CHILDREN’S “CATASTROPHIC RESPONSES” TO NEGATIVE FEEDBACK ON CANTAB’S ID/ED SET-SHIFTING TASK: RELATION TO INDICES OF A DEPRESSIVE TEMPERAMENTAL STYLE

Monica LUCIANA*1,2, Megan R. GUNNAR2,3, Elysia P. DAVIS4, Charles A. NELSON2,3, Bonny DONZELLA2,3

1 Department of Psychology, University of Minnesota, USA
2 Center for Neurobehavioral Development, University of Minnesota, USA
3 Institute of Child Development, University of Minnesota, USA
4 Department of Psychiatry and Human Behavior, University of California, Irvine, USA

ABSTRACT

This study assessed whether a specific form of motivational impairment, over-sensitivity to perceived failure, described previously in depressed adults, would impact healthy children’s performance on neuropsychological tasks as a function of individual differences in negative affectivity. Healthy children completed the Cambridge Neuropsychological Test and Automated Battery (CANTAB) as part of a large-scale study of cognitive development. The tendency to respond to perceived failure with subsequent item failure was calculated for each child on the basis of his/her performance on the CANTAB’s ID/ED set-shifting task, during which trial-by-trial feedback is provided. Children were divided into those with low versus high tendencies to react to failure with subsequent item failure. One year later, their parents completed child temperament ratings, using Rothbart’s Child Behavior Questionnaire (CBQ) as part of another study where cortisol was also measured.

Children with increased tendencies to exhibit heightened responses to failure on the ID/ED set-shifting task were rated by their parents as higher in sadness and slightly higher in overall Negative Affectivity. These children performed worse on several CANTAB subtests, including memory span, pattern recognition memory, and set shifting. Relations to daily cortisol rhythms were examined but did not yield strong effects.

Abnormally sensitive responses to negative feedback have been discussed as a trait marker of affective disorder. These findings suggest that this motivational style might impact children as a function of temperamental characteristics that might lead to vulnerability to later internalizing psychopathology. Findings are discussed in relation to neurobehavioral

* Corresponding address:
Monica Luciana, Department of Psychology, 75 East River Road, University of Minnesota, Minneapolis, MN 55455, USA
Email: lucia003@tc.umn.edu

Iunie 2005 • Cognitie, Creier, Comportament 343
models of feedback processing in affective disorder and developmental psychopathology.

**KEY-WORDS:** temperament, depression, cognition, negative affect, children

Cognitive theories of depression and anxiety are among the most influential psychological models advocated by clinicians to explain internalizing disorders (Beck, 1972; 1979; Crews & Harrison, 1995). These models posit that depression and anxiety are exacerbated by negative self-directed attributions that bias the manner in which vulnerable individuals perceive themselves, their interactions with others, and outcomes of their behaviors across situations. They also imply that individuals who have been previously depressed, but are currently remitted, remain vulnerable to relapse due to the trait-like stability of negative biases, which are automatic in nature, involuntary, and perseverative across situations (Beck, 1972). Cognitive-behavioral interventions based on these theories are perhaps the most successful non-pharmacological treatments for both depression and anxiety, particularly in children and adolescents (Compton et al., 2004).

Cognitive functions have been examined from various perspectives in internalizing psychopathology through the use of psychological inventories (e.g., the Beck Depression Inventory: Beck, Ward, Mendelsohn, Mock, & Erbaugh, 1961) and through direct measures of discrete abilities as measured by neuropsychological tests (reviewed by Elliott, Sahakian, Herrod, Robbins & Paykel 1997). These investigations have revealed, on the one hand, that depressed individuals tend to feel pessimistic or hopeless about the likelihood of positive events occurring in the future (Abramson, Seligman & Teasdale, 1978; Crews & Harrison, 1995). On the other hand, brain-based evaluations of cognition have revealed the presence of neuropsychological deficits in individuals with depression and/or anxiety (George et al., 1997; Trichard et al., 1995). These deficits include impairments in attention, memory, and executive functions under conditions where high effort is demanded. Despite some degree of consistency in the types of impairments reported, no single neurocognitive dysfunction has been described to characterize all individuals with internalizing psychopathology, and links between these deficits and negative self-attributitional biases that have traditionally been conceptualized as more “psychological” in nature have not been proposed.

