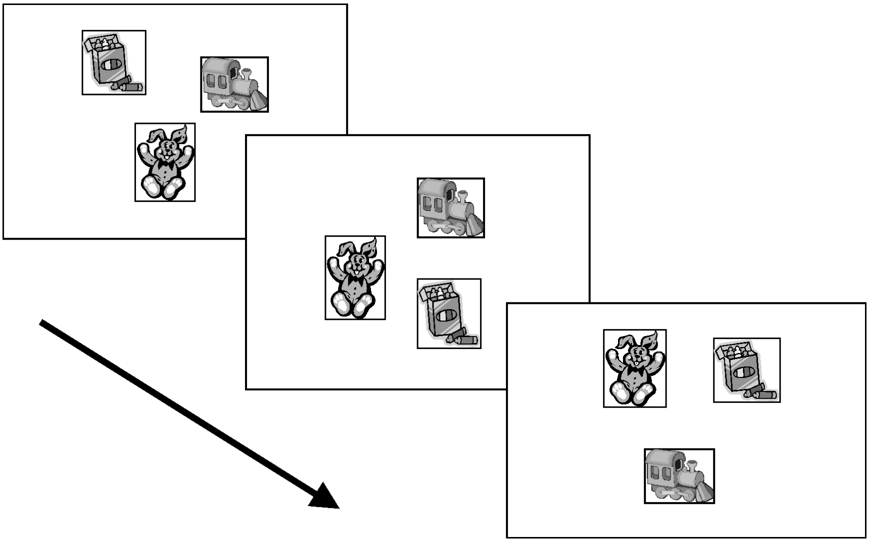

FIGURE 1 Sample picture set from the Self-Ordered Pointing task.

Test trials were presented in the same fashion as the demonstration trial(s). However, test trials started with 3 pictures, and the number of pictures per trial set was increased by 1 (up to a maximum of 10) until children erred on two consecutive trials. That is, if children erred on a particular trial set (e.g., if they chose the same picture twice), they received an additional trial set containing the same number of pictures. If children erred on this second trial set, the task was terminated. The dependent measure was the highest number of pictures in the last trial set on which children were correct.

DCCS. All children received the standard version of the DCCS. The order in which dimensions were presented (e.g., color first) was counterbalanced across age and sex. Target cards (a blue rabbit and a red boat) were first affixed to sorting trays, where they remained (visible) throughout the task (both versions). Children were then told a pair of rules (i.e., shape rules or color rules) for separating test cards, which consisted of red rabbits and blue boats. Children who sorted by color during the preswitch phase were told, “If it’s red put it here, but if it’s blue put it there.” The experimenter then sorted two test cards facedown into each tray to illustrate what children were supposed to do. After these demonstration trials, preswitch test trials commenced. On each of six preswitch trials, the experimenter stated the relevant rules, randomly selected a test card (with the constraint that the same type of card was not presented on more than two consecutive trials), labeled the card by the relevant dimension only, and asked children, "Where does this go?" Children were required to place the card facedown in one of the trays. No feedback was provided; after children sorted each card, the experimenter simply said, "Let's do another one," and then proceeded to the next trial.

When they had completed six trials, children were told to stop playing the first game and to switch to a new game. They were then given six postswitch trials, which were identical to the preswitch trials except that children were told the rules for sorting by the other dimension (e.g., shape: "If it's a boat put it here, but if it's a rabbit put it there."). As in previous research, children were considered to have passed a phase in the standard version when they correctly sorted five or more cards out of six in that phase ($p < .05$, on the basis of the binomial theorem).

To capture developmental changes across the entire age range of interest in this study, children who passed the postswitch phase of the standard version were given a new, more difficult border version (Figure 2). This version commenced immediately after completion of the postswitch phase of the standard version. To explain this version, children were first shown two test cards like those used in the standard version as well as two new test cards that had black

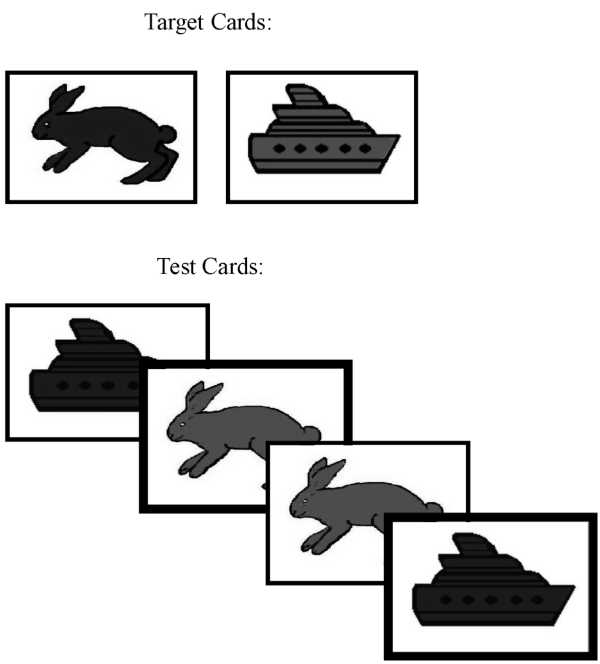

FIGURE 2 The border version. Children are told to sort test cards by one dimension (e.g., color) when there is a black border and by the other dimension (e.g., shape) when there is no border.

borders. Children were then told that the black borders indicated that they must play a particular game (e.g., “If there’s a black border, you have to play the color game; if there is no black border, you have to play the shape game”). The dimension indicated by the black border was counterbalanced across participants. Test trials then commenced. On each of 12 trials, the experimenter stated the rules, randomly selected a test card (with the constraint that no more than two cards in a row were of the same type—with borders or without), labeled it according to the relevant dimension, and asked children, “Where does this go?” As was the case for the standard version, no feedback was provided. Children were considered to have passed this phase when they correctly sorted on 9 out of these 12 trials ($p < .05$, on the basis of the binomial theorem).

