Available online at www.sciencedirect.com
SCIENCE DIRECT®
Brain and Cognition 52 (2003) 33–51
Brain and Cognition
www.elsevier.com/locate/b&c

Individual differences in children's performance during an emotional Stroop task: A behavioral and electrophysiological study

Koraly Pérez-Edgar and Nathan A. Fox*

Department of Human Development, Child Development Laboratory, University of Maryland, 3304 Benjamin Building, College Park, MD 20742, USA

Accepted 26 September 2002

Abstract

Two studies using the emotional Stroop with 11-year-old children were completed. In Study 1, children were assigned to either the "interference group" or the "facilitation group" based on their performance on the task. The interference group was slower to respond to emotion words (positive and negative) versus control words. The facilitation group was faster to respond to the emotion words. The groups were then compared on a set of cognitive, emotional, and social measures collected at ages 4, 7, and 11. The interference group showed greater signs of emotional and social, but not cognitive, maladjustment across time. Study 2 replicated the findings of Study 1. In addition, event-related potentials (ERPs) were collected in Study 2. The ERP data replicated earlier traditional Stroop studies. In addition, positive and negative words showed differences in processing across components. In particular, negative words appeared to tax attentional and processing resources more than positive words.

1. Introduction

The notion that emotion (e.g., the expression of negative affect) and cognition (e.g., the use of attentional skills for the regulation of negative affect) are functionally interdependent is supported at both the behavioral and neuroanatomical level. Traditionally, observed behavioral differences in emotional reactivity have been localized in the limbic system, particularly the amygdala, while executive function or self-regulatory processes have primarily been localized in anterior cortical areas, particularly the prefrontal cortex and the anterior cingulate (Diamond, 1990). However, recent work has detailed the extensive neuroanatomical (Alexander, Crutcher, & De Long, 1990; Masterman & Cummings, 1997) and neurofunctional (Drevets & Raichle, 1998) connections between these affective and regulatory structures, including direct connections between the amygdala and the anterior cingulate.

Given these findings, it seems plausible that tasks that effectively tap both cognitive and emotion systems may prove useful in exploring individual differences in the processing and regulation of affect. Previous findings from studies of infant and childhood temperament (Fox, Henderson, Rubin, Calkins, & Schmidt, 2001) suggest that this may be of interest in studies of negative affect. When faced with stimuli or situations that elicit negative affect, individuals appear to rely heavily on regulatory processes, particularly those involved in attention (Rothbart, Posner, & Hershey, 1995). The current paper employs this methodological strategy in two related studies of 11-year-old children. Each study has the emotional Stroop as its central cognitive-emotional task. An individual differences approach was used to see if performance in the emotional Stroop could be linked to distinguishable developmental patterns in cognition and emotion, at both the behavioral and psychophysiological level. Each component of this approach will be discussed in turn.

In over 60 years of experimental use, the Stroop effect (Stroop, 1935) has proven to be remarkably robust, withstanding variation in presentation mode (Holle, Neely, & Heimberg, 1997; Shimada, 1990), response mode (Ilan & Polich, 1999), and stimuli content (Bush et al., 1998; Whalen et al., 1998). The traditional Stroop task presents individuals with a series of words and asks them to name the color in which the word is written

* Corresponding author. Fax: 1-301-405-2891.
E-mail address: nf4@umail.umd.edu (N.A. Fox).

0278-2626/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0278-2626(03)00007-1

while disregarding the actual meaning of the word. Individuals are faster to respond when presented with congruent stimuli (the word RED in red ink) than when the stimuli are incongruent (the word RED in blue ink). The emotional Stroop substitutes emotionally charged words for the traditional color words.

In his extensive review of the emotional Stroop literature, Williams (Williams, Mathews, & MacLeod, 1996) found that individuals across a wide range of clinical populations are slow in responding to stimuli idiosyncratic to their disorder. For example, a spider-phobic individual will show slowed responses to the words "crawl" and "hairy," but not the words "dizzy" or "sweating" (Watts, McKenna, Sharrock, & Trezise, 1986). The opposite is true for an individual with panic disorder (Mathews & MacLeod, 1985).

The emotional Stroop has been used with individuals diagnosed with general anxiety disorder (Mathews & MacLeod, 1985), panic attacks (Hope, Rapee, Heimberg, & Dombeck, 1990; McNally et al., 1994), phobias (Hope et al., 1990; Watts et al., 1986), obsessive-compulsive disorder (Foa, Ilai, McCarthy, Shoyer, & Murdock, 1993; Lavy, van Oppen, & van den Hout, 1994), morbid jealousy (Intili & Tarrier, 1998), post-traumatic stress disorder (Harvey, Bryant, & Rapee, 1996; McNally, English, & Lipke, 1993), eating disorders (Lovell, Williams, & Hill, 1997), and depression (Gotlib & McCann, 1984). The Stroop effect has been shown to track the participant's degree of impairment (McNally, 1995) and amenability to treatment (Mattia, Heimberg, & Hope, 1993). It has also been successfully used with non-clinical populations when stimuli are derived from individual interviews (Logan & Goetsch, 1993; Riemann & McNally, 1995) or when stimuli match experimentally induced moods (Gilboa-Schechtman, Revelle, & Gotlib, 2000; Richards, French, Johnson, Naparstek, & Williams, 1992; but not Gotlib & McCann, 1984; Riemann & McNally, 1995).

While most research has focused on the interference effects produced by the traditional and emotional Stroop, a close inspection of the extant literature reveals that the Stroop task will also frequently produce a corresponding facilitation effect (Bauer & Hesselbrock, 1999; Ilan & Polich, 1999; MacLeod, 1991; Mathews & MacLeod, 1994; van Honk et al., 2000). Facilitation refers to the pattern of faster response times to experimental stimuli (congruent color words in the traditional Stroop; affective words in the emotional Stroop) than to the neutral control words. The presence of facilitation and interference across a group of individuals could serve as a marker for underlying differences in the way individuals process and respond to emotional stimuli.

Indeed, recent research indicates that individual neuroendocrine levels vary as a function of performance in the emotional Stroop. van Honk and colleagues (van Honk et al., 2000) used a pictorial emotional Stroop task employing masked and unmasked angry and neutral faces. They found that individuals who showed interference to angry faces had increased salivary cortisol levels, while individuals who showed facilitation had a corresponding decrease in cortisol levels. This would suggest that individuals who show interference to (negative) affective stimuli view these images as threatening and perhaps fear-inducing. This is in line with data indicating that the emotional Stroop engages the same neural systems, including the amygdala, implicated in emotional reactivity (Isenberg et al., 1999). The amygdala, in particular, is strongly tied to the HPA axis and cortisol production. In addition, van Honk et al. (2000) found that individuals who show interference also experienced a corresponding increase in testosterone levels during task participation. Again, individuals showing facilitation have decreased levels of testosterone. van Honk et al. (2000) argue that the masked condition (where the testosterone findings were concentrated) is a "clean" test of the direct thalamic-amygdala pathway, without the intervention of cortical areas (LeDoux, 1996). As such, they conclude that interference in the emotional Stroop task can be taken as a marker for a biological hardwiring to respond to threat.

At the same time, PET and MRI studies have linked both the traditional and emotional Stroop to a web of neural systems critical for the expression and self-regulation of emotion (van Honk et al., 2000; West & Alain, 2000a). Studies of the emotional Stroop involving both control (George et al., 1994; Whalen et al., 1998) and clinical (Rauch et al., 1994; Rauch et al., 1995; Rauch et al., 1996) populations have shown high levels of activation in the ventral (areas 33 and 25) and rostral (areas 32 and 24) portions of the anterior cingulate (Baker, Frith, & Dolan, 1997). These areas show strong connections to lateral orbital frontal cortex, the limbic striatum, and the amygdala (Devinsky, Morrell, & Vogt, 1995; Vogt, Nimchinsky, Vogt, & Hof, 1995), which are strongly implicated in the processing and expression of emotion (Davidson, 2000).

Analogous studies of the traditional Stroop (Bush et al., 1998; Carter, Mintun, & Cohen, 1995; Pardo, Pardo, Janer, & Raichle, 1990) have shown activation in caudal (areas 24 and 32) portions of the anterior cingulate. These areas have strong connections to the prefrontal, premotor, and supplementary motor areas (Vogt et al., 1995), all of which are linked to executive functioning in children and adults (Diamond, 2000; Goldman-Rakic, 1998; Welsh, Pennington, & Groisser, 1991). These data, coupled with the extensive clinical literature, suggest that the emotional Stroop is particularly well suited for studying the link between cognition (e.g., self-regulation or executive functioning) and emotion (e.g., emotional reactivity).

On a practical note, the relative simplicity of the emotional Stroop makes the task well suited for use in developmental studies. Overall, the findings from these studies are remarkably similar to those from adult studies. First, children diagnosed with spider phobias (Kindt, Bierman, & Brosschot, 1997; Martin, Horder, & Jones, 1992), anxiety disorder (Taghavi, 1996 in Moradi, Taghavi, Neshat-Doost, Yule, & Dagleish, 1999), conduct disorder (Bauer & Hesselbrock, 1999), and PTSD (Moradi et al., 1999) also show diagnosis-specific interference. Indeed, children will show interference to words chosen to reflect the psychological concerns (e.g., phobias, PTSD) of their parents (Moradi, Neshat-Doost, Taghavi, Yule, & Dagleish, 1998; Schneider, Unnewher, Florin, & Margraf, 1992 cited in Moradi et al., 1999). Second, non-clinical children will also show interference to idiographic stimuli. For example, children ranked low in popularity and social acceptance show greater interference to socially negative words than do popular children (Martin & Cole, 2000). This pattern even carries over to studies designed to meet the skill levels of very young children through the use of auditory (Green & Barber, 1983; Jerger, Martin, & Pirozzolo, 1988; McClain, 1983) or pictorial (Gerstadt, Hong, & Diamond, 1994) stimuli. It appears that the Stroop effect in children and adults relies on similar underlying cognitive (and perhaps neural) mechanisms, that these mechanisms are early appearing, and that the effect is broadly independent of the perceptual and response demands placed on the individual.

The current study also included two measures of executive functioning (Welsh et al., 1991): the Tower of Hanoi task and the Wisconsin Card Sorting task. These tasks, thought to measure flexible strategy implementation and planning, have both been linked to prefrontal cortex functioning (Glosser & Goodglass, 1990; Welsh et al., 1991) and were included to indicate if any differences in emotional Stroop performance were unique to the task demands or simply an indication of broader differences in executive functioning ability.

2. Study 1

2.1. Method

2.1.1. Participants

The participants in this study were taken from a cohort of 88 children initially recruited as newborns for a longitudinal study of the behavioral and physiological correlates of temperament (see Fox, Schmidt, Calkins, Rubin, & Coplan, 1996; Stifter & Fox, 1990). Selection criteria included healthy gestational age, birth weight, and Apgar score. The children were primarily Caucasian (80%), recruited from the Washington, DC, metropolitan area, and were from middle-class backgrounds. Presented in this study are the 21 children (8 male) who returned to the laboratory at age 11. An analysis comparing the children who participated at 11 versus those who did not return indicated that the children did not differ on any demographic or behavioral factors. The data presented here were collected at ages 4, 7, and 11.