Recently, it has been suggested that the general affective demeanor of anxious and/or depressed individuals might undercut their performance across structured cognitive tasks regardless of task content, perhaps contributing to failures observed during neuropsychological testing, especially when direct feedback is provided to participants in a trial-by-trial manner. For example, Beats, Sahakian and Levy (1996) administered a computerized version of the Tower of London planning task to elderly depressed individuals and to healthy controls. It was observed that although the groups were equivalent in their overall levels of

---

344 Children’s responses to negative feedback
perfect problem solutions, individuals in the depressed group tended to deteriorate in their efficiency of performance after making mistakes within a given problem. This basic effect was replicated and extended by Elliott et al. (1996) who administered the Tower of London and a Delayed Match to Sample task to depressed adults versus controls and found that depressed individuals not only performed worse than control subjects overall, but they also exhibited decrements in performance following error trials. Notably, this performance pattern was specific to depressed individuals and was not evident in individuals with other psychiatric disorders (i.e., schizophrenia) or in those with neurological impairments (i.e., individuals with Parkinson’s disease or temporal lobe lesions: Elliott et al., 1997). Moreover, previously but not currently depressed individuals who were tested while they were in states of recovery also showed oversensitivity to perceived failures (Elliott et al., 1997) suggesting that the phenomenon is trait versus state-based. Elliott et al. (1996, 1997) have therefore suggested that oversensitivity to perceived failure is a motivationally-driven trait-based disposition that renders depressed individuals vulnerable to cognitive dysfunction when cognition is measured through the use of structured neuropsychological tasks.

These findings suggest novel methodologies for assessing cognition-emotion interactions in the laboratory, because they call for a focus on the motivational processes that individuals use in the course of problem-solving versus strict adherence to categorical evaluations of performance (i.e., correct versus incorrect scores on given problems). It may be that a qualitative analysis of motivational processes is more predictive of which individuals are vulnerable to specific types of pathology. It might also be the case that dysfunctions in common brain regions account for both the distorted responses that depressed and anxious people exhibit following perceived failure as well as their difficulties with attention, memory, and executive function measures. In order to begin to address this issue, it is important to establish an association, on a behavioral level, between “catastrophic” responses and/or distorted cognition in the course of testing, measures of depressogenic traits, and neuropsychological performance. When applied to developmental samples, this type of analysis may contribute in unique ways to our understanding of internalizing disorders and their etiology. Prospective developmental studies would allow us to demonstrate that high-risk individuals who are vulnerable to pathology, but who haven’t yet expressed it, nonetheless demonstrate biased cognition. However, similar but more subtle relations might also characterize healthy (apparently low-risk) individuals who vary in temperamental traits thought to contribute to internalizing tendencies. These traits would include diminished levels of positive affect, which characterizes individuals with depressive tendencies, and high levels of negative affect, which characterizes both individuals with depression as well as those with anxiety (Clark, Watson & Mineka, 1994; Watson & Clark, 1995; Watson & Tellegen, 1985).

Accordingly, this project sought to apply the general findings of Elliott and colleagues (1996, 1997) to a developmental sample of children who were free of...
psychopathology. A low risk sample of children was tested using selected subtasks from the CANTAB neuropsychological testing battery (Sahakian & Owen, 1992), and their responses to measures where performance feedback was provided were tabulated in order to identify individual children who showed a relative tendency to respond to failure with subsequent failure (i.e., exhibiting a so-called “catastrophic response to failure”: Beats et. al., 1996). This tendency was then examined in relation to each child’s temperament as reported by a parent, with a specific focus on aspects of temperament (negative affectivity) that would suggest vulnerability to either anxiety or depression. As an additional dimension of temperament, we also examined whether children high in “catastrophic” responses to failure exhibited high levels of salivary cortisol, a recognized marker of stress-reactivity (Stansbury & Gunnar, 1994). It has been reported recently that generation of a highly variable cortisol rhythm from day to day or within a single day might be a risk factor for internalizing disorders (Halligan, Herbert, Goodyer, & Murray, 2004; Goodyer, Herbert, Tamplin, & Altham, 2000; Peeters, Nicholson, & Berkhof, 2004).

Methods

Study sample
The sample consisted of children whose parents’ names are kept on file by the University of Minnesota’s Institute of Child Development (ICD). When their children were born, parents were identified by ICD staff through the use of published birth announcements. Parents were then sent letters by ICD administrative staff and asked if they would be interested in participating in child development research. The names, addresses, and telephone numbers of interested participants are maintained in an ongoing research database. Children between the ages of four and eight years (n=235) were recruited from this database. Parents were randomly contacted by telephone and invited to participate in a study of children’s memory and problem-solving skills. The children and their parents made one visit to our laboratory where informed consent was obtained. A brief medical history was obtained from the parent through use of a structured interview questionnaire, after which the child was tested on the CANTAB. Cognitive development in this sample of children has been discussed elsewhere (Luciana & Nelson, 1998, 2002). Overall, this sample consisted of middle-to-upper middle class families, largely Caucasian, all of whom were native English speakers.

CANTAB subtasks
The details of the CANTAB task administration have been presented in extensive detail elsewhere (Luciana and Nelson, 1998) and will not be thoroughly described here. The following tests from the CANTAB (DOS version 1.71) were administered to study participants, in this order:

1) Motor Screening Test. Participants are instructed to touch X’s that appear on a computer monitor. Accuracy and response latency are recorded. This

346 Children’s responses to negative feedback
task is not discussed further, as it primarily served to habituate children to the computerized testing format.