Measures of Hot EF

Children’s Gambling Task. On each of 40 trials, children chose from one of two decks of cards (each card being 18 × 29 cm), a striped deck and a dotted deck. When turned, cards displayed a number of happy and sad faces, corresponding to the number of rewards (candies) won and lost, respectively. Cards in one deck (striped) offered more rewards per trial but were disadvantageous across trials due to occasional large losses; cards in the other deck (dotted) offered fewer rewards per trial but were advantageous overall. More specifically, cards in the advantageous deck always offered one reward (i.e., they showed one happy face) together with losses of nothing or one candy (with a net average of five candies gained per 10 cards). Cards in the disadvantageous deck always offered two rewards together with losses of nothing, four, five, or six candies (with a net average of five candies lost per 10 cards). The order of cards in each deck was fixed and followed the win–loss contingencies used by Kerr and Zelazo (2004).

At the start of the task, children were instructed that they should try to win as many candies as possible and that they could select from whichever deck they wished. They were then given a stake of 10 candies with which to begin the task. During 8 initial demonstration trials, in which the experimenter sampled four cards from each deck, children were told (and shown) that the happy faces on the cards indicated the number of candies won, whereas sad faces indicated the number of candies lost. When a card was turned over, only the happy faces were visible initially, because the sad faces were covered with a sticky note. After the number of candies won was revealed to the child and the candies distributed, the sticky note was removed, and the number of candies lost was revealed (this was done to ensure that children attended to both wins and losses). The rewards were deposited into, and removed from, a graduated cylinder situated in front of the child, at an equal distance from each of the two decks. The task proper consisted of 40 test trials. Forty rather than 50 trials (Kerr & Zelazo, 2004) were used because the task was presented as part of a relatively lengthy battery of tasks. In keeping with evidence that data from the second half of the Iowa Gambling Task provide a more reliable index of performance (Monterosso et al., 2001), the dependent variable was the net score on these trials (i.e., the number of advantageous choices minus number of disadvantageous choices made in the last 20 trials). Thus, higher scores indicated better performance.

Delay of Gratification task. The Delay of Gratification task administered in this study was taken from Prencipe and Zelazo (2005), who adapted the procedure used by C. Thompson et al. (1997). Nine test trial types, created by crossing three types of reward (stickers, pennies, candies) and three types of choice (one now vs. two later, one now vs. four later, one now vs. six later), were each depicted graphically on a separate card (14 × 11 cm). There were also two demonstration trials, one in which the option was one candy now versus one candy later and one option for one candy now versus eight candies later.

The two demonstration trials were presented first. On both trials, the experimenter read the decision aloud and made a choice herself. For each trial, the choice was explained verbally and visually by placing the two reward options in separate piles (i.e., immediate pile vs. delay pile). For the one candy now versus one candy later option, the experimenter chose the immediate reward. For the one candy now versus eight candies later option, she chose the delayed reward.

Nine test trials were then presented, involving all nine trial types presented in a random order. Test trials were presented in the same fashion as demonstration trials. However, on each trial, the experimenter asked, "What do you want to do?" The experimenter provided no feedback regarding the wisdom of children's choices, apart from administering the consequences (i.e., dispensing the rewards). When children chose the immediate option, they were allowed to eat the candy, stick their stickers on a special piece of paper, or put their pennies in a penny box. Delayed rewards were placed in an envelope and set aside. Scores were the number of times that children chose to delay.

General Intellectual Functioning

Verbal mental age. Verbal mental age (VMA) was assessed via the Peabody Picture Vocabulary Test–Third Edition–Revised (PPVT–III–R; Dunn & Dunn, 1997), which was presented in the standard fashion. The PPVT–III–R took between 5 to 20 min to complete, depending on children's performance, with better performing children taking longer due to more words being presented. VMA was obtained by determining the raw score (subtracting the number of errors from the number of items administered) and matching this score to its age equivalent from the norms provided in the testing manual.

Performance mental age. Performance mental age (PMA) was assessed via children's performance on the Bead Memory and Pattern Analysis subtests of the Stanford–Binet Intelligence Scale–Fourth Edition (Thorndike, Hagen, & Sattler, 1986). Scores for each subtest were obtained by subtracting the number of errors from the highest number of items administered, and standard age scores on each task were then derived from the norms in the scoring manual. PMA was estimated by averaging the two standard age scores. The two tasks combined took approximately 15 to 40 min, depending on children's performance.

CBQ. The CBQ consists of 195 statements describing typical child behaviors. Parents rated these behaviors on a 7-point Likert scale, ranging from *extremely untrue of your child* to *extremely true of your child*. Items from the CBQ are grouped into 15 scales, which have been shown to load reliably on three distinct factors–Effortful Control, Surgency/Positive Affect, and Negative Affectivity (Rothbart et al., 1994). Following Rothbart et al., Effortful Control scores were derived from scales of Low-Intensity Pleasure, Inhibitory Control, Perceptual Sensitivity, and Attentional Focusing. Surgency scores were computed from scores on the scales of Impulsivity, Approach, High-Intensity Pleasure, Smiling/Laughter, Shyness (negatively scored), and Activity Level. Finally, Negative Affectivity scores were calculated from scores on the scales of Discomfort, Fear, Anger/Frustration, Soothability (negatively scored), and Sadness.

RESULTS

Preliminary analyses revealed no effect of task order and no interactions between task order and any of the variables of interest (i.e., age and sex). Thus, data were combined across task orders for purposes of subsequent analyses, which first examined age-related changes in performance on measures of EF, and then assessed relations among all measures.

Age-Related Changes on Measures of Cool EF

Self-Ordered Pointing. One 3-year-old boy refused to complete the task, so analyses for this task were based on 97 children. An Age × Sex analysis of variance (ANOVA) on the highest number of items that children were able to remember yielded only a main effect for age, $F(2, 91) = 7.86, p < .001, \eta_p^2 = .15$. Tukey's honestly significant difference (HSD) tests indicated that 5-year-olds performed better than 3-year-olds ($p < .001$), as did 4-year-olds ($p = .02$; see Figure 3). Neither the effect of sex, $F(1, 91) = .018$, *ns*, $\eta_p^2 < .001$, nor the interaction between age and sex, $F(2, 91) = 2.77$, *ns*, $\eta_p^2 = .06$, were significant.