2.1.2. Procedures

2.1.2.1. Emotional Stroop. At the 11-year visit the children were presented with 45 words, 15 in each word category: positive, negative, and control (see Appendix A). The words were chosen as representatives of broad affective states. Unlike many emotional Stroop studies, the words were not tailored to target particular psychological concerns (e.g., anxiety) or idiographic factors. Post-hoc analyses (Frances & Kucera, 1982) indicated that while positive and control words did not differ in word frequency ($t(28) = 0.82$, $p = .42$) or number of syllables ($t(28) = 0.00$, $p = 1.00$), the negative words were less frequent ($t(28)$'s > 2.72, p's $< .01$) and had more syllables ($t(28)$'s > 2.15, p's $< .04$). However, an item analysis indicated that the mean reaction times in the emotional Stroop task did not correlate significantly with either word frequency ($r(45) = -.003$, $p = .99$) or with number of syllables ($r(45) = .16$, $p = .28$).

Stimuli were presented on a NANAO FlexScan 550i monitor in red, green, or blue. Word presentation (ITI = 1000 ms; time-out latency = 3500 ms) was controlled by the STIM stimulus presentation system from the James Long Company (Caroga Lake, NY). The children were asked to state the color in which the word was written, while disregarding the content of the word. Reaction times were collected using a voice-activated microphone connected to the data acquisition computer. The microphone was placed directly in front of the child at a distance of 6 in. An experimenter manually noted any errors in color naming.

2.1.2.2. Maternal ratings of adjustment and temperament. *CBCL.* As a measure of adjustment/maladjustment, mothers completed the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983) at ages 4, 7, and 11. The CBCL is a 113-item checklist in which parents use a three-point scale to rate how descriptive a series of behavior problems are of their own child. The CBCL yields a number of narrow-band factors (withdrawal, anxious/depressed, social problems, attention problems, delinquency, aggressive behaviors) that are further reduced to two broadband factors (internalizing and externalizing behavior problems). In addition, it provides an index of the overall level of difficulty a child may be experiencing through a total problems score. Additional information concerning the reliability and validity of the CBCL can be found in Achenbach (1991).

CCTI. At ages 4 and 7, mothers completed the Colorado Child Temperament Inventory (CCTI; Buss & Plomin, 1984; Rowe & Plomin, 1977). This 30-item measure yields six scales pertaining to different dimensions of child temperament. These include emotionality, activity, attention, soothability, shyness, and sociability. Two additional scales, "emotion dysregulation" (i.e., emotionality *minus* soothability) and "impulsivity" (i.e., emotionality plus activity), were computed (see Coplan, Rubin, Fox, Calkins, & Stewart, 1994; Rubin, Coplan, Fox, & Calkins, 1995, respectively). Data on the reliability and validity of the CCTI can be found in Rowe and Plomin (1977).

2.1.2.3. Social behavior during peer play. At ages 4 and 7, each child participated in a group play session with three unfamiliar, same sex, same age peers. At the 4-year visit, children were assigned to quartets based on an index of behavioral inhibition computed at age 2. At that time, the children had been presented with a series of novel stimuli (e.g., a robot, tunnel, unfamiliar adult) designed to assess their social responsiveness (see Kagan, Reznick, & Snidman, 1987). At the 7-year visit, the children were assigned to quartets based on their social play and reticence scores in the quartets at age 4. At each age, the quartets were designed so that one child was highly sociable, one child was low social, and the other two children were near the means on the measures (for details see Fox et al., 1996; Rubin et al., 1995).

To begin each quartet visit, the four children were led into a playroom where several age-appropriate toys were accessible. The visit was split into several episodes, a complete description of which may be found in Fox et al. (1995). For purposes of this study, data from two 15-min free play sessions were used, during which behaviors were coded using Rubin's (1989) Play Observation Scale (POS). Ten-second intervals were scored for social participation (unoccupied or onlooking behavior, type of solitary play, parallel play, peer conversation, and group play) and the cognitive quality of play (functional, dramatic and constructive play; exploration; games with rules). Three behavioral indices were computed: solitary-passive behavior (summing the proportion of coding intervals spent in solitary-exploratory and/or solitary-constructive play), social reticence (the sum of onlooking and unoccupied behavior), and social play (the sum of peer conversation and group play; see Coplan et al., 1994).

2.1.2.4. Cognitive tasks. Wisconsin Card Sorting Task. The computerized WCST was administered at age 11 using the standard procedures outlined in the test manual (Heaton, 1981; Wallner, 1996). The children were presented with four cards on a computer monitor. Below, a single target card was presented. The children were told to point to the card in the top row that corresponded to the target card. The children were given no indication of the matching rules they were to use. Feedback was presented for correct (high-pitched tone) and incorrect (low-pitched tone) matches.

The cards were presented with symbols that varied in type (triangle, star, plus sign, or circle), number (one, two, three, or four), and color (red, green, yellow, or blue). While the four standard cards did not vary, there were 64 unique target cards representing each combination of symbol type, number, and color. The matching rules were used in this order: color, form, number, color, form, number. The matching rule shifted after the children had correctly matched 10 consecutive target cards. Once all 64 target cards had been presented, they were reordered and presented once again if necessary. Testing continued until the children had six runs of 10 correct matches, until they had placed more than 64 cards in one category, or until they had exhausted two sets of target cards.

Coders noted the number of categories finished, the total number of trials completed, and the number of correct and incorrect responses. The incorrect responses were further subdivided into perseverative and non-perseverative errors. Errors were considered perseverative if they involved the continued use of the previously correct matching rule.

Tower of Hanoi. At age 11, the children were presented with an apparatus consisting of a wooden base (15.5 in × 5.75 in) with three pegs (4.5 in apart) at its center. Four wooden disks differing in diameter and color rested on the pegs. An identical apparatus was placed in front of the experimenter. The experimenter used her board to model disk configurations that the child was to replicate on his or her board. Disk patterns were presented to the child in order of difficulty (Wallner, 1996; Welsh, 1991). Difficulty was manipulated by either increasing the number of disks (three or four) or by designing the initial configuration so that the number of minimum moves increased (4–15). The child had to successfully complete a configuration twice before proceeding to the next level of difficulty. In order to successfully complete a configuration, the child had to reach the goal configuration in the fewest moves possible and without violating the following rules: 1. Only one disk could be moved at any time, 2. A disk had to be either on a peg or in the child's hand, and 3. A larger disk could not be placed on a smaller disk. The task ended if the child could not successfully complete two consecutive configurations.

The planning efficiency scores computed for this task took into account the number of problems completed and the number of trials it took the child to complete each of the problems.

2.1.3. Emotional Stroop data analysis

The reaction time data were edited for each child to remove error trials in the emotional Stroop, as well as any trials more than two standard deviations from his or her grand mean. Mean reaction times were then calculated for each of the three categories of words (see Fig. 1).

A repeated-measures ANOVA comparing the three word categories produced a significant main effect ($F(2,40) = 3.15$, $p = .05$) for reaction times. Pair-wise comparisons indicated that reaction times to negative words were significantly slower than to control words ($t(20) = 2.23$, $p = .04$), while there was no difference between positive words and control words ($t(20) = -0.01$, $p = .99$).

2.1.4. Group classification

In order to create two groups of children based upon their performance in the Stroop task, reaction times to the positive and negative words were averaged. Overall, there was no significant difference in reaction time between control words and the new emotion composite ($t(20) = -1.41$, $p = .17$). For each participant, reaction times to the control words were then subtracted from the new emotion word composite to calculate an index of the Stroop effect. In the interference group, reaction times to the emotion composite were slower than to the control words (positive score), while the facilitation group was faster in responding to the emotion words (negative score). Pair-wise comparisons indicated that the twelve subjects in the interference group showed significant interference with both positive ($t(11) = 2.40$, $p = .04$) and negative ($t(11) = 5.08$, $p < .001$) words. The remaining nine children, the facilitation group, were significantly faster with positive words ($t(8) = -4.17$, $p = .003$) and had means in the same direction for the negative words ($t(8) = -0.94$, $p = .37$).

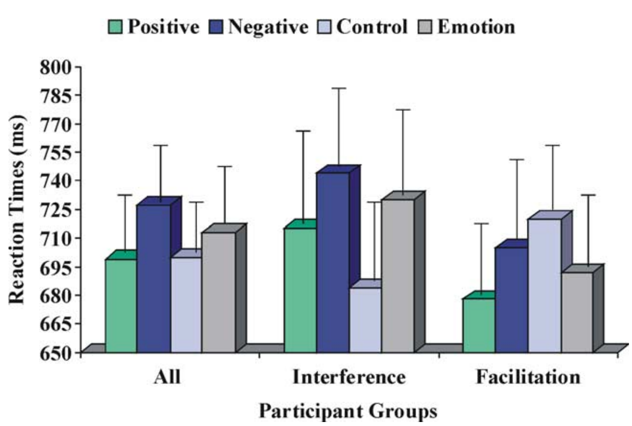

Fig. 1. Reaction times for Study 1. Reaction times (ms) are calculated for each word category within each participant group. The data for the emotion words were calculated by averaging the reaction times for positive and negative words.

2.2. Results

2.2.1. Maternal ratings of adjustment and temperament

CBCL. The children in the interference and facilitation groups differed on a number of sub-factors derived from the CBCL (see Table 1). An initial 3×2 ANOVA was used in each analysis. Age of rating (4, 7, and 11) was the within-subjects factor, while participant group (interference vs. facilitation) served as the between-subjects factor.

Parents rated the children in the interference group as significantly higher in social withdrawal ($F(1,17) = 7.71$, $p = .01$). This was significant at 4, 7, and 11 (t's > 2.30, p's $< .04$). The interference group was also rated as having more problems associated with anxiety and depression ($F(1,17) = 5.45$, $p = .03$), significant at ages 4 and 7 (t's > 2.18, p's $< .04$). Parents rated children in the interference group higher in attention problems ($F(1,17) = 4.63$, $p = .05$). This effect was significant at age 7 and 11 (t's > 2.16, p's $< .05$). In addition, the interference group was reported to show marginally greater social problems ($F(1,17) = 4.08$, $p = .059$). This was significant at age 7 ($t(18) = 2.27$, $p = .04$).

