

## Investigations of Temperament at Three to Seven Years: The Children's Behavior Questionnaire

Mary K. Rothbart, Stephan A. Ahadi, Karen L. Hershey, and Phillip Fisher

This article reviews evidence on the reliability and validity of the Children's Behavior Questionnaire (CBQ), and presents CBQ data on the structure of temperament in childhood. The CBQ is a caregiver report measure designed to provide a detailed assessment of temperament in children 3 to 7 years of age. Individual differences are assessed on 15 primary temperament characteristics: Positive Anticipation, Smiling/Laughter, High Intensity Pleasure, Activity Level, Impulsivity, Shyness, Discomfort, Fear, Anger/Frustration, Sadness, Soothability, Inhibitory Control, Attentional Focusing, Low Intensity Pleasure, and Perceptual Sensitivity. Factor analyses of CBQ scales reliably recover a three-factor solution indicating three broad dimensions of temperament: Extraversion/Surgency, Negative Affectivity, and Effortful Control. This three-factor solution also appears to be reliably recovered in ratings of children in other cultures (e.g., China and Japan). Evidence for convergent validity derives from confirmation of hypothesized relations between temperament and socialization-relevant traits. In addition, parental agreement on CBQ ratings is substantial. The CBQ scales demonstrate adequate internal consistency, and may be used in studies requiring a highly differentiated yet integrated measure of temperament for children in this age range.

### INTRODUCTION

The Children's Behavior Questionnaire (CBQ), a caregiver report measure, has been developed with the goal of providing a highly differentiated assessment of temperament in young children. In this report, we describe the development of the CBQ, and consider implications of work with this instrument for our understanding of the structure of temperament during the early childhood years. We also present evidence for the reliability and validity of the CBQ.

Most questionnaires previously developed to assess temperament in preschool and early school-age children have been based on dimensions identified by the New York Longitudinal Study (NYLS; Thomas & Chess, 1977), dimensions identified by Buss and Plomin (Emotionality–Activity–Sociability; 1975, 1984), or both combined (Rowe & Plomin, 1977).

Dimensions identified in the NYLS include Activity Level, Mood, Approach/Withdrawal, Intensity, Threshold, Rhythmicity, Distractibility, Attention Span/Persistence, and Adaptability. The nine dimensions assessed in the NYLS were originally extracted from a content analysis of interviews with mothers of 2- to 6-month-old infants (Thomas, Chess, & Birch, 1968; Thomas, Chess, Birch, Herzig, & Korn, 1963). Temperament dimensions in the NYLS system are thus based on infant data, and would therefore not include temperamental characteristics that develop later than early infancy. Because the NYLS dimensions were developed for clinical purposes, there was also no initial attempt to make the scales conceptually

independent. Thus, research using instruments based on the Thomas and Chess model has frequently identified fewer dimensions of temperamental variability than were originally posited (see reviews by Martin, Wisenbaker, & Huttunen, 1994; Rothbart & Mauro, 1990; Sanson, Smart, Prior, Oberklaid, & Pedlow, 1994).

In a second approach to the assessment of temperament, Buss and Plomin (1975, 1984) invoked criteria of early appearance and heritability to generate a list of temperament dimensions including Emotionality, Activity, and Sociability (EAS). When a selected set of items from both the NYLS and EAS scales were combined and administered to a sample of children aged 1 to 6 years, item-level factor analysis yielded dimensions of Emotionality, Soothability, Activity Level, Attention Span, and Sociability (Rowe & Plomin, 1977).

The CBQ follows a third approach, in which items are rationally generated to assess theoretically derived temperament dimensions. This approach, followed in our previous work (Capaldi & Rothbart, 1992; Derryberry & Rothbart, 1988), and based on the approach of Fiske (1966, 1971), uses contemporary theory to identify central constructs of temperament, such as emotional reactivity, arousability, and self-regulation. These central constructs are then further decomposed into subconstructs, and items generated to reflect them. This rational approach allows for precise definitions and explorations of the structure of

each domain at varying levels of breadth, and investigating patterns of intercorrelations among subconstructs to empirically characterize the more general constructs. Unlike often heterogeneous, factor-derived scales, the larger constructs are thus composed of relatively homogeneous components. This may be seen as a bottom-up approach to scale development and to an understanding of individual differences constructs.