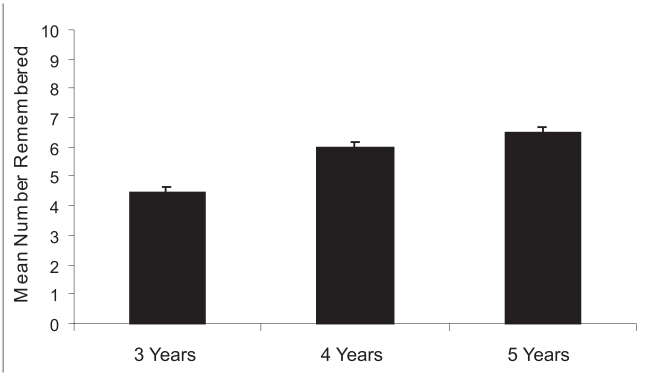

FIGURE 3 Mean (plus 1 standard error) number of items that children were able to remember in the Self-Ordered Pointing task, by age group.

DCCS. One 4-year-old boy indicated that he was deliberately placing the cards in the wrong sorting trays, so his data were excluded and analyses for this task, as with Self-Ordered Pointing, were based on 97 children. Preliminary analyses revealed no effects of dimension order (color or shape first) in the standard version, or of the sorting dimension associated with the border in the border version, so data were collapsed across these variables. Children's performance on the DCCS task was classified into four categories reflecting their highest level of performance. Children who failed the preswitch phase on the standard version were given a score of 0. Those who failed the postswitch phase on the standard version were given a score of 1. Those who passed the postswitch phase of the standard version but failed the border version were given a score of 2. Those who passed the postswitch phase of the standard version as well as the border version were given a score of 3. Due to the categorical nature of these data, a nonparametric (chi-square) analysis was conducted. This analysis demonstrated a significant effect of age across the four categories, $\chi^2(6, N = 97) = 41.88, p < .001, \phi = .66$ (see Figure 4). Further analyses revealed that the following comparisons were significant: 5-year-olds versus 4-year-olds, $\chi^2(2, N = 64) = 12.77, p < .01, \phi = .47$; 4-year-olds versus 3-year-olds, $\chi^2(3, N = 64) = 10.68, p < .05, \phi = .41$; and 5-year-olds versus 3-year-olds, $\chi^2(3, N = 66) = 28.59, p < .001, \phi = .66$. Performance did not differ as a function of sex, $\chi^2(3, N = 97) = 3.11$, *ns*, $\phi = .18$. In all cases, older children outperformed younger ones.

Children's Gambling Task. Five children (three 3-year-old boys, one 3-year-old girl, and one 4-year-old boy) stopped playing before the task was finished, and data from an additional child (a 5-year-old boy) were compromised because a fire alarm interrupted the procedure. Therefore, analyses for this task were

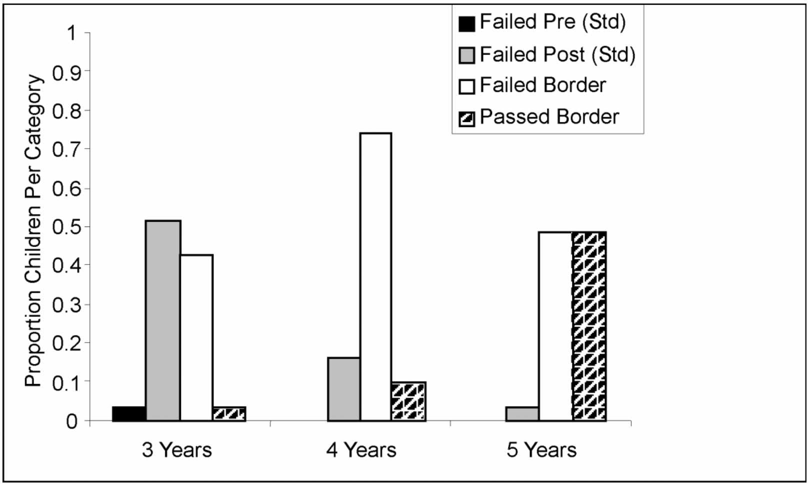

FIGURE 4 Classifications based on children's highest level of performance on the Dimensional Change Card Sort. Percentages of children at each age who were classified as (a) failing the preswitch phase of the standard version, (b) passing the preswitch phase of the standard version but failing the postswitch phase, (c) passing the standard version but failing the border version, and (d) passing both the standard and border versions.

based on 92 children. An Age × Sex ANOVA on children's net scores revealed a main effect of age, $F(2, 86) = 3.74, p < .05, \eta_p^2 = .08$; but no effect of sex, $F(1, 86) = .07$, *ns*, $\eta_p^2 = .001$; and no interaction between age and sex, $F(2, 86) = 2.40, p < .10, \eta_p^2 = .05$. Tukey's HSD test showed that 5-year-olds performed better than 3-year-olds (p = .02; see Figure 5). *T* tests indicated that 3-year-olds performed significantly worse than chance, $t(29) = -2.16, p = .04$, but that neither 4- nor 5-year-olds differed from chance, $t(30) = -.19$, *ns*, and $t(31) = 1.46$, *ns*, respectively.

Delay of Gratification task. All children in the final sample (N = 98) completed this task. Preliminary analyses indicated no significant differences in performance across type of reward (pennies, stickers, candies) and type of choice (one vs. two, one vs. four, one vs. six), so data were collapsed across these variables for further analysis, and scores were calculated out of nine. An Age × Sex ANOVA on the number of delay choices revealed a main effect of age, $F(2, 92) = 11.11, p < .001, \eta_p^2 = .20$. Post-hoc Tukey's HSD tests indicated that 3-year-olds chose to delay less often than 4-year-olds ($p < .005$) and 5-year-olds ($p < .001$), although the latter two groups did not differ (see Figure 6). Neither the effect of sex, $F(1, 92) = 1.65$, *ns*, $\eta_p^2 = .02$, nor the interaction between age and sex, $F(2, 92) = .28$, *ns*, $\eta_p^2 = .006$, were significant. *T* tests indicated that both 4- and 5-year-olds performed significantly above chance, $t(31) = 2.69, p = .01$, and $t(32) = 4.90, p < .001$, respec-

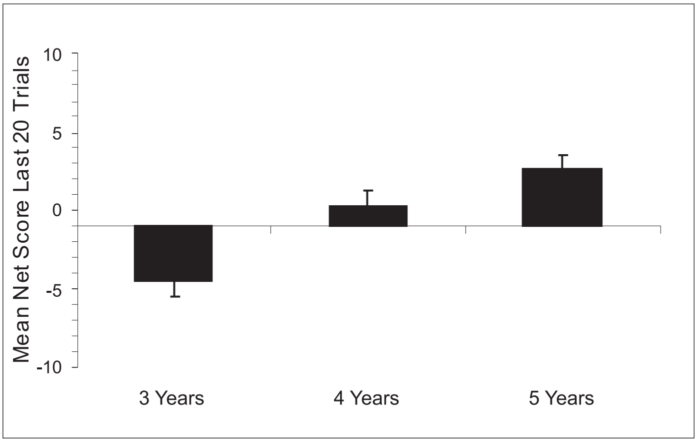

FIGURE 5 Mean (plus 1 standard error) number advantageous minus disadvantageous choices (i.e., net score) on the last 20 trials of the Children's Gambling Task, by age group.