(2) **Spatial Memory Span.** This test measures memory for a sequence of visually-presented information. Participants view a lighted sequence and must reproduce the sequence by touching items in the same order that they were originally illuminated. This task is a computer analog of the Corsi block task (Milner, 1971) and yields a measure corresponding to the length of the participant’s visual memory span.

(3) **Self-Ordered Search.** This test is an adaptation of the radial arm maze task and measures the participant’s ability to conduct an organized search of items in order to obtain tokens hidden inside of the items. Search complexity varies from searches of 2-items to searches of 8-items. If one searches a location where s/he has previously found a token, a “forgetting error” is recorded. Additionally, a strategy score is recorded that corresponds to the tendency to search through available items in an organized fashion. The test yields error scores for searches of 2, 3, 4, 6,and 8 items as well as a strategy score computed from responses to the six and eight-item searches.

(4) **Tower of London Planning Task.** This version of the Tower of London Planning Task (Shallice, 1982) is referred to as the Stockings of Cambridge. In this task, participants must solve problems by moving colored circles between three locations in a prescribed number of moves. Problems are graded in difficulty, involving 2, 3, 4, and 5 moves. For each problem set, the average number of moves and response latency are recorded.

(5) **Intradimensional/Extradimensional Set-Shift.** This task involves using computer-provided feedback to learn response contingencies (Downes et al., 1989; Owen, Roberts, Polkey, Sahakian, & Robbins, 1991). There are nine stages of the task that proceed from a very simple first stage (simple discrimination between two abstract patterns) to the eighth and ninth stages, involving between category set shifting and between-category reversals. The most frequently reported dependent variable is the number of stages reached on the task, although error rates and trials-to-criterion within each stage are also recorded.

(6) **Pattern Recognition Memory.** This task is a Delayed Match to Sample Task. Participants view a series of abstract patterns on a computer screen. Then they see pairs of stimuli, one of which was previously seen. The task is to select the familiar stimulus. The percentage of correct responses, out of a total of 24 stimuli, is the primary variable of interest.

The administration of these six measures takes approximately seventy-five minutes, depending upon the age and stamina of the child. This particular configuration of tasks was selected with the goal of including subtests that had, at the time of our assessment, received the most extensive validation in adult studies.

Of the tasks listed above, only two provide feedback to the test-taker. These two are the ID/ED set-shifting task and the Pattern Recognition Memory Task. To evaluate how failure impacts subsequent performance, it would be preferable to use a task that requires responses to complex stimuli but where easy
versus harder items are randomly interspersed. Neither the ID/ED set-shifting task nor the Pattern Recognition memory task is ideal for this purpose. The Pattern Recognition memory task uses relatively simple stimuli, a small number of trials, and a forced-choices response format. In our past analyses with this sample (Luciana & Nelson, 1998), we noted that children reach ceiling levels of performance on this task at a relatively early age. In addition, not every child in our sample was tested on the Pattern Recognition Memory Task. The ID/ED set-shifting task has the benefit of being more difficult overall, leading to more errors and a more reliable assessment of motivational factors that might impact error rate. However, the task progresses in an orderly fashion from one stage to the next, where the stages theoretically increase systematically in difficulty. Because the ID/ED set-shifting task includes more total items and because performance can be analyzed in a more sophisticated manner, we chose to focus on it for subsequent analyses of individual differences in probabilities of failure following previous failure.

Following the methodology of Elliott et al. (1996, 1997), the conditional probability of a participant failing a problem given that they had failed the previous problem (CP-F) was calculated. As discussed by Elliott et al (1997), Bayes’ theorem states that the probability of event B occurring given that event A has already occurred (labeled p(A | B)) is the probability of both events occurring [p(AB)] divided by the probability of A occurring. If events A and B are independent, this is the same as the probability of B occurring. For our purposes, event A is failing a trial (trial x) and event B is failing the subsequent trial (trial x+1). The conditional probability of failure is thus defined as:

\[
p(x+1\text{wrong} | \text{wrong}) = \frac{p(x\text{and} x+1\text{wrong})}{p(x\text{wrong})}
\]

This equation is equivalent to dividing the number of errors made immediately following an error by the total number of errors made by that participant (Elliott et al., 1997). CP-F values were computed for all participants and evaluated against temperament scores as assessed from a dimensional perspective.

**Temperament measurement**

Approximately one year after the sample was tested on the CANTAB, their parents were invited to participate in an additional study (Davis, Donzella, Krueger, & Gunnar, 1999). This subsequent study had the purpose of measuring stress reactivity in children during the first week of the school year and involved the collection of salivary cortisol measures from children as well as parental reports on each child’s temperament using Rothbart’s Children’s Behavior Questionnaire (CBQ: Ahadi, Rothbart, & Ye, 1993). The CBQ is a 195-item parent report measure of temperament that yields three higher-order factors termed Surgency/Extraversion, Negative Affectivity, and Effortful Control. These factors
are derived from weightings of the following fourteen temperament subscales: activity level, anger/frustration, attentional focusing, discomfort, soothability, fear, high-intensity pleasure, impulsivity, inhibitory control, low intensity pleasure, perceptual sensitivity, sadness, shyness, and smiling/laughter. A total of 89 children who participated in this subsequent study had previously generated usable data on the ID/ED set-shifting task. Their demographics are presented in Table 1. Table 2 includes descriptive statistics related to each CBQ super-factor, its subscales, and within-sample alpha levels for each scale calculated from data provided by these 89 participants.