Because the CBQ is a theory-derived instrument, it is appropriate to consider briefly some theoretical concerns guiding its development. The first two approaches described above assumed that individual differences in temperament reflected only characteristic styles of behavioral response (Buss & Plomin, 1975; Thomas & Chess, 1977). More recent temperament theorists, in line with Allport's (1937) position, view temperament as including the individual's characteristic affective qualities, such as the prevailing quality and intensity of emotional reactivity (Derryberry & Rothbart, 1997; Rothbart & Bates, 1998). This focus on the affective nature of temperament is related to research on the emergence of the primary emotions (Izard, 1977), and in adult subjects has led to the identification and assessment of independent dimensions of positive and negative emotionality (Diener & Emmons, 1984; Watson, Clark, & Tellegen, 1988).

However, the assumption that temperamental processes are synonymous with, or restricted to, affective processes has been challenged by the research and theory of Rothbart and colleagues (Rothbart, 1981, 1989; Rothbart, Ahadi, & Evans, 2000; Rothbart & Derryberry, 1981; Rothbart & Posner, 1985). Research on the development of attention systems, especially those involved in executive or effortful control, suggests the importance of individual differences in attentional self-regulation as a basic dimension of temperament (Posner & Rothbart, 1998; Rothbart, Ahadi, et al., 2000; Ruff & Rothbart, 1996), and these characteristics are also assessed by the CBQ.

Temperamental functioning is also seen as influenced by maturation and experience. For example, behavioral fear does not appear until about 6 to 7 months of age, the beginnings of executive attentional self-regulation are not seen until about 10 to 12 months of age, and the executive system appears to undergo rapid development during the toddler and preschool years (Posner & Rothbart, 1998; Rothbart, 1989). The executive or effortful control system, and the frontal lobe structures related to it, continues to develop into the second decade of life (Chugani, Phelps, & Mazziotta, 1987; Huttenlocher, 1990). By early childhood, we would expect temperamental

characteristics observed in both infancy and adulthood to be present. The CBQ was therefore designed to assess constructs of temperament taken upward in age from the Infant Behavior Questionnaire (IBQ; Rothbart, 1981) and downward in age from the Physiological Reactions Questionnaire (PRQ) developed for assessment of adult temperament (Derryberry & Rothbart, 1988).

The CBQ was developed to assess temperament, defined as constitutionally based, individual differences in reactivity and self-regulation. By constitution is meant the individual's relatively enduring biological make up, influenced over time by heredity, maturation, and experience (Rothbart, 1989; Rothbart & Derryberry, 1981). Reactivity refers to the arousability of motor, affective, and sensory response systems. Self-regulation refers to processes that serve to modulate (increase or decrease) reactivity, including attentional focusing and inhibitory control. This broad framework (Rothbart, 1989; Rothbart & Derryberry, 1981) allows for the investigation of dimensions of temperamental variability that are not studied in other systems, and the differentiated assessment of individual characteristics of positive emotional reactivity (smiling/laughter, pleasure) and negative emotional reactivity (fear, distress to limitations, sadness), as well as self-regulatory mechanisms of attention, activity, behavioral inhibition (fear or shyness), and inhibitory control.

## METHOD

Items written for the CBQ were based on conceptual definitions for each scale (see Appendix A), and derived from several sources. First, as described in the Introduction, items from the PRQ (Derryberry & Rothbart, 1988) and the IBQ (Rothbart, 1981) were adapted to the preschool-early elementary school-age group. Parents were also interviewed about their children, and asked about characteristics on which siblings differed; specific information was elicited about situations in which those characteristics were demonstrated. Finally, 15 parents (12 mothers, 3 fathers) of children in the appropriate age range completed the questionnaire aloud, commenting on any problems with items and suggesting possibilities for revision and addition of items. The aim was to design questions that could be used for children 3 to 8 years of age, although the initial scale development was limited to the 3- to 7-year age range. Sample items are included in Appendix A. Parents were asked to rate their child on each item, using a 7-point Likert scale ranging from "extremely untrue of your child" to "extremely true of your child." Parents were also provided

with a "not applicable" response option to be used when the child had not been observed in the situation described.

Items as written were then administered to parents of 262 children from three samples of children 3 to 7 years of age, and item statistics were computed. Items showing low item-total correlations ( $<.20$ ) with scales across these samples were eliminated, as well as items that had higher correlations with another scale than with the intended scale. This resulted in a shorter form of the CBQ that included 195 items. A long form, which has also been used in our laboratory, contains additional social-emotional scales including Aggression, Empathy, Guilt/Shame, Help-Seeking, and Negativity (Rothbart, Ahadi, & Hershey, 1994). The shorter form, with typically 12 to 13 items per scale, over time has become the standard version of the CBQ.