Table 1
Mean CBCL ratings for the children in Study 1 at ages 4, 7, and 11

	Age 4		Age 7		Age 11	
	Facilitation	Interference	Facilitation	Interference	Facilitation	Interference
Withdrawal	0.88*	2.55*	0.67*	2.55*	0.88*	3.11*
Anxious/depressed	0.63+	2.55+	1.11*	4.10*	1.88+	5.06+
Social problems	1.13	2.18	0.56*	2.36*	0.75	2.17
Attention problems	2.63	3.00	1.11*	3.91*	1.15*	4.25*
Delinquency	0.00	0.09	0.33+	1.64+	0.64	1.01
Aggressive behaviors	8.00	6.91	4.56	7.73	5.13	7.84
Internalizing	2.50+	5.82+	2.11*	7.55*	5.13*	12.08*
Externalizing	8.00	7.00	4.89+	9.36+	5.77	8.82
Total problems	16.5	20.00	9.73*	25.09*	14.71+	30.12+

Ratings are presented separately for the facilitation and interference groups.
* $p < .05$.
\+ $p < .10$.

Table 2
Mean CCTI ratings for the children in Study 1 at ages 4 and 7

	Age 4		Age 7	
	Facilitation	Interference	Facilitation	Interference
Shyness	1.80**	2.92**	3.29	3.19
Sociability	3.98*	3.23*	3.00	3.00
Emotionality	2.50	2.89	2.00	2.57
Soothability	3.06	3.04	3.49	3.16
Attention	3.60+	3.43+	2.76	2.91
Emotion dysregulation	−0.56	−0.16	−1.49	−0.59
Impulsivity	6.63	6.57	4.93	5.31

Ratings are presented separately for the facilitation and interference groups.
* $p < .05$.
** $p < .01$.
+ $p < .10$.

On the broad-band measures, the interference group was rated by parents as displaying higher internalizing scores ($F(1,17) = 5.18$, $p = .04$). This was true at ages 4 ($t(17) = 1.92$, $p = .07$), 7 ($t(18) = 2.22$, $p = .04$) and 11 ($t(16) = 2.05$, $p = .06$). While they also had higher externalizing scores, this measure never reached significance (t's < 1.67, p's $> .11$). On the measure of total problems, the interference group had higher scores at 7 ($t(18) = 2.49$, $p = .02$) and 11 ($t(16) = 1.92$, $p = .08$).

CCTI. Based on parent ratings from the CCTI, the children in the interference group were rated as more shy ($t(16) = 3.85$, $p = .001$) and less social ($t(16) = -2.81$, $p = .01$) than their counterparts in the facilitation group at age 4 (see Table 2). This pattern was not repeated at age 7 ($t(16)$'s < -0.79, p's $> .44$). There were no group differences on emotionality, soothability, or attention at either 4 or at 7 (t's < 1.63, p's $> .12$). There were also no differences on the emotion dysregulation (t's < 1.53, p's $> .14$) and impulsivity (t's < 0.85, p's $> .41$) indices.

2.2.2. Social behavior during peer play

There were no significant differences between the interference and facilitation groups in social reticence, solitary passive play, or social play (see Table 3). This was true at age 4 ($t(19)$'s < 0.62, p's $> .54$) and at age 7 ($t(18)$'s < 0.78, p's $> .45$).

Table 3
Mean ratings for social behavior in the quartets for the children in Study 1 at ages 4 and 7

	Age 4		Age 7	
	Facilitation	Interference	Facilitation	Interference
Reticence	0.23	0.18	0.13	0.18
Solitary passive	0.17	0.13	0.11	0.13
Social play	0.13	0.16	0.40	0.37

Ratings are presented separately for the facilitation and interference groups.

2.2.3. Cognitive tasks

WCST. In the WCST, the interference and facilitation groups did not differ in the number of categories finished, the total number of trials performed, or the number of correct trials ($t(16)$'s < 1.73, p's $> .10$) (see Table 4). While there was a weak trend for the facilitation group to show greater errors (21.3 vs. 39.0; $t(16) = -1.87$, $p = .08$), this was driven solely by the number of non-perseverative errors (9.4 vs. 19.6; $t(16) = -2.71$, $p = .02$). The groups did not differ in the number of perseverative errors (12.0 vs. 19.4; $t(16) = -1.01$, $p = .29$).

TOH. In the TOH, the two groups did not differ in any condition, regardless of the complexity of the configuration ($t(17)$'s < -1.46, p's $> .16$).

2.3. Discussion

Although a simple reaction time measure, performance on the emotional Stroop task was linked to a

Table 4
Mean scores on the Wisconsin Card Sorting Task and Tower of Hanoi at age 11 for Study 1

	Facilitation	Interference
Wisconsin Card Sort		
Categories finished	4.43	5.36
Number of trials	109.00+	92.73+
Number correct	70.00	71.36
Number errors	39.00+	21.36+
Perseverative errors	19.43	12.00
Nonperseverative errors	19.57*	9.36*
Tower of Hanoi		
Three disks, five moves	5.67	4.40
Three disks, six moves	3.78	4.30
Three disks, seven moves	3.78	2.90
Four disks, seven moves	3.89	3.00
Four disks, 11 moves	3.11	2.50
Four disks, 15 moves	1.78	1.90

Scores are presented separately for the facilitation and interference groups.
* $p < .05$.
+ $p < .10$.

consistent pattern of behavior across time on a broad range of social and emotional measures. At every age, the children in the interference group showed greater signs of maladjustment than the children in the facilitation group. It is important to note, however, that at no time did any child reach a clinical cutoff. Rather, the interference and facilitation groups simply clustered at different points along the normal distribution. However, it is striking that reliable results can be documented over a seven-year span with children who were not selected for particular physiological or behavioral characteristics.

The current data indicate that the social and emotional dispositions found in the children are not transient in nature. However, within this stability, the constellation of findings was most acute at age 7. Since the expression of a behavioral predisposition is dependent on a number of external and internal forces, it is not surprising that the pattern peaks at particular points in development. At age 7, there is a shift in children's responsibilities as they enter formal schooling. As well, social and emotional forces at play are felt quite acutely in the children as they begin to navigate increasingly complex social roles at school and at home. By age 11, these same children may be better equipped for self-regulation, thus attenuating the behavioral markers.

The null results in the Wisconsin Card Sort Task and the Tower of Hanoi are in line with previous studies comparing social withdrawal and performance on these tasks (Fox & Henderson, 1997). In addition, the groups did not differ in overall response speeds to the control words, indicating that they were both comfortable with the task demands of the Stroop. Rather, differences in self-regulation ability appear only when affective and regulatory systems are called on in tandem, either temporally or functionally, as in the emotional Stroop.

3. Study 2

A second study was designed with two issues in mind. First, we attempted to replicate these initial findings with an independent sample. Second, while behavioral (reaction time) differences were quite evident in the emotional Stroop, little information was available concerning the neural and cognitive processes underlying the data.

While MRI and PET studies have greatly expanded our understanding of the neuroanatomical substrates of the emotional Stroop, these tools cannot offer real-time temporal resolution of the neural processes involved. This level of temporal specificity is necessary in order to address the cognitive processes that fill the void between stimulus presentation and response production. To fill this requirement, event-related potentials (ERPs) were employed in Study 2.

The earliest ERP-Stroop study (Duncan-Johnson & Kopell, 1981) found no consistent differences across conditions in the traditional Stroop. They concluded that Stroop interference was related to conflicts in response selection rather than in stimulus evaluation processes. However, more recent studies have indicated that there are detectible differences in ERPs generated by the traditional (color-word) Stroop (Ilan & Polich, 1999; Liotti, Woldroff, Pérez III, & Mayberg, 2000; Schack, Chen, Mescha, & Witte, 1999; West & Alain, 1999). Early in the ERP wave, studies have found a distinct N1–P2–N2 complex (e.g., West & Alain, 2000a). These components are thought to index early sensory processing and low-level attention allocation (Hillyard, Luck, & Mangun, 1994). Traditional Stroop studies have not found any differences across their two conditions: congruent vs. non-congruent color words. Rather, these studies have focused on more endogenous components, the P3 and N4 (e.g., Ilan & Polich, 1999). The larger amplitudes noted in the incongruent condition are thought to reflect the stimulus evaluation time and attentional requirements needed to ultimately repress the information carried in the incongruent trials (West & Alain, 2000b). This interpretation also carries over to the positive slow wave prominently seen in traditional ERP studies (West & Alain, 2000a).

While the ERP Stroop literature is growing steadily, the currently published studies have focused exclusively on the traditional Stroop. As such, this second study was among the first to directly examine the ERPs generated by an emotional Stroop task. It therefore allowed us to see if the relationships outlined above for congruent versus incongruent color-word stimuli were replicated in the affective and neutral stimuli employed in the emotional Stroop. In addition, this study is one of only a small number (Barch et al., 1999; Bauer & Hesselbrock, 1999) employing an individual differences approach.

3.1. Method

3.1.1. Participants

The participants in this study were 48 children selected at age four for a longitudinal study of the behavioral and physiological correlates of temperament (see Fox et al., 1996). Families were recruited from the metropolitan Washington, DC, area via a general mailing. The children were primarily Caucasian (88.6%) and of middle-class background. Children were excluded on the basis of maternal pre-natal health complications or infant postnatal health problems. In this paper we report on the thirty-one children (13 male) who returned to the laboratory at age 11. An analysis comparing the children who participated at 11 versus those who did not return indicated that the children did not differ on any demographic or behavioral factors.

3.1.2. Procedures

The procedures employed for the maternal ratings, WCST, and TOH were identical to those in Study 1.

3.1.2.1. Emotional Stroop. At the 11-year visit, the children were presented with 63 words, 21 in each word category: positive, negative, and control (see Appendix A). The words were chosen as representatives of broad affective states. Unlike many emotional Stroop studies, the words were not tailored to target particular psychological concerns (e.g., anxiety) or idiographic factors. Post hoc analyses (Frances & Kucera, 1982) indicated that while positive and control words did not differ in word frequency ($t(40) = 1.07$, $p = .29$) or number of syllables ($t(40) = 1.28$, $p = .21$), the negative words were less frequent ($t(40)$'s > 2.48, p's $< .02$) than the positive and control words. In addition the negative words had significantly more syllables ($t(40) = 2.43$, $p = .02$) than the control words. An item analysis indicated that the mean reaction times in the emotional Stroop task did not correlate significantly with number of syllables ($r(63) = .237$, $p = .06$). There was, however, a significant correlation between word frequency and average reaction times ($r(63) = -.289$, $p = .02$). Separate analyses conducted for the three word categories showed that this held only for the control words ($r(21) = -.508$, $p = .02$). There was no such relationship for the positive and negative words ($r(21)$'s $< -.349$, p's $> .12$).

3.1.2.2. EEG collection. EEG was collected from each child while performing the emotional Stroop task.

Ten sites (F3, F4, Fz, C3, C4, P3, P4, Pz, O1, and O2) were collected using a stretch Lycra cap with electrodes sewn in according to the 10–20 international system (Jasper, 1958). Impedances were kept below 5 kΩ. During collection the signal was digitized at 512 Hz referenced to vertex (Cz). The signal was later converted to average reference for analysis.