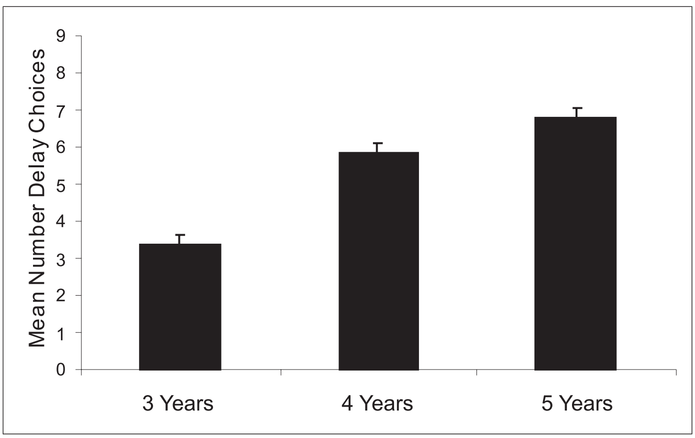

FIGURE 6 Mean (plus 1 standard error) number of times children chose to delay (out of nine) on the modified Delay of Gratification task, by age group.

tively. There was a trend toward 3-year olds performing worse than chance, $t(32) = -1.85$, $p = .07$.

VMA and PMA. An Age × Sex multivariate analysis of variance (MANOVA) was conducted on estimates of VMA and PMA. Significant effects of age were obtained for both VMA, $F(2, 92) = 40.78$, $p < .001$, $\eta_p^2 = .47$, and

PMA, $F(2, 92) = 47.98$, $p < .001$, $\eta_p^2 = .51$ (see Table 1). Tukey's HSD tests revealed differences between each of the three age groups ($p < .005$ for all comparisons). Neither the effect of sex, $F(1, 92) = .15$, *ns*, $\eta_p^2 = .002$, for VMA and $F(1, 92) = .09$, *ns*, $\eta_p^2 = .001$, for PMA nor the Age × Sex interaction, $F(2, 92) = 1.17$, *ns*, $\eta_p^2 = .025$, for VMA, and $F(2, 92) = .08$, *ns*, $\eta_p^2 = .002$, for PMA were significant.

CBQ. Parents of six children (one boy from each age group, one 4-year-old girl, and two 5-year-old girls) failed to complete the CBQ, leaving data from 92 children for analysis. Parent ratings on the CBQ were used to derive scores for the three temperament factors of Effortful Control, Surgency, and Negative Affectivity. An Age × Sex MANOVA on all three scores indicated effects of age, $F(2, 86) = .59$, $p = .001$, $\eta_p^2 = .14$, and sex, $F(1, 86) = 19.64$, $p < .001$, $\eta_p^2 = .15$, that were limited to Effortful Control (see Table 1). There was no interaction between these variables. More specifically, girls scored higher than boys, and a Tukey's HSD test showed that 5-year-olds scored higher than 3-year-olds on this measure ($p = .001$). No other effects were detected.

TABLE 1
Verbal Mental Age (VMA), Performance Mental Age (PMA), and Children's Behavior Questionnaire (CBQ) Factors (Effortful Control, Surgency, and Negative Affectivity)

Variable	*M*	*SD*	*Range*
VMA (months)			
3 years	48.90	19.11	21.6–86.4
4 years	65.88	12.84	49.2–88.8
5 years	82.44	12.39	44.4–105.6
PMA (months)			
3 years	44.76	7.32	36.0–60.0
4 years	55.63	14.10	45.6–84.0
5 years	74.02	13.69	51.6–98.4
CBQ–Effortful Control			
3 years	19.05	2.57	14.97–24.62
4 years	19.89	2.22	14.70–23.43
5 years	21.01	1.84	17.22–25.43
CBQ–Surgency			
3 years	22.58	3.65	12.92–28.54
4 years	20.52	3.81	13.31–26.90
5 years	21.22	2.91	14.85–26.62
CBQ–Negative Affectivity			
3 years	10.96	2.76	6.69–18.31
4 years	12.10	3.83	5.14–20.24
5 years	12.49	2.94	6.31–18.08

Relations Among Measures

Pearson correlations were conducted to assess whether performance on the measures of hot and cool EF were related to each other and to individual differences in general intellectual functioning and temperament. Due to missing data (see previous sections), correlations were based on total sample sizes ranging from 86 to 98.

Relations among measures of EF. As shown in Table 2, there was a highly significant positive relation between performance on the two measures of cool EF—the DCCS and Self-Ordered Pointing task. This relation remained significant even when chronological age (CA) in months was partialled out and even when both CA and mental age (MA; i.e., both VMA and PMA) were partialled out. Simple correlations failed to reveal a relation between the two measures of hot EF, but after partialling, a significant negative relation emerged. The only other relation to remain significant after partialling out CA (and CA and MA) was a positive relation between performance on the Self-Ordered Pointing task and performance on the Children's Gambling Task.

Relations between EF and measures of general intellectual function and temperament. As can be seen in Table 3, performance on both of the two measures of cool EF was highly positively related to both VMA and PMA. Relations with VMA remained significant even when CA was partialled out, as did the relation between the DCCS and PMA. No other relations with VMA or PMA were significant after partialling out CA.