Collection of cortisol data

Families of participating children were also asked to collect cortisol samples through saliva collection kits. Briefly, this procedure involved giving the child a small amount (~1/8 teaspoon) of Kool-Aid crystals to stimulate saliva secretion. The child was instructed to simultaneously place a cotton dental roll in his/her mouth until the roll was saturated. The saturated roll was then placed by the parent into an empty syringe and squeezed so that all liquid (saliva) was collected into a small plastic vial. Collected samples were stored in the families’ refrigerators until all sampling for that child was completed. The families then mailed the samples to the investigator’s laboratory where they were stored at -20°C until assayed. Additional details on the methodology related to cortisol analyses can be found in Davis et al. (1999).

Parents collected saliva from their children using this technique on four separate days, three times per day. The three times were (i) in the morning, shortly after the child awoke but before breakfast, (ii) at noon but before lunch, and (iii) in the evening within 30 minutes of bedtime (Davis et al., 1999). The four collection days were the first and fifth days of school and two weekend days approximately one month after the start of a new school year. For the purpose of this analysis, only the weekend cortisol values were examined, since these values were available for a maximal number of children (regardless of age and/or school status). Although disruptions in cortisol secretion would be expected to be maximal under novel or psychologically-challenging conditions (Kirschbaum & Hellhammer, 1994), it is also the case that anxious and inhibited individuals are more likely to exhibit elevations in cortisol secretion under conditions that others do not perceive to be psychologically difficult (Rosen & Schulkin, 1998).

Examination of cognition-temperament interactions

We were interested in associating specific patterns of cognitive performance on the CANTAB with specific features of temperament as measured by the CBQ. Specifically, we questioned whether children who exhibited extreme reactions to cognitive failures on the CANTAB (high CP-F scores on the ID/ED set-shifting task) were also those rated high on measures of anxiety, sadness, and other components of negative affect on the CBQ.
Table 1. Characteristics of Participant Sample

<table>
<thead>
<tr>
<th>Age Group (Years)</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=</td>
<td>13</td>
<td>19</td>
<td>19</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Male</td>
<td>6</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td># Female</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Caucasian</td>
<td>83</td>
<td>87</td>
<td>98</td>
<td>97</td>
<td>98</td>
</tr>
<tr>
<td>Parent Years of Education [Mean (SD)]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>15.9 (2.6)</td>
<td>15.4 (1.8)</td>
<td>16.4 (2.2)</td>
<td>15.8 (1.7)</td>
<td>16.1 (2.2)</td>
</tr>
<tr>
<td>Father</td>
<td>16.6 (2.0)</td>
<td>16.1 (2.1)</td>
<td>16.6 (2.3)</td>
<td>15.9 (2.1)</td>
<td>16.4 (1.9)</td>
</tr>
<tr>
<td>Conditional Probability of Failure Scores (mean +/- S.D.)</td>
<td>0.34 (.19)</td>
<td>0.39 (.17)</td>
<td>0.29 (.10)</td>
<td>0.19 (.14)</td>
<td>0.22 (.14)</td>
</tr>
</tbody>
</table>