EOG was recorded via electrodes above and below the right eye and at the outer canthi of each eye. The data from 24 children were amplified through individual Grass AC bioamplifiers (Model 78D) using high- and low-pass filters of 1 and 100 and a 60 Hz notch filter. The signal was digitized using Snapshot-Snapstream acquisition software (HEM Data). The data from the remaining six children were collected with SA Instruments (San Diego, CA) isolated bioelectric amplifiers using high- and low-pass filters of .10 and 100 Hz. The signal was digitized with the Snapmaster Data Acquisition System (HEM Data). Before each visit a 50 μV 10 Hz signal was input into each of the channels for calibration purposes.

The digitized EEG data were manually edited for eye-blink or movement related artifact. Eye blinks were regressed out using software provided by James Long Company (Caroga Lake, NY). All other artifact was expunged from the files.

Event-related potentials (ERPs) were collected simultaneously with the presentation of each word, referenced to a 100 ms pre-stimulus baseline. Each of the included trials was artifact free for the 1000 ms following word presentation. Grand ERPs for each of the three word categories were calculated by averaging across the digitized and edited EEG from individual trials.

On average, there were 15.7 positive, 15.6 negative, and 15.0 control trials available for each participant. There were no statistically significant differences across the word categories ($t(30)$'s < 1.41, p's $> .17$). A repeated-measures ANOVA with word category as the within-subject factor and participant group (Interference vs. Facilitation) as the between-subject factor showed neither a main effect for word category nor a word category by group interaction. However, there was a main effect of group ($F(1, 29) = 9.08$, $p = .01$). This was due to the interference group having more useable trials than the facilitation group (49.6 vs. 42.3, total).

3.1.2.3. Social behavior during peer play. At ages 4 and 7, each child participated in a group play session with three unfamiliar, same sex, same age peers. At the 4-year visit, children were assigned to quartets based on maternal assessments of sociability and shyness on the CCTI. At the 7-year visit, children were assigned to quartets based on their social play and reticence scores in the quartets at age 4. At each age, the quartets were designed so that one child was highly sociable, one child was low social, and the other two children were near the means on the measures (for details see Fox et al., 1996; Rubin et al., 1995).

The testing and coding procedures were identical to those in Study 1.

3.1.3. Emotional Stroop data analysis

Reaction times for each child were edited to remove errors and any trials more than two standard deviations from his or her grand mean. Mean reaction times were then calculated for each of the three categories of words (see Fig. 2).

A repeated-measures ANOVA comparing the three word categories indicated no significant differences ($F(2, 56) = 0.52$, $p = .60$) in reaction times. Pair-wise comparisons also failed to show any differences in responses to positive ($t(29) = 0.34$, $p = .73$) and negative words ($t(29) = 0.96$, $p = .49$) versus controls.

3.1.4. Group classification

As in Study 1, an index of the emotional Stroop effect was calculated for each child using reaction times to the control words and the mean of reaction times to the positive and negative words. Overall, there was no significant difference in reaction times between the emotion word composite and the control words ($t(29) = -0.70$, $p = .49$). The 16 subjects in the interference group were

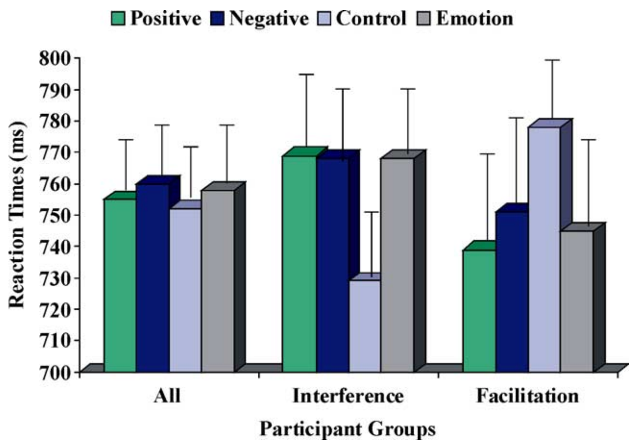

Fig. 2. Reaction times for Study 2. Reaction times (ms) are calculated for each word category within each participant group. The data for the emotion words were calculated by averaging the reaction times for positive and negative words.

slower in responding to positive ($t(15) = 4.72$, $p < .001$) and negative ($t(15) = 4.85$, $p < .001$) words versus control words. The remaining fourteen children, the facilitation group, were significantly faster with positive ($t(13) = -5.61$, $p < .001$) and negative ($t(13) = -3.35$, $p = .01$) words.

3.2. Results

3.2.1. Maternal ratings of adjustment and temperament

Based on parent ratings of the child on the CBCL (see Table 5), the children in the interference group showed more attention problems ($F(1, 17) = 5.70$, $p = .03$). This was significant at age 7 ($t(25) = 3.04$, $p = .01$). The interference group also tended to show more social problems ($F(1, 17) = 4.11$, $p = .059$). Again, the effects appeared to peak at age 7, at which time the interference group was more anxious and depressed ($t(25) = 1.97$, $p = .05$) and more aggressive ($t(25) = 2.77$, $p = .01$).

On the broad-band measures, the interference group showed higher internalizing scores at 4, 7, and 11, although these did not reach significance (t's < 1.30, p's $> .26$). The interference group also had higher externalizing scores throughout, reaching significance at age 7 ($t(25) = 2.71$, $p = .01$). On the measure of total problems, the interference group had higher scores at all three ages ($F(1, 17) = 4.45$, $p = .05$), again reaching significance at age 7 ($t(25) = 2.89$, $p = .01$).

Based on ratings from the CCTI noted above, the children in the interference group showed more impulsivity ($t(25) = 2.14$, $p = .04$) (see Table 6). Unlike in Study 1, here the interference group was *less* shy than the facilitation group at age 7 ($t(25) = -2.41$, $p = .02$).

3.2.2. Social behavior during peer play

There were no significant differences between the interference and facilitation groups in social reticence, solitary passive play, or social play at age 4 ($t(27)$'s < 0.76, p's $> .46$) (see Table 7). At age 7, the interference group showed less solitary passive behavior ($t(25) = -2.06$, $p = .05$) and more social play ($t(25) = 2.58$, $p = .02$). These data are in line with the maternal ratings on the CCTI. There were no significant differences between the groups in social reticence ($t(25) = -0.75$, $p = .46$).

3.2.3. Cognitive measures

In the WCST, the groups did not differ on any of the coding measures ($t(25)$'s < 1.92, p's $> .07$) (see Table 8).

In the TOH, the two groups did not differ in any condition, regardless of the complexity of the configuration ($t(28)$'s < 1.85, p's $> .08$).

3.2.4. ERP data

The ERP findings did not follow the pattern of results noted in the behavioral data. Therefore, these data will be address in a separate section after the discussion of the current results.

Table 5
Mean CBCL ratings for the children in Study 2 at ages 4, 7, and 11

	Age 4		Age 7		Age 11	
	Facilitation	Interference	Facilitation	Interference	Facilitation	Interference
Withdrawal	2.00	2.00	2.17	1.67	2.40	2.20
Anxious/depressed	2.50	2.65	2.58^{+}	4.47^{+}	1.72	3.91
Social problems	1.64	1.31	1.17	2.00	0.60	1.70
Attention problems	2.07	2.35	1.25**	3.73**	1.52	3.05
Delinquency	0.65	0.88	1.25^{+}	2.20^{+}	1.10	1.30
Aggressive behaviors	7.71	10.41	5.33**	10.20**	4.50	5.75
Internalizing	5.36	6.00	5.42	7.20	5.33	8.41
Externalizing	8.36	11.35	6.58**	12.40**	5.60	6.92
Total problems	20.21	23.63	14.79**	25.85**	14.41	22.33

Ratings are presented separately for the facilitation and interference groups.
$^{**}p < .01$.
$^{+}p < .10$.

Table 6
Mean CCTI ratings for the children in Study 2 at ages 4 and 7

	Age 4		Age 7	
	Facilitation	Interference	Facilitation	Interference
Shyness	2.41	2.41	2.60	2.03
Sociability	3.70	3.98	3.57*	3.87*
Emotionality	3.07	2.83	2.35	2.69
Soothability	3.39	3.30	3.47	3.56
Attention	3.39	3.34	3.55	3.51
Emotion dysregulation	−0.32	−0.48	−1.12	−0.87
Impulsivity	6.71	6.75	5.93*	6.53*

Ratings are presented separately for the facilitation and interference groups.
* $p < .05$.

3.3. Discussion

The behavioral findings in Study 2 complement those of Study 1. As in Study 1, children in the interference group in Study 2 showed greater signs of maladjustment, scoring higher on measures of both externalizing and internalizing problems. This pattern was consistent across time, again peaking at age 7. There were, however, some subtle differences in the pattern of results across the two studies. While the findings for the interference group children in Study 1 tended to cluster around internalizing measures (e.g., withdrawal, anxiety), the children in Study 2 had greater difficulty with externalizing behaviors (e.g., aggression). This is in line with the questionnaire and behavioral data in Study 2 indicating that the interference group was less shy than the facilitation group. It is important to note that the interference group had higher scores for both internalizing and externalizing problems in both studies and that this was reflected in the significantly higher total problems scores (see Fig. 3).

Table 7
Mean ratings for social behavior in the quartets for the children in Study 2 at ages 4 and 7

	Age 4		Age 7	
	Facilitation	Interference	Facilitation	Interference
Reticence	0.16	0.15	0.16	0.11
Solitary passive	0.29	0.24	0.18*	0.09*
Social play	0.29	0.32	0.34*	0.57*

Ratings are presented separately for the facilitation and interference groups.
* $p < .05$.

Table 8
Mean scores on the Wisconsin Card Sorting Task and tower of Hanoi at age 11 for Study 2

	Facilitation	Interference
Wisconsin Card Sort		
Categories finished	5.92+	5.20+
Number of trials	94.58	107.67
Number correct	71.50	72.07
Number errors	23.08+	35.60+
Perseverative errors	11.67	17.27
Nonperseverative errors	11.42	18.27
Tower of Hanoi		
Three disks, five moves	5.50	5.69
Three disks, six moves	4.93	5.56
Three disks, seven moves	4.93	4.81
Four disks, seven moves	5.29	4.56
Four disks, 11 moves	3.86	2.63
Four disks, 15 moves	2.43+	1.06+

Scores are presented separately for the facilitation and interference groups.
+ $p < .10$.

3.4. Conclusion

The emotional Stroop has been used extensively in clinical populations. Since early studies often involved individuals with anxiety disorders (Williams et al., 1996), the initial explanations for the phenomena centered on issues central to theories of anxiety. In particular, Stroop interference was thought to grow out of the extended practice anxious individuals have with anxiety-related concepts as they ruminate on particular themes (Segal, Truchon, Horowitz, Gemar, & Guirguis, 1995).