The dimensions of temperament chosen for assessment in the CBQ are defined in Appendix A, with sample items included. *Activity Level*, widely assessed by temperament researchers, is presently included in NYLS-derived questionnaires (e.g., McDevitt & Carey, 1981), the EAS, and the IBQ. *Anger/Frustration* is assessed as part of the Emotionality construct in the EAS, and as part of the Mood construct in the NYLS. The same construct is assessed as Distress to Limitations in the IBQ. *Anger/Frustration* has been identified as a primary emotion (Izard, 1977) and is related to both strength of expectation of reward (and thus approach-anticipation) and to aggressive self-regulation (Rothbart & Bates, 1998).

*Attentional Focusing* corresponds to the NYLS dimension of Attention Span/Persistence, and, to a degree, with the IBQ dimension of Duration of Orienting. In adults, related but distinct scales of Attentional Focusing and Attentional Shifting were found to be highly intercorrelated (Derryberry & Rothbart, 1988). For this reason, in the early version of the CBQ, Attentional Focusing and Shifting were combined into a single scale; however, the two variables did not hold together in item analysis. There were enough items for an Attentional Focusing Scale, but not for an Attentional Shifting Scale; thus, only the former scale is included in the present analyses. (In our laboratory, we are currently working on developing an Attentional Shifting Scale.)

The *Discomfort* aspect of negative emotion corresponds to the primary emotion of distress (Izard, 1977). *Falling Reactivity/Soothability* is not assessed directly in infant scales, with the exception of its affective component in the IBQ scale of Soothability. *Fear* is subsumed under the Emotionality dimension of the EAS, and corresponds to the IBQ Fear dimension and the Withdrawal pole of the NYLS Approach/

Withdrawal dimension. Fear has also been identified as a primary emotion (Izard, 1977).

*High Intensity Pleasure* is assessed in the PRQ (Derryberry & Rothbart, 1988), and it appears to correspond to the construct of sensation seeking studied by Zuckerman (1979, 1990). *Impulsivity* is defined as speed of response initiation. Although Impulsivity was assessed in the EASI (Emotionality-Activity-Sociability-Impulsivity; the precursor of the EAS), Buss and Plomin (1984) dropped this dimension from the EAS due to a lack of strong evidence for its heritability. Impulsivity or behavioral activation, however, measured as latency to contact small objects, has been found to be an important aspect of children's approach in our laboratory (Rothbart, 1988) and is included in other theoretical models (see review by Rothbart, Derryberry, & Posner, 1994). It is therefore included in the CBQ. Because *Inhibitory Control* develops after early infancy (Rothbart, 1989), it is not assessed in either the IBQ or the NYLS system; however, it is assessed in the PRQ (Derryberry & Rothbart, 1988) and is appropriate to this age range (3-7 years) in children. *Low Intensity Pleasure* is also assessed in the PRQ (Derryberry & Rothbart, 1988).

*Perceptual Sensitivity* corresponds to the NYLS dimension of Threshold and the PRQ scale of External sensitivity, and is related to attentional systems of orienting (Posner & Raichle, 1994). *Positive Anticipation* is assessed under the Approach pole in the Approach/Withdrawal dimension of the NYLS system, although chiefly with respect to novel situations. Positive Anticipation is not assessed in either the EAS or IBQ, but is assessed in the adult PRQ. This dimension corresponds to one of the major dimensions in Gray's (1982, 1987), Panksepp's (1982, 1998), and Depue and associates' (Depue & Collens, 1999; Depue & Iacono, 1989) biological models of temperament (see review by Rothbart, Derryberry, & Posner, 1994).

*Sadness* is not assessed in the IBQ or other infant and child measures, but is assessed in the PRQ and has been identified as a primary emotion (Izard, 1977). *Shyness* corresponds to one facet of Kagan and colleagues' behavioral inhibition construct (Kagan, Reznick, & Snidman, 1988; Rothbart & Mauro, 1990). *Smiling/Laughter* is expected to be related to the primary emotion of joy (Izard, 1977). It is not assessed in the EAS, but would constitute the positive pole of the NYLS Mood dimension. It is assessed in the IBQ (Rothbart, 1981).

This article reports on a number of initial steps basic to the development of an individual differences measure. Results of tests for internal consistency and temporal stability were obtained, the factor structures of the instrument, both for U.S.

samples and for samples from the People's Republic of China and Japan were investigated, and evidence for convergent validity was derived from both reports of parent agreement and prediction of social behavior patterns.