Table 2. CBQ Subscales and Factors and their reliabilities in this sample

<table>
<thead>
<tr>
<th>Scale</th>
<th>Description</th>
<th>Mean (+/- S.D.)</th>
<th>Within-Sample Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative Affectivity</td>
<td>Negative Engagement; Anxiety; Fear</td>
<td>-0.00 (0.51)</td>
<td>0.77</td>
</tr>
<tr>
<td>Anger/Frustration</td>
<td>Negative affect when goals are blocked</td>
<td>4.58 (0.78)</td>
<td>0.78</td>
</tr>
<tr>
<td>Discomfort</td>
<td>Negative affect related to high sensory stimulation</td>
<td>3.97 (0.85)</td>
<td>0.78</td>
</tr>
<tr>
<td>Soothability</td>
<td>Rate of recovery from peak distress, excitement, arousal</td>
<td>0.23 (0.06)</td>
<td>0.82</td>
</tr>
<tr>
<td>Fear</td>
<td>Unease, worry, nervousness related to anticipated pain/threat</td>
<td>3.56 (0.88)</td>
<td>0.75</td>
</tr>
<tr>
<td>Sadness</td>
<td>Negative affect, lowered mood, and energy related to suffering, disappointment, or loss</td>
<td>4.26 (0.72)</td>
<td>0.68</td>
</tr>
<tr>
<td>Surgency</td>
<td>Extraversion, Positive Emotionality</td>
<td>-0.01 (0.82)</td>
<td>0.84</td>
</tr>
<tr>
<td>Activity Level</td>
<td>Level of gross motor activity including rate and extent of locomotion</td>
<td>4.59 (0.86)</td>
<td>0.84</td>
</tr>
<tr>
<td>High intensity pleasure</td>
<td>Pleasure &amp; enjoyment related to situations that are intense, complex and novel</td>
<td>4.97 (0.98)</td>
<td>0.88</td>
</tr>
<tr>
<td>Impulsivity</td>
<td>Speed of response initiation</td>
<td>4.37 (0.88)</td>
<td>0.82</td>
</tr>
<tr>
<td>Shyness</td>
<td>Slow or inhibited approach in situations involving novelty or uncertainty</td>
<td>3.68 (1.29)</td>
<td>0.94</td>
</tr>
<tr>
<td>Effortful Control</td>
<td>Attentional control over other processing resources</td>
<td>-0.00 (0.68)</td>
<td>0.71</td>
</tr>
<tr>
<td>Attentional Focusing</td>
<td>Tendency to maintain focus and to stay on-task</td>
<td>4.68 (0.92)</td>
<td>0.79</td>
</tr>
<tr>
<td>Inhibitory Control</td>
<td>Capacity to plan and to suppress inappropriate approach responses</td>
<td>4.96 (0.81)</td>
<td>0.84</td>
</tr>
<tr>
<td>Perceptual Sensitivity</td>
<td>Amount of detection of slight, low intensity stimuli from the external environment</td>
<td>5.06 (0.74)</td>
<td>0.80</td>
</tr>
<tr>
<td>Smiling/Laughter</td>
<td>Amount of positive affect in response to changes</td>
<td>5.60 (0.59)</td>
<td>0.77</td>
</tr>
<tr>
<td>Low intensity pleasure</td>
<td>Pleasure &amp; enjoyment related to situations involving low stimulus intensity, complexity, and novelty</td>
<td>5.57 (0.66)</td>
<td>0.79</td>
</tr>
</tbody>
</table>

For descriptive statistics, values presented for each subscale are based on raw scores. Those presented for each of the three superfactors are based on within-sample z-scores.

Children’s responses to negative feedback
The CBQ subscales most relevant to this hypothesis include anger/frustration, discomfort, soothability, fear, and sadness. These subscales cohere to form the CBQ’s Negative Affectivity superfactor. We did not expect subscales that form the Surgency and Effortful Control factors to be related to CP-F scores.

Results

Data were analyzed using the Statistical Package for the Social Sciences, Windows version 11.5. Alpha levels less than 0.05 were viewed as statistically significant. Values between 0.05 and 0.10 are described as statistical trends.

Conditional Probability of Failure (CP-F) scores

CP-F scores ranged in the sample from 0 to 0.71 and were not randomly distributed by age. When children were grouped into five age groups (ages 4, 5, 6, 7 and 8 years based on when the CANTAB was administered), CP-F scores differed significantly by age group \( F(4,88)=5.59, p<.01 \). Four-to-six year-olds did not differ from another. Seven and eight-year-olds did not differ from one another or from six year-olds but both had lower CP-F scores than five year-olds.

In addition, males and females (regardless of age) differed in their CP-F scores, with females exhibiting higher values \( F(1,88)=4.62, p<.05 \). Females had a mean (+/- S.D.) score of 0.32 ± 0.16; males had a value of 0.25 ± 0.16. There was no significant interaction between age and gender in the magnitude of CP-F scores.

Associations between CP-F scores and cognitive performance

To examine cognition-CP-F score relations, CP-F scores were first correlated with variables representing each CANTAB subtask. If a significant bivariate association was found, partial correlations, controlling for chronological age (as expressed in months) were computed to determine if the observed relationship survived age correction. Gender differences were also considered when significant bivariate associations were observed.

CP-F scores did not relate to performance on the Tower of London or Spatial Self-Ordered Search tasks. However, CP-F scores were marginally associated with spatial memory spans \( r= -0.20, p<.10 \) indicating that children with lower memory spans had detrimental responses to perceived failure. This association was significant in females but non-significant in males.

CP-F scores were also marginally correlated with pattern recognition memory scores \( r= -0.22, p<.10 \) suggesting that recognition memory for patterned stimuli was associated with higher sensitivity to perceived failure. There were no gender differences in this relation.

Finally, CP-F scores were more strongly related to error rates on several stages of the ID/ED set-shifting task, notably stages 1 (Simple Discrimination: \( r= 0.23, p<.05 \)); stage 3 (Compound Discrimination: \( r=0.34, p<.01 \)); stage 6 (Intradimensional Shift: \( r=0.47, p<.01 \)); stage 7 (Intradimensional Reversal: \( r=0.56, p<.01 \)); and stage 8 (Extradimensional Shift: \( r=0.42, p<.01 \)). In all cases, high likelihoods of failures following failures were associated with poor
performance within a stage. Indeed, the overall stage reached on the task was inversely related to this response tendency \((r= -0.68, p<.01)\). These associations were significant for both genders but were more strongly observed in females.