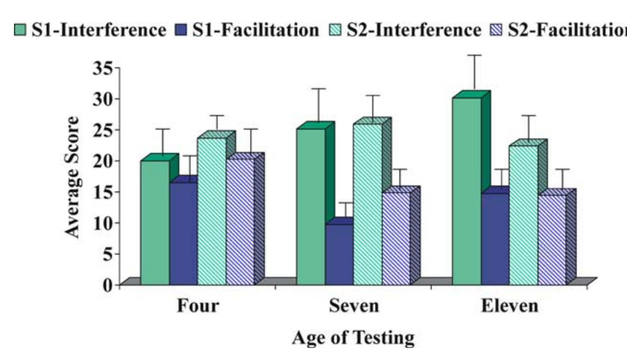

Fig. 3. CBCL total problems scores. Total problem scores are noted for the facilitation and interference groups in each study. The children in the interference group showed a generally stable pattern of higher scores across the three ages.

However, the emotional Stroop has proven robust in populations that do not suffer from anxiety-related disorders, or any disorder at all. Clearly, the mechanisms involved can not be unique to anxiety disorders. A related class of explanations also focused on the fact that most studies use carefully constructed stimuli designed to target the idiosyncratic concerns of the population of interest. Gilboa-Schechtman (Gilboa-Schechtman et al., 2000) noted that this selective processing could be attributed to three distinct factors: (1) the concordance between the person's affective state and the valence of the stimulus, (2) the relevance of the stimuli to the individual's current concerns, and (3) the emotional impact of the stimulus. Depending on the facts at hand, each of these explanations has been invoked, alone or in combination, to interpret data from the emotional Stroop.

This study differed from much of the literature in that neither the stimuli nor the participants were chosen to target a particular psychological or social concern. Yet, two independent samples found clear differences in performance that were linked to distinct patterns of social behavior across a seven-year span. In order to interpret the current Stroop data, we can attempt to apply the three broad explanations noted above.

First, there is the possibility that the findings were due to mood congruence. This cannot adequately explain the current data since the children were tested in a relatively affect-neutral state and the pattern of results (interference vs. facilitation) held for both positive and negative words. Valence was relevant only for the ERP data, which will be discussed in the special section below.

Second, it may be that personal relevance fuels the emotional Stroop effect. Again, this reasoning does not seem to fit the current data. As noted above, the stimuli were chosen to represent broad positive and negative emotions without regard to a particular psychological or emotional concern. Secondly, the literature stresses that the presence and magnitude of the Stroop effect are closely linked to the individual's concerns at that moment in time (Riemann & McNally, 1995). Here we see that performance on the Stroop is linked to behavior across a wide span of time. While we do not know how the children would have performed at ages 4 and 7, it is unlikely that the emotional, psychological, and social concerns of these children were static during this developmental period. Indeed, since this was an unselected sample, one cannot make the secondary argument that the children's concerns were expected to be similar across the participant pool.

Third, the Stroop effect found in the studies may reflect the emotional impact of the stimuli. Given the motivations behind most emotional Stroop studies, the unspoken assumption is that the individual is reacting to a negative emotional state. In this study, however, the children were grouped based on responses to emotional stimuli without regard to valence. In addition, since the stimuli were randomly presented, it seems highly unlikely that the children vacillated between strong positive and negative emotional states as the task progressed. Taken together, the data suggest that the current Stroop findings do not reflect an acute emotional state in the children. Rather, the mechanisms involved in interference may reflect an emotional/regulatory style on the part of the children. The notion that personality/temperament might play a role in the Stroop effect is supported in the work of Mogg and Marden (1990). They found that individuals with high trait anxiety showed interference to both positive and negative words.

In the developmental literature the presence of an enduring behavioral style is thought to reflect underlying temperamental traits in children. This literature has focused primarily on the expression and regulation of negative affect and much of this work may be useful for understanding the current findings.

There are a growing number of studies in the developmental literature exploring individual differences in the expression and control of emotions. Much of this research has focused on temperamental differences in behavioral inhibition and social withdrawal (Fox et al., 1996; Kagan et al., 1987). As infants, behaviorally inhibited children show high levels of motor activity and marked negative affect when presented with novel sensory stimulation (Calkins, Fox, & Marshall, 1996; Kagan, Reznick, Clarke, Snidman, & Garcia-Coll, 1984). Behaviorally inhibited children also display signs of fear and wariness in response to unfamiliar stimuli (Schmidt et al., 1997). By pre-school, many of these same children are reluctant to interact with unfamiliar peers and often appear shy and withdrawn in social situations (Fox et al., 2001). Behaviorally inhibited children also appear anxious and may be at greater risk for anxiety disorders as adults (Kagan, 1994).

Children who are behaviorally inhibited or socially withdrawn often experience negative affect in unfamiliar or novel social environments. Once experiencing negative emotions, these children may have greater difficulty regulating and stabilizing their affective experiences (Rothbart, Ahadi, & Hershey, 1994; Ruff & Rothbart, 1996). This dynamic balance between the experience of negative affect and its regulation has led researchers to explore the cognitive and neuropsychological components that may underlie individual differences in the expression and regulation of emotions among behaviorally inhibited or socially withdrawn children (Fox, Henderson, & Marshall, 2001; Fox et al., 1995).

The origins of negative affect and distress among behaviorally inhibited children may be linked to hyperexcitability of certain limbic centers. Kagan (Kagan, Reznick, Snidman, Gibbons, & Johnson, 1988) has argued that inhibited children show greater arousal in selected hypothalamic and limbic sites, particularly the amygdala. As noted above, these same regions have also been implicated in the Emotional Stroop (Isenberg et al., 1999; Vogt et al., 1995). The work of Michael Davis (1992) and Joseph LeDoux (1998) has indicated that the amygdala plays a central role in the expression of conditioned fear in animals.

Building on this work, Kagan speculated that these same systems might be involved in the display of behavioral inhibition in human infants and children. Both Davis (Davis & Shi, 1999) and LeDoux (1990) found that activity within the central nucleus of the amygdala was critical for the expression of conditioned fear. Heightened amygdala activity was also associated with several behavioral and physiological outputs including increased startle response, autonomic changes, and heightened activity in the HPA axis. Kagan and Fox have conducted a series of studies with independent samples and have found that inhibited children display a similar pattern of responses. Inhibited or reticent children showed increased heart rate and decreased heart rate variability (Calkins et al., 1996), enhanced startle (Schmidt & Fox, 1998, but not Schmidt et al., 1997), and elevated morning basal cortisol levels (Schmidt et al., 1997, but not Schmidt, Fox, Schulkin, & Gold, 1999a, 1999b). Fox and colleagues further report increased alpha desynchronization particularly from right anterior scalp leads in inhibited and reticent children (Fox et al., 2001). This pattern of physiological and behavioral findings suggests that the negative affect and distress seen in inhibited and reticent children may be a function of the proposed limbic hyper-arousability.

Although inhibited and reticent children may share an underlying biology that predisposes them to express greater levels of negative affect and distress, there are clear individual differences among these children in the actual expression of inhibition or reticence. Such differences in the expression of temperament may be a function of external factors such as caregiving experience or learning, as well as internal processes such as the maturation of neural systems involved in the regulation of both positive and negative affect. Emotion regulation involves the use of executive control to shape affective processes and their corresponding behavioral markers. Fox (Fox et al., 2001) argues that successful regulation depends on the development of cognitive processes (e.g., selective attention skills) and on the child's experiences within his or her daily environment, which may or may not support the efficacy of these skills for affect regulation. Thus, for example, behaviorally inhibited children may develop the attentional skills important for modulation of affect but still be unable to efficiently utilize these skills for adaptive prosocial behavior due to poor support from their caregiving environment.

The notion of temperamental style is particularly important when one considers that the current data is not unidirectional. That is, the groups were not chosen for having different degrees of Stroop interference. Instead, roughly one-half of the children preformed "better" with the emotion stimuli. This finding has not been dealt with extensively in the literature since most emotional Stroop studies have been designed to elicit the largest Stroop effect possible. van Honk (van Honk et al., 2000) found both interference and facilitation effects when presenting masked and unmasked pictures of angry faces. He argued that individuals are biologically prepared to exhibit either a dominant or submissive stance when confronted with a threat cue. In his study, the increase in testosterone and cortisol for the facilitation group readied them for an active, aggressive response to the perceived threat. By the same token, the drop in these hormones for the interference group hinted that this group may be inhibited in terms of their behavioral and physiological responses. In the current study, children in the interference group were disposed to showing signs of social and emotional maladjustment. However, it remains to be seen if children who showed facilitation in the emotional Stroop are by comparison merely less prone to these difficulties, if they are actively inoculated against social-emotional difficulties, or if they are temperamentally predisposed to exhibit positive or exuberant behavior.

Clearly, further research will be needed in order to clarify the promising data from the current studies. The data suggest that the role of the emotional Stroop as an empirical and theoretical tool may be broader than initially thought. Rather than being limited to acute tests of emotionality within strict diagnostic boundaries, the emotional Stroop may help reveal broad styles of functioning that remain relatively stable across time. In addition, the emotional Stroop appears to tap into the interaction between emotion and cognition at an early age, perhaps allowing for greater insight into the development of self-regulation. If the emotional Stroop proves to be useful as both a passive dependent measure and as an independent factor in the classification and study of individuals, it may prove particularly helpful to researchers applying an individual differences approach to the issues at hand. This will allow researchers to examine both the causes and consequences of differing response patterns in the emotional Stroop.

4. Event-related potentials in the emotional Stroop

4.1. Results

A grand ERP was calculated across participants and conditions for each site in order to select prominent components for analysis. Based on these ERPs, peak amplitude and latency were calculated at each site for each of the following components: P1 (0–100 ms), N1 (50–150 ms), P2 (150–250 ms), N2 (250–350 ms), P3 (350–450 ms), and N4 (400–500 ms). In addition, mean amplitude was calculated for the positive slow wave from 600 to 1000 ms.

A separate ANOVA was calculated with Word Category (Positive, Negative, and Control), Region (Frontal, Central, Parietal, and Occipital), and Hemisphere (Right vs. Left) as within-subject factors and the pattern of Stroop effect (Interference vs. Facilitation) as a between-subjects factor. The individual ERP components served as the dependent measures. After the initial ANOVA, separate 3 (Word Category) × 2 (Hemisphere) × 2 (Group) ANOVAs were calculated for each scalp region.

The data analyses were organized across three central questions.