### Samples

Both published and original results of tests for reliability and validity of the CBQ are included in the present review. Original results are based on CBQ protocols obtained from a number of sources. Protocols from 497 children were obtained from studies of temperament conducted by Goldsmith and Rothbart at the University of Oregon (Goldsmith, Buss, & Lemery, 1997). These protocols were divided into three groups according to children's age. The first group consisted of 99 three-year-olds ( $M = 41.8$  months), including 60 boys and 38 girls (1 child was unidentified by gender). The second group consisted of 240 four- and five-year-olds ( $M = 57.6$  months), including 135 boys and 105 girls. The third group consisted of 158 six- and seven-year-olds ( $M = 82.7$  months), including 94 boys and 60 girls (4 children were unidentified by gender). These protocols were included in the factor analyses of the CBQ scales. In addition, CBQ protocols for 171 children were obtained from Kochanska (Kochanska, DeVet, Goldman, Murray, & Putnam, 1994) at the University of Iowa. These children were, on average, much younger ( $M = 39.95$  months), and included 79 boys and 92 girls. Only protocols from children who were at least 3 years old were included in the factor analysis of these scales.

CBQ protocols were also obtained from the parents of several samples of children who participated in studies conducted by Fagot and Fisher at the Oregon Social Learning Center (OSLC; Fagot & Leve, 1998; Fisher, 1993). Three samples were obtained when the children were between 5 and 6 years of age (Sample 1:  $N = 77$ ,  $M = 70.38$  months; Sample 2:  $N = 81$ ,  $M = 67.16$  months; and Sample 3:  $N = 135$ ,  $M = 65.87$  months). Of the 293 children in these three samples, 154 were boys and 139 were girls. Participants in the first two samples were followed up when the children were about 7 years of age. At the second assessment, Sample 1  $N = 59$ ,  $M = 90.1$  months, and Sample 2  $N = 64$ ,  $M = 85.42$  months of age. Protocols from OSLC participants were also included in the factor analyses of the CBQ scales. The protocols from these participants allowed us to examine the stability of CBQ scale scores over time. In addition, because, in many cases, protocols were obtained from both parents, we were also able to examine parental agreement on ratings of children's temperament.

## RESULTS

### CBQ Internal Consistency Reliability

Internal consistency estimates of the CBQ scales have been reported in a number of sources. Ahadi, Rothbart, and Ye (1993) presented internal consistency coefficients for CBQ scales administered to a sample of 262 parents in the northwest region of the United States. Coefficient  $\alpha$ s for the CBQ scales ranged from .67 to .94, with a mean internal consistency estimate of .77 across all 15 scales. Kochanska and colleagues (Kochanska et al., 1994) obtained similar reliability estimates for the CBQ scales in a study of temperament that included parent reports on 171 children. They obtained internal consistency estimates ranging from .68 to .93, with a mean reliability estimate of .78 across the 15 scales.

Table 1 presents internal consistency estimates based on coefficient  $\alpha$ s for the CBQ scales obtained from all available mother ratings in the OSLC samples, regardless of test administration period. This yielded a total of 411 CBQ protocols for children between the ages of 4 and 8 years of age (9 children were over 95 months of age). Internal consistency estimates for the CBQ scale ratings of 4- and 5-year-olds ranged from .64 to .92, with a mean of .73. Reliability estimates for scale ratings of 6- and 7-year-olds ranged from .67 to .92, with a mean of .75. Table 1 also presents average item-total correlations for the CBQ scales, correcting for the contribution of the items to the scale. The average item-total correlations suggest adequate consistency of item content within the CBQ scales.

### Structure of CBQ Scales

To examine the factor structure of the CBQ scales, parent ratings of children aged 4 to 7 years were obtained from the studies on child temperament conducted in the northwestern and midwestern regions of the United States. Mean item response scores on the 15 CBQ scales were obtained for 341 six- and seven-year-olds and 516 four- and five-year-olds. The matrix of intercorrelations among the CBQ scale scores was submitted to a principal axis factor analysis, iterating to communalities, and the extracted factors were rotated obliquely via the Oblimin algorithm (Norusis, 1994).