Of these significant or marginally-significant associations, the only ones to survive age-correction were associated with the ID/ED set-shifting task in terms of the overall stage reached and errors made on stages 6, 7, and 8.

**Associations between CP-F scores and temperament ratings**

We next considered whether CP-F scores related to parent-rated temperament scores on the CBQ. First, within-sample z-scores were calculated for each CBQ subscale. Then factor scores were computed on the basis of these z-score transformations.

As with the cognitive data, several age-related differences in parent ratings of temperament were observed. These included the following subscales and/or factors: **Effortful Control, Soothability, Low Intensity Pleasure, Shyness** and **Smiling/Laughter**. In all cases, seven year-olds differed from various other age groups, although no consistent differences were observed across scales in terms of which other age groups were maximally different from 7-year-olds.

When CP-F scores were correlated with temperament scores in the sample as a whole, they were found to be significantly and negatively associated only with the Soothability subscale \([r=-0.25, p<.05]\). The other subscales and the superfactors of Negative Affectivity, Surgency, and Effortful Control did not relate continuously to CP-F scores. Soothability reflects a child’s rate of recovery from extreme distress, excitement, or general arousal (Ahadi et al., 1993). While this finding would seem to be very much in keeping with the study’s overall hypothesis, when age was controlled, the association between Soothability and CP-F scores was no longer significant. Whether this association was similarly observed in each gender was then examined.

When males and females were considered separately, the Soothability-CP-F score association was noted exclusively in females \([r= -0.37, p<.05]\) but not in males \([r=0.15, \text{ns}]\). This association was still significant in females even after age was controlled. Males exhibited a different pattern of CP-F/temperament association. In males, high CP-F scores were associated with high activity levels \([r=0.32, p<.05]\) and at a trend level with high levels of high-intensity pleasure \([r=0.25, \text{ns}]\). The association with high-intensity pleasure remained significant when age was controlled.

Overall, from a dimensional perspective, these findings suggest possible inter-relations between “catastrophic responses to failure”, absolute levels of cognitive performance on memory and set-shifting tasks (especially those within which feedback is provided), and a child’s ability to be soothed in times of distress. These inter-relations appear to better-characterize females versus males. However, maximal confidence in these interpretations is confounded in this sample by the possibility that age accounts for several of the observed associations.
Categorical examination of CP-F / temperament associations

Accordingly, we next considered whether dividing the sample based on more extreme CP-F scores and restricting it to achieve less extensive variation in age would yield a more interpretable pattern of findings. To eliminate substantial age-related variability (and because seven year-olds in the sample seemed to be deviant from the rest of the group in terms of temperament scores), only 4-to-6 year-olds were considered for subsequent analyses, yielding a total sample of 51 children. Gender was not considered in these analyses because of small sample sizes.

Within this group, the sample was divided into children who exhibited a less-than-50% probability of failure-following-failure on the ID/ED set-shifting task (n=43) and those who exhibited a greater-than-50% probability of failure-following-failure on the ID/ED set-shifting task (n=8). A value of 50% was selected, because participants are instructed to select one of two response options on each trial of the ID/ED set-shifting task, and thus have a 50% probability of success on any given trial. The mean CP-F score in the sample as a whole was 0.29 +/- 0.16. Use of the 0.50 cutoff to create high versus low score groups restricted the high end of the CP-F variable to the upper 10% of the total distribution.

Overall cognitive task performance differences. We then revisited whether CANTAB performance differed generally between the two CP-F groups. In terms of the ID/ED set-shifting task, the maximum stage reached on the task, which represents its primary outcome variable (Luciana & Nelson, 1998) was compared between groups, yielding only a marginally significant difference [F(1,50)=3.52, p<.10]. Individuals with low CP-F scores reached stage 7 of the task (mean=7.25 +/- 2.22: the Intradimensional Shift-Reversal stage) on average, while those higher in CP-F scores reached stage 5 (mean=5.63 +/- 2.4: the Compound Discrimination-Reversal stage).

The groups did not differ in their Self-Ordered Search performance, lengths of memory span, average moves to complete Tower of London problems, or percentage correct on the Pattern Recognition memory task.

Temperament Scale Differences. In terms of temperament scores, the groups differed on the CBQ’s Negative Affectivity factor at a trend level [F(1,48)=3.19, p<.10] but specifically and more significantly on the Sadness subscale [F(1,48)=4.13, p<.05]. Children with higher CP-F scores were rated one year later as higher in negative affective tendencies and higher in sadness by their parents. These findings are graphed in Figure 1. There were no other significant differences in temperament subscale ratings between the two groups.