4.1.1. *Are ERPs generated by the emotional Stroop similar in overall morphology to ERPs collected during the traditional Stroop task?*

The ERPs generated by the emotional Stroop task produced a number of distinct wave components (see Fig. 4). Early in the waveform, there is a distinct

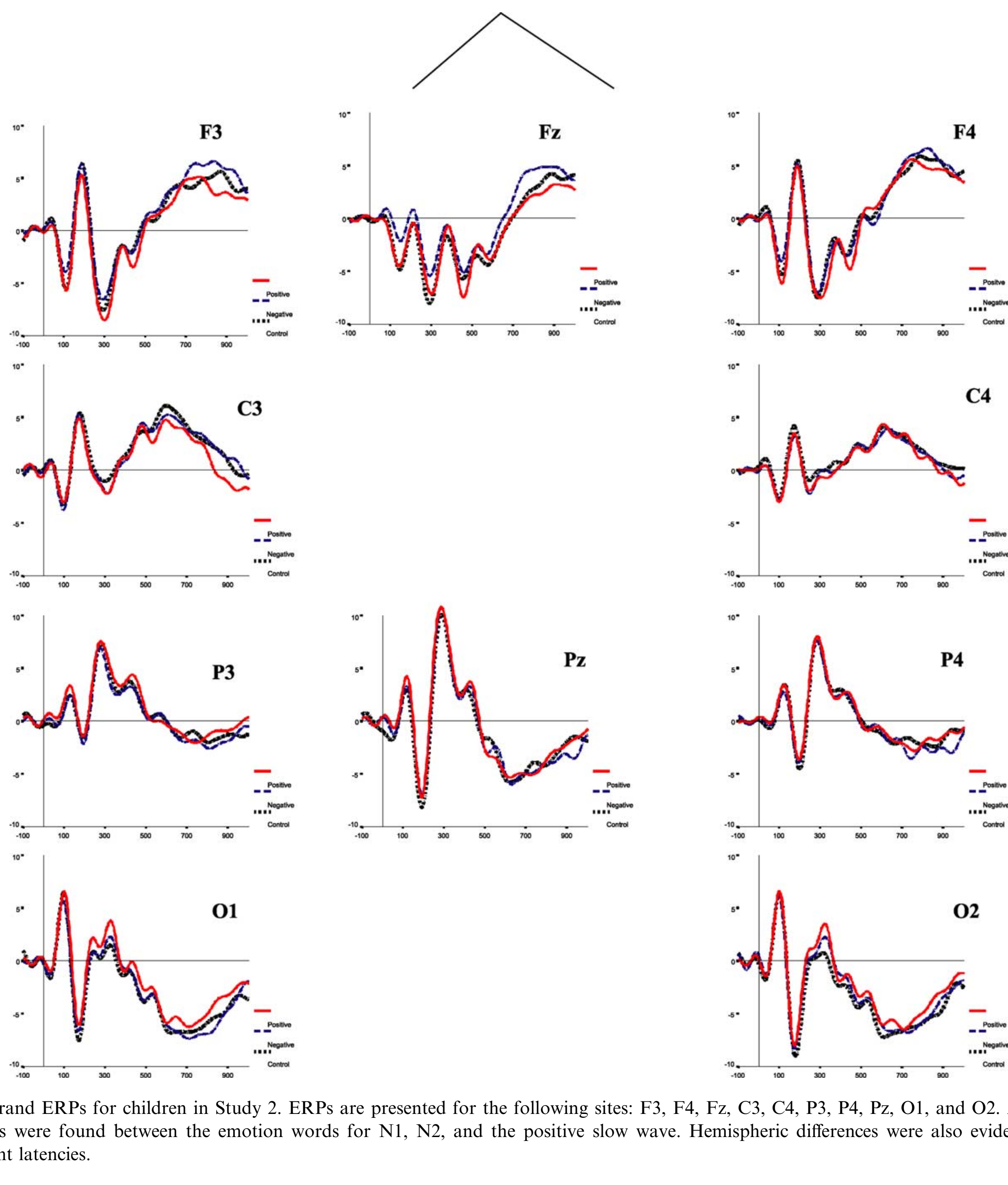

Fig. 4. Grand ERPs for children in Study 2. ERPs are presented for the following sites: F3, F4, Fz, C3, C4, P3, P4, Pz, O1, and O2. Amplitude differences were found between the emotion words for N1, N2, and the positive slow wave. Hemispheric differences were also evident for the component latencies.

P1–N1–P2–N2 complex. This closely mirrors the N1–P2–N2 complex seen in the ERP studies of the traditional Stroop (Bauer & Hesselbrock, 1999; Ilan & Polich, 1999; Liotti et al., 2000; West & Alain, 1999; West & Alain, 2000a, but not West & Alain, 2000b). The traditional Stroop studies also found prominent P3 and N4 components. While these components are discernible in the emotional Stroop, they are attenuated. Interestingly, the only traditional Stroop study not to find strong P3 and N4 effects (Bauer & Hesselbrock, 1999) also involved a fairly young (late adolescence) participant population. Finally, each traditional Stroop study finds a marked positive slow wave (except for West & Alain, 2000b, where it is negative going). The positive slow wave was particularly prominent in the current study.

4.1.2. Can ERPs distinguish the different word categories used in the emotional Stroop?

For the early P1–N1–P2–N2 complex, only the negative components produced significant word category effects (see Fig. 5).

For N1, the initial ANOVA indicated a trend distinguishing among positive, negative, and control words ($F(2,58) = 2.92$, $p = .06$). This was significant only for the frontal sites ($F(2,58) = 6.95$, $p = .002$). Pair-wise comparisons indicated that N1 amplitudes for negative words were significantly smaller than for positive ($t(30)$'s > 2.28, p's $< .03$) or control ($t(30)$'s > 2.10, p's $< .04$) words.

At N2, the initial ANOVA indicated a significant effect for word category ($F(2,58) = 4.02$, $p = .02$). This held for the frontal ($F(2,58) = 2.87$, $p = .07$), central ($F(2,58) = 3.54$, $p = .04$), and occipital ($F(2,58) = 5.78$, $p = .01$) sites. Pair-wise comparisons indicated that the amplitudes to negative words were smaller than those for positive words ($t(30)$'s > 1.91, p's $< .06$).

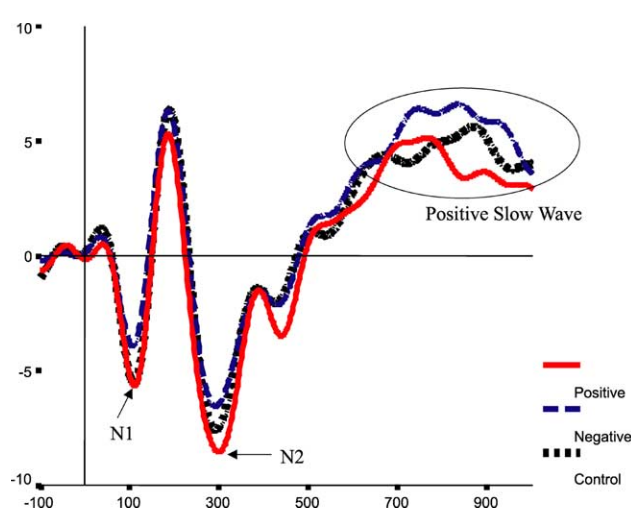

Fig. 5. ERP from F3. Individual components for the grand ERP are noted. Of particular interest are N1, N2, and the positive slow wave.

The amplitude of a positive component is often thought to reflect the amount of processing resources an individual can allocate among tasks and task components. However, some have argued that a reciprocal change is found in negative components (Mecklinger, Kramer, & Strayer, 1992). If correct, this may help explain why the negative words produced the smallest component amplitudes.

These findings do not concur with the traditional Stroop studies, which generally do not find significant effects in the early sensory components (but see Ilan & Polich, 1999).

There were no significant word category effects for the P3 and N4 components. Again, this is in contrast to the significant differences across experimental conditions found in the traditional Stroop.

For the positive slow wave, the initial ANOVA showed that the negative words had the largest mean amplitude across scalp regions. This finding was more pronounced in the left hemisphere ($F(2,58) = 3.66$, $p = .03$). Here, the data mirror the traditional Stroop.

Except for Ilan and Polich (1999), each of the currently published ERP studies focuses exclusively on peak amplitude. Ilan and Polich (1999) found no significant P3 latency differences across experimental conditions. In the current study, there were also no consistent main effects for word category. However, there was a strong pattern of data indicating hemispheric differences in the processing of positive and negative words.

At P1, the main effect for hemisphere was significant at the frontal sites ($F(1,29) = 4.02$, $p = .05$). This was driven by the fact that negative words tended to have shorter latencies than positive words in the right hemisphere ($t(30) = 1.91$, $p = .07$).

While there were no significant effects at N1 and P2, the N2 component produced a significant word category by hemisphere interaction in the initial ANOVA ($F(2,58) = 4.02$, $p = .02$). For the left hemisphere sites, positive words had a shorter latency than the negative words (293 vs. 296 ms), while in the right hemisphere, the reverse was true (296 vs. 289 ms). This pattern was repeated for the frontal sites, although it did not reach significance ($F(2,58) = 2.68$, $p = .08$). For the frontal, central, and parietal sites there was also a significant main effect of hemisphere ($F(1,29)$'s > 4.20, p's $< .05$), such that the latencies were shorter for sites on the right.

For P3, the central sites showed a significant word category by hemisphere interaction ($F(2,58) = 3.22$, $p = .05$), indicating that negative words showed delayed latencies in the left hemisphere.

At N4, the initial ANOVA produced a marginally significant main effect for word category ($F(2,58) = 2.99$, $p = .06$). Here, negative words had significantly shorter latencies than positive words ($t(30) = 2.10$, $p = .04$). While the main effect of word category was not significant for the frontal sites, there was a significant interaction between word category and hemisphere. In the frontal sites, positive words had shorter latencies in the left hemisphere, while negative words had shorter latencies in the right hemisphere ($F(2, 58) = 3.66$, $p = .03$).

This pattern of data is striking since the subject classification criterion, the emotional Stroop index score, was derived without regard to valence. Indeed, pair-wise comparisons indicated the children within each group treated positive and negative words in the same manner. However, the ERP data center on differences between words of different valence. It remains to be seen how the transition between word processing (ERP) and response selection (Reaction Time) is carried out.

4.1.3. *Are there discernible ERP differences between the interference and facilitation groups?*

There were no consistent group differences across ERP components produced by the emotional Stroop. However, there were two findings of note. First, larger mean amplitudes were found for the positive slow wave in the parietal sites for the facilitation group ($F(2, 58) = 3.07$, $p = .05$). Second, there was a significant word category by group interaction for P2 latencies at the parietal sites ($F(2, 58) = 3.81$, $p = .03$) and a trend in the frontal sites ($F(2, 58) = 2.79$, $p = .07$). At both locations, the interference group had shorter latencies to negative words versus positive words. The reverse was true for the facilitation group.

4.2. *Discussion*

The morphology of the ERPs generated by the emotional Stroop was remarkably similar to those studies of the traditional Stroop reported in the literature (West & Alain, 2000a). This was despite differences in stimuli content, participant populations, and response mode. There were, however, differences in particular ERP components. In previous Stroop studies, condition effects were concentrated in the middle (P3 and N4) and late (positive slow wave) components (Ilan & Polich, 1999; West & Alain, 1999). The current study, however, also produced findings in the early (P1–N1–P2–N2) components.