Temperament factors (as assessed by parent ratings on the CBQ) were also related to the measures of cool EF, whereas no correlation was found with measures of hot EF (see Table 3). Among the three factors, significant relations were found for both Effortful Control and Surgency. Effortful Control (which is related to age; see previously) was significantly positively related to both Self-Ordered Pointing and the DCCS, although both relations become nonsignificant after controlling for

TABLE 2
Simple and Partial Correlations Among Measures of Executive Function

Variable	*DCCS*	*CGT*	*DoG*
SOP	.50*** (.34**) [.28**]	.33** (.25*) [.29**]	–.01 (–.19) [–.17]
DCCS		.07 (–.15) [–.09]	.21* (–.07) [.00]
CGT			–.07 (–.22*) [–.24*]

Note. Chronological-age-partialled correlations appear in parentheses. Chronological-age- and mental-age-partialled correlations appear in brackets. SOP = Self-Ordered Pointing; DCCS = Dimensional Change Card Sort; CGT = Children's Gambling Task; DoG = Delay of Gratification.

*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 3
Simple (and Chronological-Age-Partialled) Correlations Among Measures of Executive Function, Measures of General Intellectual Functioning, and Measures of Temperament

Variable	*SOP*	*DCCS*	*CGT*	*DoG*
VMA	.41*** (.28***)	.70*** (.44***)	.20 (–.04)	.18 (–.15)
PMA	.37** (.15)	.69*** (.46***)	.08 (–.20)	.23* (–.10)
Effortful Control	.27** (.17)	.21* (–.01)	.14 (.05)	.12 (–.01)
Surgency	–.25* (–.23*)	–.07 (–.01)	–.01 (.01)	–.08 (–.05)
Negative Affect	.10 (.04)	.14 (.05)	.09 (.06)	.06 (–.01)

Note. SOP = Self-Ordered Pointing; DCCS = Dimensional Change Card Sort; CGT = Children's Gambling Task; DoG = Delay of Gratification; VMA = verbal mental age; PMA = performance mental age.

*$p < .05$. **$p < .01$. ***$p < .001$.

CA. Surgency was significantly negatively related to performance on the Self-Ordered Pointing task, and this relation remained significant after controlling for CA.

DISCUSSION

Assessments of EF in children have relied almost exclusively on the cool, cognitive aspects of EF associated with DL–PFC. Much less is known about the hot, affective aspects associated with VM–PFC, but interest in the topic is growing. Our study investigated the development of both hot and cool EF in the same sample of children to assess whether these aspects of EF were related and whether they showed similar patterns of relations with measures of general intellectual functioning and temperament.

Age-related improvements in performance were found for all four measures of EF, consistent with the notion that both hot and cool aspects of EF develop rapidly during the preschool years (e.g., Zelazo & Müller, 2002)—a period during which PFC as a whole undergoes considerable growth (e.g., see Diamond, 2002; Fuster, 2002, for reviews). In particular, reliable age-related improvements were seen on the Self-Ordered Pointing task, the DCCS, the Children's Gambling Task, and the Delay of Gratification task. This version of the Self-Ordered Pointing task had never been used with preschoolers, and the DCCS included a new, more difficult border version that proved sensitive to developmental changes between 4 and 5 years of age.

Despite the finding that all four measures were developmentally sensitive within this age range, different patterns of relations were found among measures of EF. The two measures of cool EF were significantly positively related to one another, even after controlling for CA and MA. This finding suggests that performance on these tasks relies on common processes and indicates that although cool EF was related to MA, it is not synonymous with general intelligence as measured by standardized tests. Surprisingly, the two hot EF tasks (the Children's Gambling Task and the Delay of Gratification task) correlated with one another in a negative direction, suggesting that they may differ in important ways and that the construct of hot EF may need to be further refined. This finding with preschool-age children differs from previous findings with cocaine-dependent adults, showing that the Iowa Gambling Task is positively related to delay discounting (e.g., Monteresso et al., 2001).

There are several fundamental differences between the two measures of hot EF that may account for the lack of a positive relation between them. In contrast to the Children's Gambling Task, which requires learning to avoid disadvantageous decks across a series of 40 trials, the Delay of Gratification task involves a smaller series of nine independent choices. Thus, although both tasks require consideration of affectively significant future consequences (relative pros and cons associated with each of two options), only the Children's Gambling Task requires children to track the wins and losses associated with each deck across trials. Tracking wins and losses may require working memory, consistent with the finding that the Children's Gambling Task was related to the Self-Ordered Pointing task in this study and with recent evidence from adults suggesting that working memory may play a role in the Iowa Gambling Task. For example, Manes et al. (2002) found that patients with DL–PFC damage performed worse than a control group on the Iowa Gambling Task. More direct evidence of the role of working memory on the Iowa Gambling Task comes from Hinson, Jameson, and Whitney (2002), who found that performance declined when participants' working memory was taxed (via dual task conditions). It is possible, therefore, that the Children's Gambling Task and the Iowa Gambling Task entail *both* the flexible appraisal of emotional stimuli (a VM–PFC function) *and* working memory (a DL–PFC function) to track the feedback provided by previous selections and, hence, that neither task provides a "pure" measure of hot EF.

It is also important to note, however, that other studies have failed to find a relation between working memory and performance on the Iowa Gambling Task or variants of this task (see Crone & van der Molen, 2004; Hooper, Luciana, Conklin, & Yarger, 2004; Wilder, Weinberger, & Goldberg, 1998). Resolving this issue will be important in understanding the processes underlying affective decision making. For example, Schoenbaum and Setlow (2001) suggested that orbitofrontal cortex (OFC; a subregion of VM–PFC) may support its own working memory function. That is, although DL–PFC-mediated working memory may involve the storage and manipulation of objects and places (absent affective content), OFC-mediated working memory may involve the storage and manipulation of the incentive value of objects. However, much of Schoenbaum and Setlow's evidence was based on relatively simple tasks with nonhuman animals, and it is still unclear whether OFC plays a role in working memory in human beings. Further work is required to refine the construct of working memory and explore its relation to different aspects of EF at different points in development.

Another difference between Delay of Gratification and the Children's Gambling Task is the extent to which one can be certain that prudent choices will yield subsequent rewards. That is, whereas choice contingencies are clear in the Delay of Gratification task (e.g., one now vs. two later), the Children's Gambling Task, like the Iowa Gambling Task, was designed to make the contingencies associated with each choice uncertain. According to Bechara (2004), VM–PFC is most likely to be implicated in decisions under uncertainty. Once again, further investigation is required to explore the possibility that hot EF, although perhaps distinct from cool EF, may itself be a heterogeneous construct.