Table 2 presents the factor pattern matrix for the sample of 6- and 7-year-olds. The principal axis factor analysis of the CBQ scale scores of the 6- and 7-year-olds identified three factors with eigenvalues greater than 1. The first factor was defined primarily by loadings for the scales of Discomfort, Sadness, Fear, Anger/Frustration, and, loading negatively, Soothability. This pattern of loadings was consistent with the

**Table 1** Internal Consistency and Item-Total Correlations of the Children's Behavior Questionnaire Scales

Scale	No. Items	Internal Consistency Estimates (coefficient $\alpha$ )		Average Corrected Item-Total Correlations	
		4- and 5-Year-Olds <sup>a</sup>	6- and 7-Year-Olds <sup>b</sup>	4- and 5-Year-Olds <sup>a</sup>	6- and 7-Year-Olds <sup>b</sup>
Positive Anticipation	13	.74	.77	.38	.40
High Intensity Pleasure	13	.79	.77	.43	.40
Smiling/Laughter	13	.75	.80	.40	.46
Activity Level	13	.75	.75	.38	.38
Impulsivity	13	.74	.78	.36	.40
Shyness	13	.92	.92	.66	.65
Discomfort	12	.73	.67	.38	.31
Fear	12	.70	.70	.34	.34
Anger/Frustration	13	.80	.81	.43	.45
Sadness	12	.69	.71	.32	.34
Soothability	12	.66	.67	.32	.34
Inhibitory Control	13	.76	.78	.39	.40
Attentional Focusing	14	.67	.69	.29	.31
Low Intensity Pleasure	13	.64	.73	.29	.37
Perceptual Sensitivity	12	.64	.71	.31	.37

<sup>a</sup>  $N = 228$ .<sup>b</sup>  $N = 183$ .

superfactor of Negative Affectivity. The second factor was defined primarily by loadings for the scales of Impulsivity, High Intensity Pleasure, Activity Level, and, loading negatively, Shyness. Positive Anticipation and Smiling/Laughter also loaded substantially

**Table 2** Factor Pattern of Children's Behavior Questionnaire Scales in a U.S. Sample of 6- and 7-Year-Olds

Scale	Factor		
	1 Negative Affectivity	2 Extraversion/ Surgency	3 Effortful Control
Positive Anticipation	.48 <sup>a</sup>	.46 <sup>a</sup>	.25 <sup>a</sup>
High Intensity Pleasure	-.12	.69 <sup>a</sup>	.09
Smiling/Laughter	-.01	.35 <sup>a</sup>	.72 <sup>a</sup>
Activity Level	.15	.61 <sup>a</sup>	-.13
Impulsivity	-.02	.87 <sup>a</sup>	-.07
Shyness	.31 <sup>a</sup>	-.60 <sup>a</sup>	-.11
Discomfort	.64 <sup>a</sup>	-.06	.06
Fear	.53 <sup>a</sup>	-.24	.01
Anger/Frustration	.64 <sup>a</sup>	.29 <sup>a</sup>	-.07
Sadness	.68 <sup>a</sup>	-.08	.10
Soothability	-.52 <sup>a</sup>	-.07	.28 <sup>a</sup>
Inhibitory Control	-.23	-.41 <sup>a</sup>	.60 <sup>a</sup>
Attentional Focusing	-.30 <sup>a</sup>	-.34 <sup>a</sup>	.41 <sup>a</sup>
Low Intensity Pleasure	-.08	-.11	.77 <sup>a</sup>
Perceptual Sensitivity	.11	.07	.51 <sup>a</sup>

Note:  $N = 341$ , from the University of Oregon ( $n = 158$ ) and Oregon Social Learning Center ( $n = 183$ ) datasets.

<sup>a</sup> Loadings greater than or equal to .25.

on this factor. The constructs defining this factor were consistent with the superfactor dimension of Extraversion identified in structural examinations of adult temperament and personality (Eysenck & Eysenck, 1985; Tellegen, 1985). The third factor was defined by loadings for the scales of Low Intensity Pleasure, Smiling/Laughter, Inhibitory Control, Perceptual Sensitivity, and Attentional Control. This factor appeared to correspond to the superfactor variously labeled Effortful Control (Rothbart & Bates, 1998) or Constraint (Tellegen, 1985). For normative purposes, the means, *SDs*, and ranges for the mean item responses on the CBQ scales for these three samples are provided in Appendix B.

The factor pattern matrix for the 4- and 5-year-olds is presented in Table 3. The pattern of loadings for the first factor conformed closely to the first extracted factor in the 6- and 7-year-old sample, and suggested underlying Negative Affectivity. The pattern of loadings on the second factor was quite similar to the second factor extracted for the 6- and 7-year-olds, and conformed to the broad dimension of Extraversion/Surgency. The third factor's pattern of loadings was quite similar to that of the third extracted factor identified in the analysis of the older sample, and was labeled Effortful Control.