ERP amplitude is thought to reflect the amount of cognitive processing that an individual is allocating to a particular task (Rugg & Coles, 1995). Word category differences were found for early components (N1 and N2) presumed to reflect automatic attentional processes (Hillyard et al., 1994), indicating that early perceptual processing may activate latent biases in attentional allocation. This is in line with the clinical emotional Stroop literature, which argues that the idiographic pattern of Stroop interference is fueled by attentional biases that focus resources on self-referential environmental stimuli (McNally, 1995, 1996).

Late in the ERP wave, the positive slow wave was most pronounced for the negative words. West and Alain (2000a) have speculated that this slow wave may mark the additional processing needed by perceptual-level color information in order to counteract the ongoing cognitive processing of word meaning. Presumably, this task was more difficult for the negative words than for the positive. However, this presumed processing difference produced neither significant between-group differences nor reaction time differences between positive and negative words.

The latency data also revealed differences in how emotion words are processed. Only here did hemispheric differences appear to play an important role. As early as 50 ms after presentation (at P1), differences between word categories were apparent. In particular, negative words produced shorter ERP latencies in the right hemisphere versus the left hemisphere. This pattern reoccurred across components, late into the ERP wave, and coincides with a growing literature concerning the relationship between EEG asymmetry and emotional processing. In particular, right frontal EEG activation has been linked to negative or withdrawal emotions, while left frontal EEG activation has been linked to positive or approach emotions (Davidson & Fox, 1989; Fox, 1991). In addition, both the P3 and N4 components showed shorter latencies for negative words. This may indicate that the negative words held a "privileged" status and were afforded greater resources for stimulus evaluation, semantic processing, and response selection (Schack et al., 1999; West & Alain, 1999).

Acknowledgments

We would like to thank Stacey Barton-Bowers, Genevieve Erb, Cindy Polak, Ariana Shahinfar, and Katherine Wallner for their assistance in data collection. We would especially like to thank the parents of the children who participated and continue to participate in our studies. Portions of these data were presented at the Society for Psychophysiological Research Annual Meeting in October 2000. This research was partially supported by grants from the National Institute of Health (HD# 32666 and HD# 17899) to Nathan A. Fox.

Appendix A

Words used for the emotional Stroop in Studies 1 and 2.

Positive	Negative	Control
Study 1		
brave	afraid	apple
calm	alone	book
cool	awkward	cake
friends	bully	chair
glad	cry	grass
happy	embarrass	house
laugh	lonely	paper
liked	nervous	phone
loud	quiet	plane
popular	sad	plum
proud	scared	table
relax	shy	tree
share	tease	truck
smile	upset	water
sure	worry	wheel
Study 2		
awesome	afraid	cake
brave	alone	central
calm	anxious	chair
cheerful	awkward	clap
comfortable	bully	climb
confident	cry	daily
cool	embarrass	heavy
friend	failure	hollow
glad	frighten	melt
happy	hopeless	movie
laugh	lonely	nature
liked	loser	outdoor
perfect	miserable	pen
popular	nervous	phone
proud	rejected	pine
relax	sad	tape
share	scared	tasty
smile	shy	truck
success	tease	walking
talkative	upset	wheel
trust	worry	window

References

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4-18 and 1991 profile*. Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the child behavior checklist and revised child behavior profile*. Burlington, VT: University of Vermont.

Alexander, G. E., Crutcher, M. D., & De Long, M. R. (1990). Basal ganglia-thalamo-cortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Progress in Brain Research, 85*, 119–146.

Baker, S. C., Frith, C. D., & Dolan, R. J. (1997). The interaction between mood and cognitive function studied with PET. *Psychological Medicine, 27*, 565–578.

Barch, D. M., Carter, C. S., Perlstein, W., Baird, J., Cohen, J. D., & Schooler, N. (1999). Increased Stroop facilitation effects in schizophrenia are not due to increased automatic spreading activation. *Schizophrenia Research, 39*, 51–64.

Bauer, L. O., & Hesselbrock, V. M. (1999). Subtypes of family history and conduct disorder: Effects on P300 during the Stroop test. *Neuropsychopharmacology, 21*, 51–62.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The Counting Stroop: An interference task specialized for functional neuroimaging: Validation study with functional MRI. *Human Brain Mapping, 6*, 270–282.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.

Calkins, S. D., Fox, N. A., & Marshall, T. R. (1996). Behavioral and physiological antecedents of inhibited and uninhibited behavior. *Child Development, 67*, 523–540.

Carter, C. S., Mintun, M., & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: An H_2150 PET study of Stroop task performance. *Neuroimage, 2*, 264–272.

Coplan, R. J., Rubin, K. H., Fox, N. A., Calkins, S. D., & Stewart, S. L. (1994). Being alone, playing alone, and acting alone: Distinguishing among reticence and passive and active solitude in young children. *Child Development, 65*, 129–137.

Davidson, R. J. (2000). Cognitive neuroscience needs affective neuroscience (and vice versa). *Brain and Cognition, 42*, 89–92.

Davidson, R. J., & Fox, N. A. (1989). Frontal brain asymmetry predicts infants' response to maternal separation. *Journal of Abnormal Psychology, 98*, 127–131.

Davis, M. (1992). The role of the amygdala in conditioned fear. In J. P. Aggleton (Ed.), *The amygdala: Neurobiological aspects of emotion, memory, and mental dysfunction* (pp. 255–306). New York: Wiley–Liss.

Davis, M., & Shi, C. (1999). The extended amygdala: Are the central nucleus of the amygdala and the bed nucleus of the stria terminalis differentially involved in fear versus anxiety? In J. F. McGinty (Ed.), *Advancing from the ventral striatum to the extended amygdala: Implications for neuropsychiatry and drug use: In honor of Lennart Heimer. Annals of the New York Academy of Sciences* (Vol. 877, pp. 281–291). New York: New York Academy of Sciences.

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of the anterior cingulate cortex to behavior. *Brain, 118*, 279–306.

Diamond, A. (1990). The development and neural bases of memory functions as indexed by the AB and delayed response tasks in human infants and infant monkeys. *Annals of the New York Academy of Sciences, 608*, 267–317.

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development, 71*, 44–56.

Drevets, W. C., & Raichle, M. E. (1998). Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: Implications for interactions between emotion and cognition. *Cognition and Emotion, 12*, 353–385.

Duncan-Johnson, C. C., & Kopell, B. S. (1981). The Stroop effect: Brain potentials localize the source of interference. *Science, 214*, 938–940.

Foa, E. B., Ilai, D., McCarthy, P. R., Shoyer, B., & Murdock, T. (1993). Information processing in obsessive-compulsive disorder. *Cognitive Therapy and Research, 10*, 477–486.

Fox, N. A. (1991). If it's not left, it's right: Electroencephalograph asymmetry and the development of emotion. *American Psychologist, 46*, 863–872.

Fox, N. A., & Henderson, H. A. (1997). Emotion regulation: Distinguishing between subtypes of behavioral inhibition. In J. Gross (Chair), *Emotion regulation: Theory and research*. Symposium conducted at the 9th annual convention of the American Psychological Society, Washington, DC.

Fox, N. A., Henderson, H. A., & Marshall, P. J. (2001). The biology of temperament: An integrative approach. In C. A. Nelson, & M. Luciana (Eds.), *The handbook of developmental cognitive neuroscience* (pp. 631–645). Cambridge, MA: MIT Press.

Fox, N. A., Henderson, H. A., Rubin, K., Calkins, S. D., & Schmidt, L. A. (2001). Continuity and discontinuity of behavioral inhibition and exuberance: Psychophysiological and behavioral influences across the first 4 years of life. *Child Development, 72*, 1–21.

Fox, N. A., Rubin, K. H., Calkins, S. D., Marshall, T. R., Coplan, R. J., Porges, S. W., Long, J., & Stewart, S. (1995). Frontal activation asymmetry and social competence at four years of age. *Child Development, 66*, 1770–1784.

Fox, N. A., Schmidt, L. A., Calkins, S. D., Rubin, K. H., & Coplan, R. J. (1996). The role of frontal activation in the regulation and dysregulation of social behavior during the preschool years. *Development and Psychopathology, 8*, 89–102.

Frances, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

George, M. S., Ketter, T. A., Parekh, P. I., Rosinsky, N., Ring, H., & Casey, B. J. (1994). Regional brain activity when selecting a response despite interference: An H_2150 PET study of the Stroop and the emotional Stroop. *Human Brain Mapping, 1*, 194–209.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children $3\frac{1}{2}$–7 years old on a Stroop-like day-night task. *Cognition, 53*, 129–153.

Gilboa-Schechtman, E., Revelle, W., & Gotlib, I. H. (2000). Stroop interference following mood induction: Emotionality, mood congruence, and concern relevance. *Cognitive Therapy and Research, 24*, 491–502.

Glosser, G., & Goodglass, H. (1990). Disorders in executive control functions among aphasic and other brain-damaged patients. *Journal of Clinical and Experimental Neuropsychology, 12*, 485–501.

Goldman-Rakic, P. S. (1998). The prefrontal landscape: Implications of functional architecture for understanding human mentation and the central executive. In A. C. Roberts, & T. W. Robbins (Eds.), *The prefrontal cortex: Executive and cognitive functions* (pp. 87–102). New York: Oxford University Press.

Gotlib, I. H., & McCann, C. D. (1984). Construct accessibility and depression: An examination of cognitive and affective disorders. *Journal of Personality and Social Psychology, 47*, 427–439.

Green, E. J., & Barber, P. J. (1983). Interference effects in an auditory Stroop task: Congruence and correspondence. *Acta Psychologica, 53*, 183–194.

Harvey, A. G., Bryant, R. A., & Rapee, R. M. (1996). Preconscious processing of threat in posttraumatic stress disorder. *Cognitive Therapy and Research, 20*, 613–623.

Heaton, R. K. (1981). *Wisconsin card sorting test manual*. Odessa, FL: Psychological Assessment Resources.

Hillyard, S. A., Luck, S. J., & Mangun, G. R. (1994). The cueing of attention to visual field locations: An analysis with ERP recordings. In H. J. Heinze, T. F. Munte, & G. R. Mangun (Eds.), *Cognitive electrophysiology* (pp. 1–25). Boston: Birhäuser.

Holle, C., Neely, J. H., & Heimberg, R. G. (1997). The effects of blocked versus random presentation and semantic relatedness of stimulus words on response to a modified Stroop task among social phobics. *Cognitive Therapy and Research, 21*, 681–697.

Hope, D. A., Rapee, R. M., Heimberg, R. G., & Dombeck, M. J. (1990). Representations of the self in social phobia: Vulnerability to social threat. *Cognitive Therapy and Research, 14*, 177–189.

Ilan, A. B., & Polich, J. (1999). P300 and response time from a manual Stroop task. *Clinical Neurophysiology, 110*, 367–373.

Intili, R., & Tarrier, N. (1998). Attentional bias in morbid jealousy. *Behavioural and Cognitive Psychotherapy, 26*, 323–338.