The measures of hot and cool EF also showed different patterns of relations with general intellectual functioning. Strong positive (age-partialled) relations were found between VMA and the two measures of cool EF and between PMA and performance on the DCCS. These findings suggest that there may be substantial overlap between the construct of cool EF and the construct of intelligence (as conceptualized in standardized tests; e.g., Duncan et al., 2000), and they are consistent with previous research using the DCCS (e.g., Bialystok, 1999) and tests of working memory (e.g., Carlson, Moses, & Breton, 2002). It is noted that the Bead Memory subscale used to assess PMA is a measure of short-term memory (involving passive storage of content) rather than working memory (involving both storage and manipulation of content), which may partly explain why PMA was not more highly related to the Self-Ordered Pointing task than to the DCCS.

The finding that measures of hot EF were not related to general intellectual functioning is consistent with the suggestion that hot EF may be related to a different form of intelligence than that assessed in this study. More specifically, in contrast to standard measures of general intellectual function, hot EF, with its more emotional nature, may be more related to emotional intelligence (e.g., Goleman, 1995; Salovey & Mayer, 1990), emotional and social intelligence (Bar-On, 2000), or personal intelligence (Gardner, 1983). Consistent with this possibility, Rosso, Young, Femia, and Yurgelun-Todd (2004) found that a measure of emotional intelligence was unrelated to abstract reasoning (matrix reasoning) in 9- to 18-year-olds. It is of interest that these authors found significant age differences on matrix reasoning but not emotional intelligence, which they interpret in terms of "differential developmental trajectories across various cognitive and emotional domains of frontal lobe functioning" (Rosso et al., 2004, p. 355). Indeed, whereas DL–PFC may be more closely linked to standard measures of intelligence (Duncan et al., 2000), VM–PFC and related structures may play a greater role in emotional intelligence. This suggestion is supported by a recent study showing that performance on the Iowa Gambling Task was positively related to scores on a measure of emotional and social intelligence but not to scores on the Wechsler Adult Intelligence Scale (Bar-On, Tranel, Denburg, & Bechara, 2003). From this perspective, one may or may not expect some aspects of temperament to be related to performance on hot EF tasks; the relation between temperament and emotional intelligence in childhood is unclear.

In fact, although cool EF was related to temperament, hot EF was not. First, a negative relation between Surgency and Self-Ordered Pointing (even after controlling for age) was found and may reflect the fact that Surgency has an impulsive component (Davis, Bruce, & Gunnar, 2002). For example, in this working memory task, it was necessary to avoid selecting the same picture twice despite the fact that having responded to a particular picture once may have primed responding to that picture again in the future (repetition priming).

The two measures of cool EF were positively related to Effortful Control before controlling for age, although the relations became nonsignificant after controlling for age. The relations were expected based on previous research and theory (e.g., Rothbart et al., 1994; Ruff & Rothbart, 1996; Wolfe & Bell, 2003), and given that there was a significant effect of age on Effortful Control, it is perhaps not surprising that these relations were attenuated when age was controlled. Parent ratings of Effortful Control may well be based in large part on parents' observations of children's developing cool EF.

In contrast, neither measure of hot EF was related to Effortful Control or any other temperament factor. Effortful Control is conceptually and empirically related to EF, but previous research has focused mainly on cool EF, and the developmental correlates of hot EF remain largely unknown. Although some previous work has considered delay of gratification in relation to Effortful Control (e.g., Davis et al., 2002; Kochanska, Murray, & Harlan, 2000), these studies used delay versions of delay of gratification that differ in important respects from the choice version used in our study (i.e., in delay versions, one measures how long children wait before engaging in a desired activity). Overall, the finding that hot EF was unrelated to temperament is consistent with the suggestion that hot and cool EF reflect distinct aspects of EF.

CONCLUSION

Although researchers traditionally have treated EF as a domain-general cognitive function and have focused almost exclusively on cool, cognitive aspects of EF, the findings from this study support a distinction between hot and cool aspects of EF, and they underscore the need for more comprehensive assessments of EF that include measures of hot EF. The findings also suggest that whereas cool EF seems to be a reasonably coherent functional construct, hot EF is not. Much less is known about hot EF than cool EF, and further research is clearly warranted. This research will fill an important gap in our understanding of EF and its development, and it has the potential to shed light on the complex interactions between emotion and cognition. In addition, it may have practical implications for children with disorders such as autism and conduct disorder, where social and emotional deficits are prominent.

ACKNOWLEDGMENT

This research was supported by grants from the Natural Sciences and Engineering Research Council and the Canada Research Chairs Program to Philip David Zelazo.

REFERENCES

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience, 2*, 1032–1037.

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology, 5*, 115–129.

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology, 15*, 31–36.

Arffa, S., Lovell, M., Podell, K., & Goldberg, E. (1998). Wisconsin Card Sorting Test performance in above average and superior school children: Relationship to intelligence and age. *Archives of Clinical Neuropsychology, 13*, 713–720.

Baddeley, A. (1986). *Working memory*. New York: Clarendon Press/Oxford University Press.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65–94.

Bar-On, R. (2000). Emotional and social intelligence: Insights from the Emotional Quotient Inventory. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 363–388). San Francisco: Jossey-Bass.

Bar-On, R., Tranel, D., Denburg, N. L., & Bechara, A. (2003). Exploring the neurological substrate of emotional and social intelligence. *Brain, 126*, 1790–1800.

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition, 55*, 30–40.

Bechara, A., Damasio, A., Damasio, H., & Anderson, S. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50,* 7–15.

Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision-making within the human prefrontal cortex. *The Journal of Neuroscience, 18*, 428–437.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain, 123*, 2189–2202.

Bechara, A., Tranel, D., Damasio, H., & Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex, 6*, 215–225.

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

Bialystok, E., & Martin, M. (2003). Notation to symbol: Development in children's understanding of print. *Journal of Experimental Child Psychology, 86*, 223–243.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*, 111–127.

Carlson, S. M., Davis, A. C., & Leach, J. G. (in press). Less is more: Executive function and symbolic representation in preschool children. *Psychological Science*.

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

Carlson, S. M., Moses, l. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73–92.