Cortisol differences in children high versus low in CP-F. Cortisol values for each weekend day and at each time point were averaged to yield three values: an average morning value, an average noon value, and an average evening value. These findings should be considered exploratory, since only forty children had all of these values available for analysis. These values were entered into a repeated measures ANOVA with three levels of Time as the within-subjects variable and CP-F group as the between-subjects variable. Findings are graphed in Figure 2. The analysis yielded a main effect of Time [F(2,76)=99.93, p<.001], no significant
Figure 1. A: Parent-rated sadness on the CBQ in children who were high (>50% tendency) versus low (< 50% tendency) in responding to failure with subsequent failure on the ID/ED set-shifting task. Sadness ratings are within-sample z-scores. B: Parent-rated Negative Affectivity on the CBQ, expressed in z-score units, in children high versus low in responding to failure with subsequent failure. Values represent means +/- one standard deviation. See text for statistical comparisons.
effect of CP-F Group $[F(1,38)=1.86, p<.20]$, and no significant Time x CP-F Group interaction $[F(2,76)=1.72, p<.20]$. Consistent with expectations, cortisol values significantly decreased throughout the day in both groups. The slope of this decrease (morning-to-noon as well as noon-to-evening) did not differ between groups. As suggested by Figure 2, there was a tendency (though non-significant: $F(1,39)=2.45, p<.15$) for morning cortisol values to be slightly higher in children with higher CP-F scores. The average weekend cortisol values at each time point were unrelated to CBQ Negative Affectivity and Sadness ratings. Variability in cortisol levels in the morning and evening and from morning-to-evening was also examined but did not yield significant effects.

![Figure 2](image-url)

**Figure 2.** Weekend cortisol levels (morning, noon, and evening) averaged across two days of sampling for children high versus low in CP-F scores. Values represent means +/- one standard deviation.

**Discussion**

In this report, we attempted to associate a specific cognitive response style, termed by others as a “catastrophic response to failure” (Beats et al., 1996) with parental ratings of negative emotional traits and salivary cortisol levels in a low-risk sample of 4-to-8 year-old children. We found that as children show a tendency to be increasingly derailed by perceived failure in the course of a set-shifting task, performance is also worse on measures of memory span and recognition memory.
These children also exhibit relatively poor set-shifting performance overall. In addition, these children are rated by their parents as less soothable during times of high arousal or distress. These findings appear to be more characteristic of females than of males, which is interesting given females’ greater tendencies toward internalizing psychopathology in adolescence and young adulthood, a relation that may be at least partially mediated by females’ increased levels of neuroticism (Goodwin & Gotlib, 2004; Hayward & Sanborn, 2002; Margalit & Eysenck, 1991). However, interpretation of these findings in the current sample was complicated by the fact that younger children tended to exhibit more catastrophic responses to failure than did older children, making it difficult to ascertain whether this tendency reflected a sampling bias or whether the phenomenon is somehow restricted only to younger children. Younger children also exhibited less-well-developed cognitive skills in the domains of memory, attention, and executive function (Luciana & Nelson, 1998, 2002).

To partially alleviate this confound, we restricted our sample to the youngest age cohorts (ages 4 to 6 years), dichotomized our response-to-failure variable, and re-examined the associations of interest. Within this restricted sample, we found that children who respond to failure with subsequent failures more than 50% of the time were those rated by their parents one year later as higher in Sadness as well as, at a trend level, Negative Affectivity more broadly construed. These children might also exhibit higher morning cortisol levels under relatively low-stress conditions (weekend versus school-day sampling), although this latter conclusion is tentative given that this finding eluded statistical significance even at a marginal level and was based on a small sample.

The processes through which individuals cope with perceived errors or failures might serve, then, as important indicators of their levels of depression-related personality tendencies and markers of their vulnerability to subsequent clinical disorders. Moreover, an increased understanding of the neural mechanisms of error processing during development might provide us with important clues regarding the pathophysiology of internalizing psychopathology.

Feedback processing has been linked in recent studies to a network of structures including the anterior cingulate cortex, the amygdala, and the ventromedial prefrontal cortex (Bechara, Damasio & Damasio, 2000; Gehring & Willoughby, 2002). Although interactions among these regions have not been studied extensively in children, a number of adult studies indicate that these regions are critical for motivated decision-making. Specifically, the amygdala appears to mark or tag situations according to their motivational salience. It would presumably define the affective tone that should characterize a given stimulus context in conjunction with other limbic and ventral striatal structures that modulate reinforcement learning (Depue & Collins, 1999). In the context of neurocognitive testing, the affective “set” that should guide performance is based on an individual’s desire to perform well and, if feedback is provided, to achieve positive reinforcement as frequently as possible. This hedonic goal would be relayed to the ventromedial prefrontal cortex, which may, in concert with the
anterior cingulate cortex, assume the task of monitoring the outcomes of actions that are incentive-guided. If outcomes are positive (as with successful performance on a given trial), then behavior can continue according to the contingencies that have maintained performance up to that point. However, if the outcome on a given trial is negative, this processing network is put on a “state of alert”. The anterior cingulate cortex should signal the ventromedial prefrontal cortex that an error has been made and that a change in response is necessary. Flexible control over responding is presumably achieved among a number of prefrontal cortical regions including the ventromedial prefrontal cortex but also the dorsolateral prefrontal cortex, which orchestrates larger-scale shifts in behavior (Robbins, 1996), and specifically of the type demanded by the more difficult stages of the ID/ED set-shifting task that was used here (Dias, Robbins, & Roberts, 1996).