Isenberg, N., Silbersweig, D., Engelien, A., Emmerich, S., Malavade, K., Beattie, B., & Leon, A. C. (1999). Linguistic threat activates the human amygdala. *Proceedings of the National Academy of Sciences, 96*, 10456–10459.

Jasper, H. H. (1958). The 10–20 electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10*, 371–375.

Jerger, S., Martin, R. C., & Pirozzolo, F. J. (1988). A developmental study of the auditory Stroop effect. *Brain and Language, 35*, 86–104.

Kagan, J. (1994). *Galen's prophecy*. New York: Basic Books.

Kagan, J., Reznick, J. S., Clarke, C., Snidman, N., & Garcia-Coll, C. (1984). Behavioral inhibition to the unfamiliar. *Child Development, 55*, 2212–2225.

Kagan, J., Reznick, S., & Snidman, N. (1987). The physiology and psychology of behavioral inhibition in children. *Child Development, 58*, 1459–1473.

Kagan, J., Reznick, S., Snidman, N., Gibbons, J., & Johnson, M. O. (1988). Childhood derivatives of inhibition and lack of inhibition to the unfamiliar. *Child Development, 59*, 1580–1589.

Kindt, M., Bierman, D., & Brosschot, J. F. (1997). Cognitive bias in spider fear and control children: Assessment of emotional interference by a card format and a single-trial format of the Stroop task. *Journal of Experimental Child Psychology, 66*, 163–179.

Lavy, E. H., van Oppen, P., & van den Hout, M. A. (1994). Selective processing of emotional information in obsessive-compulsive disorder. *Behaviour Research and Therapy, 32*, 243–246.

LeDoux, J. E. (1990). Fear pathways in the brain: Implications for a theory of the emotional brain. In P. F. Brain, & S. Parmigiani (Eds.), *Ettore Majorana international life sciences series* (Vol. 8, pp. 163–177) *Fear and defence*. Amsterdam: Harwood Academic Publishers.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Simon and Schuster.

LeDoux, J. E. (1998). Fear and the brain: Where have we been, and where are we going? *Biological Psychiatry, 44*, 1229–1238.

Liotti, M., Woldroff, M. G., Pérez III, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia, 38*, 701–711.

Logan, A. C., & Goetsch, V. L. (1993). Attention to external threat cues in anxiety states. *Clinical Psychology Review, 13*, 541–559.

Lovell, D. M., Williams, J. M. G., & Hill, A. B. (1997). Selective processing of shape-related words in women with eating disorders, and those who have recovered. *British Journal of Clinical Psychology, 36*, 421–432.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An intergrative review. *Psychological Bulletin, 109*, 163–203.

Martin, J. M., & Cole, D. A. (2000). Using the personal Stroop to detect children's awareness of social rejection by peers. *Cognition and Emotion, 14*, 241–260.

Martin, M., Horder, P., & Jones, G. V. (1992). Integral bias in naming of phobia-related words. *Cognition and Emotion, 6*, 479–486.

Masterman, D. L., & Cummings, J. L. (1997). Frontal-subcortical circuits: The anatomical basis of executive, social, and motivated behaviors. *Journal of Psychopharmacology, 11*, 107–114.

Mathews, A. M., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy, 31*, 57–62.

Mathews, A. M., & MacLeod, C. (1994). Cognitive approaches to emotions and emotional disorders. *Annual Review of Psychology, 45*, 25–50.

Mattia, J. I., Heimberg, R. G., & Hope, D. A. (1993). The revised Stroop color-naming task in social phobics. *Behaviour Research and Therapy, 31*, 305–313.

McClain, L. (1983). Stimulus-response compatibility affects auditory Stroop interference. *Perception and Psychophysics, 33*, 266–270.

McNally, R. J. (1995). Automaticity and the anxiety disorders. *Behaviour Research and Therapy, 33*, 747–754.

McNally, R. J. (1996). Cognitive bias in the anxiety disorders. In D. A. Hope (Ed.), *Nebraska symposium on motivation, 1995: Perspectives on anxiety, panic, and fear. Current theory and research in motivation* (pp. 211–250). Lincoln, NE: University of Nebraska Press.

McNally, R. J., Amir, N., Louro, C. E., Lukach, B. M., Riemann, B. C., & Calamari, J. E. (1994). Cognitive processing of ideographic emotional information in panic disorder. *Behaviour Research and Therapy, 32*, 119–122.

McNally, R. J., English, G. E., & Lipke, H. J. (1993). Assessment of intrusive cognition in PTSD: Used of the modified Stroop paradigm. *Journal of Traumatic Stress, 6*, 33–41.

Mecklinger, A., Kramer, A. F., & Strayer, D. L. (1992). Event related potentials and EEG components in a semantic memory search task. *Psychophysiology, 29*, 104–119.

Mogg, K., & Marden, B. (1990). Processing of emotional information in anxious participants. *British Journal of Clinical Psychology, 29*, 227–229.

Moradi, A. R., Neshat-Doost, H. T., Taghavi, M. R., Yule, W., & Dagleish, T. (1998). Performance of children of adults with PTSD on the Stroop colour-naming task. *Journal of Traumatic Stress, 12*, 663–671.

Moradi, A. R., Taghavi, M. R., Neshat-Doost, H. T., Yule, W., & Dagleish, T. (1999). Performance of children and adolescents with PTSD on the Stroop colour-naming task. *Psychological Medicine, 29*, 415–419.

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences, 87*, 256–259.

Rauch, S. L., Jenike, M. A., Alpert, N. M., Baer, L., Breiter, H. C., & Savage, C. R. (1994). Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using 15-O-labeled CO_2 and positron emission tomography. *Archives of General Psychiatry, 51*, 62–70.

Rauch, S. L., Savage, C. R., Alpert, N. M., Miguel, E. C., Baer, L., & Breiter, H. C. (1995). A positron emission tomography study of simple phobic symptom provocation. *Archives of General Psychiatry, 52*, 20–28.

Rauch, S. L., van der Kolk, B. A., Fisler, R. E., Alpert, N. M., Orr, S. P., & Savage, C. R. (1996). A symptom provocation study of posttraumatic stress disorder using positron emission tomography and script-driven imagery. *Archives of General Psychiatry, 53*, 380–387.

Richards, A., French, C. C., Johnson, W., Naparstek, J., & Williams, J. (1992). Effects of mood manipulation and anxiety on performance on an emotional Stroop task. *British Journal of Psychology, 83*, 479–491.

Riemann, B., & McNally, R. J. (1995). Cognitive processing of personally relevant information. *Cognition and Emotion, 9*, 325–340.

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

Rothbart, M. K., Posner, M. I., & Hershey, K. L. (1995). Temperament, attention, and developmental psychopathology. In D. Cicchetti, & D. Cohen (Eds.), *Developmental psychopathology, Vol. 1: Theory and methods. Wiley series on personality processes*. New York: Wiley.

Rowe, D. C., & Plomin, R. (1977). Temperament in early childhood. *Journal of Personality Assessment, 41*, 150–156.

Rubin, K. H. (1989). The Play Observation Scale (POS). Waterloo, Canada: University of Waterloo.

Rubin, K. H., Coplan, R. J., Fox, N. A., & Calkins, S. D. (1995). Emotionality, emotion regulation, and preschoolers' social adaptation. *Developmental Psychopathology, 7*, 49–62.

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

Rugg, M. D., & Coles, M. G. H. (1995). The ERP and cognitive psychology: Conceptual issues. In M. D. Rugg, & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 27–39). Oxford: Oxford University Press.

Schack, B., Chen, A. C. N., Mescha, S., & Witte, H. (1999). Instantaneous EEG coherence analysis during the Stroop task. *Clinical Neurophysiology, 110*, 1410–1426.

Schmidt, L. A., & Fox, N. A. (1998). Fear-potentiated startle responses in temperamentally different human infants. *Developmental Psychobiology, 32*, 113–120.

Schmidt, L. A., Fox, N. A., Rubin, K. H., Sternberg, E., Gold, P. W., Smith, C., & Schulkin, J. (1997). Behavioral and neuroendocrine responses in shy children. *Developmental Psychobiology, 30*, 127–140.

Schmidt, L. A., Fox, N. A., Schulkin, J., & Gold, P. W. (1999a). Behavioral and psychophysiological correlates of self-presentation in temperamentally shy children. *Developmental Psychobiology, 35*, 119–135.

Schmidt, L. A., Fox, N. A., Sternberg, E. M., Gold, P. W., Smith, C. C., & Schulkin, J. (1999b). Adrenocortical reactivity and social competence in seven-year-olds. *Personality and Individual Differences, 26*, 977–985.

Schneider, S., Unnewher, S., Florin, I., & Margraf, J. (1992, September). *Cognitive characteristics of children of panic patients*. Paper presented at the International Congress of Psychology, Brussels.

Segal, Z. V., Truchon, C., Horowitz, L. M., Gemar, M., & Guirguis, M. (1995). A priming methodology for studying self-representation in major depressive disorder. *Journal of Abnormal Psychology, 104*, 205–213.

Shimada, H. (1990). Effects of auditory presentation of words on color naming: The intermodal Stroop effect. *Perceptual and Motor Skills, 70*, 1155–1161.

Stifter, C. A., & Fox, N. A. (1990). Infant reactivity and regulation: Physiological correlates of newborn and five-month temperament. *Developmental Psychology, 26*, 582–588.

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

Taghavi, S. M. R. (1996). *Cognitive aspects of clinical anxiety in children and adolescents*. Unpublished dissertation. University of London.

van Honk, J., Tuiten, A., van den Hout, M., Koppeschaar, H., Thijssen, J., de Han, E., & Verbaten, R. (2000). Conscious and preconscious selective attention to social threat: Different neuroendocrine response patterns. *Psychoneuroendocrinology, 25*, 577–591.

Vogt, B. A., Nimchinsky, E. A., Vogt, L. J., & Hof, P. R. (1995). Human cingulate cortex: Surface features, flat maps, and cytoarchitecture. *Journal of Comparative Neurology, 359*, 490–506.

Wallner, K. E. (1996). *Executive functioning and its relation to planning skill in seven-year-old children*. Unpublished dissertation. University of Maryland.

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology, 77*, 97–108.

Welsh, M. C. (1991). Rule guided behavior and self-monitoring on the Tower of Hanoi disk-transferring task. *Cognitive Development, 6*, 59–76.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7*, 131–149.

West, R., & Alain, C. (1999). Event-related neural activity associated with the Stroop task. *Cognitive Brain Research, 8*, 157–164.

West, R., & Alain, C. (2000a). Effects of task context and fluctuations of attention on neural activity supporting performance of the Stroop task. *Brain Research, 873*, 102–111.

West, R., & Alain, C. (2000b). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology, 37*, 179–189.

Whalen, P. J., Bush, G., McNally, R. J., Wilhelm, S., McInerney, S. C., Jenike, M. A., & Rauch, S. L. (1998). The emotional counting Stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry, 44*, 1219–1228.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120*, 3–24.