We were also interested in determining whether the three-factor structure of temperament would emerge for a sample of 3-year-olds, and thus obtained CBQ protocols from a sample of 149 three-year-olds

**Table 3** Factor Pattern of Children's Behavior Questionnaire Scales in a U.S. Sample of 4- and 5-Year-Olds

Scale	Factor		
	1 Negative Affectivity	2 Extraversion/ Surgency	3 Effortful Control
Positive Anticipation	.50 <sup>a</sup>	.49 <sup>a</sup>	.29 <sup>a</sup>
High Intensity Pleasure	-.10	.73 <sup>a</sup>	.12
Smiling/Laughter	.02	.52 <sup>a</sup>	.66 <sup>a</sup>
Activity Level	.10	.72 <sup>a</sup>	-.20
Impulsivity	-.06	.83 <sup>a</sup>	-.17 <sup>a</sup>
Shyness	.29 <sup>a</sup>	-.46 <sup>a</sup>	-.04
Discomfort	.60 <sup>a</sup>	-.20	.02
Fear	.55 <sup>a</sup>	-.20	.05
Anger/Frustration	.70 <sup>a</sup>	.21	-.04
Sadness	.69 <sup>a</sup>	-.02	.07
Soothability	-.59 <sup>a</sup>	-.01	.30 <sup>a</sup>
Inhibitory Control	-.23	-.26 <sup>a</sup>	.70 <sup>a</sup>
Attentional Focusing	-.31 <sup>a</sup>	-.17	.58 <sup>a</sup>
Low Intensity Pleasure	.08	.14	.80 <sup>a</sup>
Perceptual Sensitivity	.20	-.05	.42 <sup>a</sup>

Note:  $N = 517$ , from the University of Oregon ( $n = 240$ ), Oregon Social Learning Center ( $n = 228$ ), and the University of Iowa ( $n = 49$ ) datasets.

<sup>a</sup>Loadings greater than or equal to .25.

(65 girls) from the midwestern and northwestern samples for analysis. The analysis identified four factors with eigenvalues greater than 1. The four-factor solution failed to converge, however, because a communality estimate exceeded 1 during iteration. We then forced a three-factor solution, which is presented in Table 4. The three-factor structure for this sample was substantially similar to that obtained with the older children, with just a few differences.

The first factor was defined primarily by loadings on the scales of Impulsivity, High Intensity Pleasure, Activity Level, and, loading negatively, Shyness. This factor was identifiable as the global dimension of Extraversion/Surgency. The Positive Anticipation scale loaded moderately on all three factors, consistent with the older samples. The second factor appeared to correspond to the global dimension of Effortful Control, with primary loadings of Low Intensity Pleasure, Inhibitory Control, Attentional Focusing, and Perceptual Sensitivity, although the loading for Perceptual Sensitivity appeared attenuated in this sample. As with the older children, Smiling/Laughter loaded more strongly on this Effortful Control factor. The third factor was defined primarily by loadings for the scales reflecting negative emotions, and was recognizable as the global dimension of Negative Affectivity. The loading for Fear appeared attenuated in this sample, however, and Shyness did not load on

this factor as it did in the older samples. In addition, although Soothability loaded moderately on the Effortful Control factor in the older samples, this relation did not appear in the 3-year-old samples. Thus, there appeared to be substantial congruence in the structure of temperament across ages 3 to 7, although some differences suggest the presence of developmental shifts. It is important to note, however, that the 3-year-old sample was considerably smaller than the samples of older children, and thus differences in structure may not be reliable.

In deciding whether to rotate the extracted factors orthogonally versus obliquely, we were more interested in identifying possible differences in structural relations among the factors across populations than in ensuring the comparability of the derived structure across populations. For example, we previously reported that in a sample of Chinese respondents, the Effortful Control factor correlated negatively with the Extraversion factor, but was uncorrelated with the Negative Affectivity factor (Ahadi et al., 1993). This pattern of structural relations was potentially very interesting because in other studies of U.S. adults and children, Effortful Control was negatively correlated with Negative Affectivity, but uncorrelated with Extraversion. It may be the case, for example, that effortful self-regulation was employed to inhibit culturally discouraged behavioral tendencies—negative affect in the United States, and Extraversion in China.