Clark, L., Cools, R., & Robbins, T. W. (2004). The neuropsychology of ventral prefrontal cortex: Decision-making and reversal learning. *Brain and Cognition, 55*, 41–53.

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, 25*, 251–279.

Curtis, C. E., Zald, D. H., & Pardo, J. V. (2000). Organization of working memory within the human prefrontal cortex: A PET study of self-ordered object working memory. *Neuropsychologia, 38*, 1503–1510.

Damasio, A. R. (1994). *Descartes' error: Emotion, rationality, and the human brain*. New York: Putnam.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review, 3,* 422–433.

Davis, E. P., Bruce, J., & Gunnar, M. R. (2002). The anterior attention network: Associations with temperament and neuroendocrine activity in 6-year-old children. *Developmental Psychobiology, 40*, 43–56.

Dawson, G., Meltzoff, A. N., Osterling, J., & Rinaldi, J. (1998). Neuropsychological correlates of early symptoms of autism. *Child Development, 69*, 1276–1285.

Deák , G. O. (2000). The growth of flexible problem solving: Preschool children use changing verbal cues to infer multiple word meanings. *Journal of Cognition and Development,1*, 157–191.

Demakis, G. J. (2003). A meta-analytic review of the sensitivity of the Wisconsin Card Sorting Test to frontal and lateralized frontal brain damage. *Neuropsychology, 17*, 255–264.

Denckla, M. B., & Reiss, A. L. (1997). Prefrontal-subcortical circuits in developmental disorders. In N. A. Krasnegor, G. R. Lyon, & P. S. Goldman-Rakic (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 283–293). Baltimore: Brookes.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London: Oxford University Press.

Diamond, A., Prevor, M., Callender, G., & Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development, 62*(4, Whole No. 252).

Dias, R., Robbins, T. W., & Roberts, A. (1996, March 7). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380*, 69–72.

Dinn, W. M., Robbins, N. C., & Harris, C. L. (2001). Adult attention-deficit/hyperactive disorder: Neuropsychological correlates and clinical representation. *Brain and Cognition, 46*, 114–121.

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000, July 21). A neural basis for general intelligence. *Science, 289*, 457–460.

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

Elliott, R., Frith, C. D., & Dolan, R. J. (1997). Differential neural response to positive and negative feedback in planning and guessing tasks. *Neuropsychologia, 35*, 1395–1404.

Epsy, K. A., Kaufman, P. M., Glisky, M. L., & McDiarmid, M. (2001). New procedures to assess executive functions in preschool children. *Clinical Neuropsychologist, 15,* 46–58.

Frye, D., Zelazo, P. D., Brooks, P. J., & Samuels, M. C. (1996). Inference and action in early causal reasoning. *Developmental Psychology, 32*, 120–131.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*, 483–527.

Fuster, J. M. (2002). Frontal lobe and cognitive development. *Journal of Neurocytology, 31*, 373–385.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Garon, N., & Moore, C. (2004). Complex decision-making in early childhood. *Brain and Cognition, 55,* 158–170.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 39*, 3–27.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdendos, A., et al. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2*, 861–863.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004, May 25). Dynamic mapping of human cortical development during childhood through early adulthood. *Procedings of the National Academy of Science, USA*, 101, 8174–8179.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses to Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38,* 404–411.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6,* 207–208.

Green, L., Myerson, J., & Ostaszewski, P. (1999). Discounting of delayed rewards across the lifespan: age differences in individual discounting functions. *Behavioural Processes, 46*, 89–96.

Happaney, K. R., Zelazo, P. D., & Stuss, D. T. (2004). Development of orbitofrontal function: Current themes and future directions. *Brain and Cognition, 55*, 1–10.

Hauser, M. D. (1999). Perseveration, inhibition and the prefrontal cortex: A new look. *Current Opinion in Neurobiology, 9*, 214–222.

Hinson, J. M., Jameson, T. L., & Whitney, P. (2002). Somatic markers, working memory, and decision-making. *Cognitive, Affective, & Behavioral Neuroscience, 2*, 341–353.

Hooper, C. J., Luciana, M., Conklin, H. M., & Yarger, R. S. (2004). Adolescents' performance on the Iowa Gambling Task: Implications for the development of decision-making and ventromedial prefrontal cortex. *Developmental Psychology, 40*, 1148–1158.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16,* 233–253.

Hughes, C., Dunn, J., & White, A. (1998). Trick or treat?: Uneven understanding of mind and emotion and executive dysfunction in "Hard-to-manage" preschoolers. *Journal of Child Psychology and Psychiatry, 39*, 981–994.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia, 32*, 477–492.

Jacobson, C. F. (1936). Studies of cerebral functions in primates: I. The functions of the frontal association areas in monkeys. *Comparative Psychology Monographs, 13*, 1–30.

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

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*, 637–671.

Kerr, A., & Zelazo, P. D. (2004). Development of "hot" executive function: The Children's Gambling Task. *Brain and Cognition, 55,* 148–157.

Kheramin, S., Brody, S., Ho, M.-Y., Velazquez-Martinez, D. N., Bradshaw, C. M., Szabadi, E., et al. (2003). Role of the orbitofrontal cortex in choice between delayed and uncertain reinforcers: A quantitative analysis. *Behavioural Processes, 64*, 239–250.

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

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in

early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology, 36*, 220–232.

Krawczyk, D. C. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neuroscience and Biobehavioral Reviews, 26*, 631–664.

Lombardi, W. J., Andreason, P. J., Sirocco, K. Y., Rio, D. E., Gross, R. E., Umhau, J. C., et al. (1999). Wisconsin Card Sorting Test performance following head injury: Dorsolateral fronto-striatal circuit activity predicts perseveration. *Journal of Clinical & Experimental Neuropsychology, 21*, 2–16.

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight- year old children. *Neuropsychologia, 36*, 273–293.

Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12- year old children. *Developmental Neuropsychology, 22*, 595–624.

Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., et al. (2002). Decision-making processes following damage to prefrontal cortex. *Brain, 125,* 624–639.