Responding to perceived failure with subsequent failure suggests that this self-monitoring system functions in a relatively intact way up until the point when a response shift is required. Dysfunction at this point in the processing stream could be due to dysregulation in the anterior cingulate cortex, the ventromedial prefrontal cortex, or the dorsolateral prefrontal cortex. Notably, dysfunction in all three regions has been discussed in relation to depression (see reviews by Drevets, 2000; Rajkowska, 2000), and anterior cingulate dysfunction has been specifically associated with obsessive anxiety (see review by Allman, Hakeem, & Watson, 2002). The amygdala might also be involved if, in the course of error-related feedback, the motivational valence attached to the stimulus context is altered from one that is positively incentive-guided to one that is viewed as threatening. These regions are also intimately connected with other limbic structures, including the hippocampus and hypothalamus, and are thus positioned to impact (and be impacted by) memory-based cognition but also stress reactivity as achieved through endocrine responses of the hypothalamic-pituitary-adrenal axis (Swaab, Fliers, Hoogendijk, Veltman, & Zhou, 2000).

How these structures interact to mediate feedback processing in healthy developmental samples has only recently been considered, and findings suggest that similar structures as well as physiological mechanisms regulate error-monitoring and cognitive control functions in children as well as adults (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Davies, Segalowitz, & Gavin, 2004) albeit at decreased levels of efficiency. These methods have not been applied, however, to children with internalizing disorders. Future studies that incorporate neuroimaging of higher-risk samples in the context of error monitoring might be more fruitful in eliciting interactions among these behavioral phenomena and establishing their neural correlates.

The analysis presented here could be viewed as a first step in that direction, although we must acknowledge several limitations of our approach. First, we may have been less likely to observe affect-cognition associations in the current sample, because they were free of psychopathology and presumably low in risk. That said, it must also be acknowledged that we did not comprehensively assess these children’s socioeconomic situations or parental histories of
psychopathology. In this regard, prospective studies that include high-risk samples and larger numbers of children in each age cohort are clearly needed. It has been shown, for instance, that the development of anxiety and depression symptoms in children, as exemplified by a recent longitudinal study of 4th to 11th grade children, is predicted by their levels of positive and negative affect (Lonigan, Phillips, & Hooe, 2003). Second, it could be argued that parental ratings are insufficient to capture internalizing tendencies in children (Faraone, Biederman, & Milberger, 1995). Given that the sample studied here was expected to be low overall in internalizing tendencies, it is unclear how the use of parental ratings impacted our findings. Third, to examine the tendency to respond to failure with subsequent failure in a manner that will allow maximal sensitivity to the phenomenon but provide information regarding specificity of the effect, tasks are needed that present challenges to the test-taker so that sufficient numbers of errors will be generated but where difficulty level is evenly interspersed across trials. Ideally, several tasks would be used, each of which addresses a different cognitive function that is mediated by a different brain region. This design would permit inferences to be made about the nature of feedback processing as a general motivational influence over performance that is not linked to a specific task, especially if that task activates some of the same prefrontal regions that might contribute to feedback processing in a more general sense. For instance, the ID/ED set-shifting task has very discrete brain activity correlates (in this case, the ventromedial and dorsolateral prefrontal cortices: Dias et al., 1996). Fourth, our assessment of cortisol activity might have yielded stronger findings had we measured those levels while children were actually performing a task such as the ID/ED set-shifting task and while they were in the midst of coping with task failure. Finally, targeted investigations of this type should include larger sample sizes that permit multivariate models of performance to be examined.

Despite these caveats, it is encouraging that we were able to obtain even subtle effects in our sample in the context of a series of analyses that were admittedly post hoc. Better designed approaches to investigating interactions among temperament, error monitoring, and cognitive performance have the strong potential to inform developmental cognitive neuroscientific investigations of motivated decision-making, how deficiencies in this process might be associated with the cognitive biases that have been used for decades to describe individuals with anxiety and/or depression, and how such deficiencies might render individuals vulnerable to internalizing psychopathology.

Acknowledgments

The CANTAB data collection phase of this research was supported by the MacArthur Research Network on Psychopathology and Development, David Kupfer M.D. (Chair). The collection of temperament and cortisol data was supported by a grant from the National Institute of Health (MH56958) and a National Institute of Mental Health Research Science Award (MH00946) to Megan R. Gunnar.
REFERENCES


*Iunie 2005 ● Cogniție, Creier, Comportament* 359