**Table 4** Factor Pattern of Children's Behavior Questionnaire Scales in a U.S. Sample of 3-Year-Olds

Scale	Factor		
	1 Negative Affectivity	2 Extraversion/ Surgency	3 Effortful Control
Positive Anticipation	.33 <sup>a</sup>	.52 <sup>a</sup>	.44 <sup>a</sup>
High Intensity Pleasure	-.17	.76 <sup>a</sup>	.05
Smiling/Laughter	.00	.42 <sup>a</sup>	.74 <sup>a</sup>
Activity Level	.13	.77 <sup>a</sup>	-.15
Impulsivity	-.02	.89 <sup>a</sup>	-.04
Shyness	.14	-.45 <sup>a</sup>	-.09
Discomfort	.62 <sup>a</sup>	-.26 <sup>a</sup>	.03
Fear	.36 <sup>a</sup>	-.06	-.08
Anger/Frustration	.67 <sup>a</sup>	.20	.20
Sadness	.67 <sup>a</sup>	.00	.17
Soothability	-.55 <sup>a</sup>	-.02	.17
Inhibitory Control	-.25 <sup>a</sup>	-.31 <sup>a</sup>	.49 <sup>a</sup>
Attentional Focusing	-.25 <sup>a</sup>	-.30 <sup>a</sup>	.39 <sup>a</sup>
Low Intensity Pleasure	-.03	.01	.82 <sup>a</sup>
Perceptual Sensitivity	.16	.03	.34 <sup>a</sup>

Note:  $N = 149$ , from the University of Oregon ( $n = 99$ ) and the University of Iowa ( $n = 50$ ) datasets.

<sup>a</sup>Loadings greater than or equal to .25.

In the current samples, the correlations among the three factors were uniformly small. In the sample of 6- and 7-year-olds, the Negative Affectivity factor correlated positively with the Extraversion/Surgency factor,  $r(339) = .11$ , and negatively with the Effortful Control factor,  $r(339) = -.13$ . The Extraversion/Surgency and Effortful Control correlation was near 0,  $r(339) = -.04$ . In the 4- and 5-year-old sample, all three correlations were near 0: Negative Affectivity correlated with the Extraversion/Surgency and Effortful Control factors,  $r(515) = -.01$  and  $-.04$ , respectively, and there was a 0 correlation between the Effortful Control and Extraversion factors. With respect to the sample of 3-year-olds, the Negative Affectivity factor correlated positively with the Extraversion/Surgency factor,  $r(147) = .13$ , and negatively with the Effortful Control factor,  $r(147) = -.07$ . The Extraversion/Surgency and Effortful Control factors were also positively correlated,  $r(147) = .08$ . Because the correlations among the factors were near 0, it is unlikely that comparability of factor structures across age groups would be meaningfully compromised.

#### Confirmatory Factor Analysis

To further investigate the structure of the CBQ scales, a maximum likelihood confirmatory factor analysis (CFA) was performed for both the 4- and 5-year-old and the 6- and 7-year-old samples used in the exploratory factor analysis (EFA). For each sam-

ple, the structure identified in the EFA was used as the initial hypothesized structure in a model-building process. In the initial model, CBQ scales were allowed to load on factors in which their EFA-derived loading was .25 or greater, and scales with EFA-derived loadings less than .25 on a factor were set to 0 in the initial CFA model. The initial CFA models did not, however, demonstrate adequate fit. Consequently, modification indices from each run were used to refine the structural model, which was then tested in a subsequent run in an iterative fashion until a "best-fitting" model was identified. The best-fitting models for both the 4- and 5-year-old and the 6- and 7-year-old samples were three-factor solutions, and these are presented in Tables 5 and 6, respectively.

The adequacy of the final models was evaluated via several goodness-of-fit indices generated by the EQS package (Bentler, 1989). For the 4- and 5-year-old sample, the Bentler and Bonnet Normed Fit Index (NFI) was .96, the Comparative Fit Index (CFI) was .98, and the ratio of C2 (127) to degrees of freedom ( $df$ ; 61) was 2.08, all indicating acceptable fit. Goodness-of-fit indices for the 6- and 7-year-old sample included a Bentler and Bonnett NFI value of .94, a CFI value of .97, and a C2 (125) to  $df$  (66) ratio of 1.89, again all indicating acceptable fit.

Although the factor models derived from the EFA and CFA were highly similar, they were not identical. Tables 7 and 8 present the differences between factor loadings identified in the EFA and those identified in

**Table 5** Confirmatory Factor Analysis Model of Children's Behavior Questionnaire Scales in a U.S. Sample of 4- and 5-Year-Olds

Scale	Factor		
	1 Negative Affectivity	2 Extraversion/ Surgency	3 Effortful Control
Positive Anticipation	.46	.60	.35
High Intensity Pleasure	.00	.76	.20
Smiling/Laughter	.00	.74	.76
Activity Level	.11	.77	-.21
Impulsivity	.00	.79	-.16
Shyness	.26	-.31	.00
Discomfort	.53	-.16	.00
Fear	.47	-.14	.00
Anger/Frustration	.75	.18	.00
Sadness	.71	.00	.00
Soothability	-.64	.10	.34
Inhibitory Control	-.26	.00	.65
Attentional Focusing	-.33	.00	.52
Low Intensity Pleasure	.00	.44	.93
Perceptual Sensitivity	.17	.00	.38

Note:  $N = 517$ .