Matsuzawa, J., Matsui, M., Konishi, T., Noguchi, K., Gur, R. C., Bilker, W., et al. (2001). Age-related changes in brain grey and white matter in healthy infants and children. *Cerebral Cortex, 11*, 335–342.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review, 106*, 3–19.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K Akert (Eds.), *The frontal granular cortex and behavior* (pp. 313–334). New York: McGraw-Hill.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989, May 26). Delay of gratification in children. *Science, 244*, 933–938.

Mobini, S., Body, S., Ho, M.-Y., Bradshaw, C. M., Szabadi, E., Deakin, J. F. W., et al. (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology, 160*, 290–298.

Monterosso, J., Ehrman, R., Napier, K. L., O'Brien, C. P., & Childress, A. R. (2001). Three decision-making tasks in cocaine-dependent patients: Do they measure the same construct? *Addiction, 96,* 1825–1837.

Munakata, Y., & Yerys, B. E. (2001). All together now: When dissociations between knowledge and action disappear. *Psychological Science, 12*, 335–337.

Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Oyanagi, C., et al. (2001). Dissociable mechanisms of attentional control within the human prefrontal cortex. *Cerebral Cortex, 11*, 85–92.

Orzhekhovskaya, N. S. (1981). Fronto-striatal relationships in primate ontogeny. *Neuroscience & Behavioral Physiology, 11,* 379–385.

O'Sullivan, L. P., Mitchell, L. L., & Daehler, M. W. (2001). Representation and perseveration: Influences on young children's representational insight. *Journal of Cognition and Development, 2*, 339–365.

Pennington, B. F. (1997). Dimensions of executive functions in normal and abnormal development. In N. A. Krasnegor, G. R. Lyon, & P. S. Goldman-Rakic (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 265–281). Baltimore: Brookes.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 37*, 51–87.

Pérez-Edgar, K., & Fox, N. A. (in press). A behavioral and electrophysiological study of children's selective attention under neutral and affective conditions. *Journal of Cognition and Development*.

Perner, J., & Lang, B. (1999). Development of theory of mind and cognitive control. *Trends in Cognitive Science, 3*, 337–344.

Perner, J., Stummer, S., & Lang, B. (1999). Executive functions and theory of mind: Cognitive complexity or functional dependence? In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 133–152). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20*, 249–262.

Pfefferbaum, A., Mathalon, D. H., Sullivan, E. V., Rawles J. M., Zipursky, R. B., & Lim, K. O. (1994). A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Archives of Neurology, 51,* 874–887.

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,* 501–505.

Rahman, S., Sahakian, B. J., Cardinal, R. N., Rogers, R. D., & Robbins, T. W. (2001). Decision making and neuropsychiatry. *Trends in Cognitive Sciences, 5*, 271–277.

Rogers, R. D., Andrews, T. C., Grasby, P. M., Brooks, D. J., & Robbins, T. W. (2000). Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *Journal of Cognitive Neuroscience, 12*, 142–162.

Rogers, R. D., Everitt, B. J., Baldacchino, A., Blackshaw, A. J., Swainson, R., Wynne, K., et al. (1999). Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: Evidence for monoaminergic mechanisms. *Neuropsychopharmacology, 20*, 322–339.

Rogers, R. D., Owen, A. M., Middleton, H. C., Williams, E. J., Pickard, J. D., Sahakian, B. J., et al. (1999). Choosing from small, likely rewards and large, unlikely rewards activated inferior and orbital prefrontal cortex. *The Journal of Neuroscience, 20*, 9029–9038.

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain and Cognition, 55,* 11–29.

Rosso, I. M., Young, A. D., Femia, L. A., & Yurgelun-Todd, D. A. (2004). Cognitive and emotional components of frontal lobe functioning in childhood and adolescence. In R. E. Dahl & L. P. Spear (Eds.), *Adolescent brain development: Vulnerabilities and opportunities* (pp. 355–362). New York: New York Academy of Sciences.

Rothbart, M. K., Ahadi, S. A., & Hershey, K. L. (1994). Temperament and social behavior in childhood. *Merrill Palmer Quarterly, 40*, 21–39.

Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development: themes and variations*. London: Oxford University Press.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9,* 185–211.

Schoenbaum, G., & Setlow, B. (2001). Integrating orbitofrontal cortex into prefrontal theory: Common processing themes across species and subdivisions. *Learning and Memory, 8*, 134–147.

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*, 620–625.

Sohn, M. H., Ursu, S., Anderson, J. R., Stenger, V. A., & Carter, C. S. (2000). The role of prefrontal cortex and posterior parietal cortex in task switching. *Proceedings of the National Academy of Sciences, USA, 97*, 13448–13453.

Thompson, C., Barresi, J., & Moore, C. (1997). The development of future-oriented prudence and altruism in preschoolers. *Cognitive Development, 12*, 199–212.

Thompson, P. M., Giedd, J. N., Woods, R. P., MacDonald, D., Evans, A. C., & Toga, A. W. (2000, March 9). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature, 404*, 190–193.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale, Fourth edition* [Technical manual]. Chicago: Riverside.

Toner, I. J., Holstein, R. B., & Hetherington, E. M. (1977). Reflection-impulsitivity and self-control in preschool children. *Child Development, 48*, 239–245.

Tsujimoto, S., Yamamoto, T., Kawaguchi, H., Koizumi, H., & Sawaguchi, T. (2004). Prefrontal cortical activation associated with working memory in adults and preschool children: an event-related optical topography study. *Cerebral Cortex, 14*, 703–712.

Wang, L., Kakigi, R., & Hoshiyama, M. (2001). Neural activities during Wisconsin Card Sorting Test–MEG observation. *Cognitive Brain Research, 12*, 19–31.

Wilder, K. E., Weinberger, D. R., & Goldberg, T. E. (1998). Operant conditioning and the orbitofrontal cortex in schizophrenic patients: Unexpected evidence for intact functioning. *Schizophrenia Research, 30*, 169–174.

Wolfe, C. D., & Bell, M. A. (2003). Working memory and inhibitory control in early childhood: Contributions from physiology, temperament, and language. *Developmental Psychobiology, 44*, 68–83.

Zald, D. H., Mattson, D. L., & Pardo, J. V. (2002). Brain activity in ventromedial prefrontal cortex correlates with individual differences in negative affect. *Proceedings of the National Academy of Sciences, USA, 99*, 2450–2454.

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1*, 198–226.

Zelazo, P. D., Frye, D. & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37–63.

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.

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