**Table 6** Confirmatory Factor Analysis Model of Children's Behavior Questionnaire Scales in a U.S. Sample of 6- and 7-Year-Olds

Scale	Factor		
	1 Negative Affectivity	2 Extraversion/ Surgency	3 Effortful Control
Positive Anticipation	.34	.52	.37
High Intensity Pleasure	-.27	.68	.15
Smiling/Laughter	-.15	.22	.67
Activity Level	.00	.72	.00
Impulsivity	.00	.77	.00
Shyness	.37	-.38	.00
Discomfort	.69	.00	.00
Fear	.59	.00	.00
Anger/Frustration	.56	.43	.00
Sadness	.57	.00	.00
Soothability	-.51	-.20	.29
Inhibitory Control	-.21	-.61	.56
Attentional Focusing	-.28	-.49	.37
Low Intensity Pleasure	.00	-.23	.63
Perceptual Sensitivity	.00	.00	.53

Note:  $N = 341$ .

**Table 7** Discrepancies between Confirmatory and Exploratory Factor Analysis Parameter Estimates in a Sample of 4- and 5-Year-Olds

Scale	Factor		
	1 Negative Affectivity	2 Extraversion/ Surgency	3 Effortful Control
Positive Anticipation	-.04	.11	.06
High Intensity Pleasure	.10	.03	.08
Smiling/Laughter	-.02	.22	.10
Activity Level	.01	.05	-.01
Impulsivity	.06	-.04	.01
Shyness	-.03	.15	.04
Discomfort	-.07	.04	-.02
Fear	-.08	.06	-.05
Anger/Frustration	.05	-.03	.04
Sadness	.02	.02	-.07
Soothability	-.05	.11	.04
Inhibitory Control	-.03	.26	-.05
Attentional Focusing	-.02	.17	-.06
Low Intensity Pleasure	-.08	.30	.13
Perceptual Sensitivity	-.03	.05	-.04

**Table 8** Discrepancies between Confirmatory and Exploratory Factor Analysis Parameter Estimates in a Sample of 6- and 7-Year-Olds

Scale	Factor		
	1 Negative Affectivity	2 Extraversion/ Surgency	3 Effortful Control
Positive Anticipation	-.14	.06	.12
High Intensity Pleasure	-.15	-.01	.06
Smiling/Laughter	-.14	-.13	-.05
Activity Level	-.15	.11	.13
Impulsivity	.02	-.10	.07
Shyness	.06	.22	.11
Discomfort	.05	.06	-.06
Fear	.06	.24	-.01
Anger/Frustration	-.08	.14	.07
Sadness	-.11	.08	-.10
Soothability	.01	-.13	.01
Inhibitory Control	.02	-.20	-.04
Attentional Focusing	.02	-.15	-.04
Low Intensity Pleasure	.08	-.12	-.14
Perceptual Sensitivity	-.11	-.07	.02

the CFA for the 4- and 5-year-old and 6- and 7-year-old samples, respectively. Inspection of Tables 7 and 8 indicates that differences between the two models were generally small. The average magnitude of difference between loadings estimated between the two analytic approaches was .07 for the 4- and 5-year-old sample and .09 for the 6- and 7-year-old sample. In addition, the largest discrepancies were confined to the Extraversion/Surgency factor in both samples, indicating, perhaps, that parameter estimates for this factor may be somewhat less reliable than those for the other factors.

#### Structure of CBQ Scales Cross-Culturally

Although cultural processes can be seen as having important consequences with respect to mean levels on trait scores, the basic structure of temperament characteristics should be relatively invariant across cultures. To examine this hypothesis, a Chinese translation of the CBQ was administered to mothers of 468 six- and seven-year-old children in Shanghai (Ahadi et al., 1993). The CBQ was translated by Ye and colleagues at the Shanghai Normal University and distributed to mothers of 6- and 7-year-old children attending a local kindergarten and elementary school. Internal consistencies for the CBQ scales in this sample ranged from .43 to .85, with a mean coefficient  $\alpha$  of .65. Although the internal consistency estimates were generally attenuated relative to the U.S. samples, two

scales (Attentional Focusing and Soothability) were particularly unreliable, with coefficient  $\alpha$ s of .43. Further work will be necessary to determine whether the poor reliability was due to problems with the translation of the CBQ or with the applicability of the item content in these scales across cultures.

A factor analysis of CBQ scale scores for this sample of Chinese 6- and 7-year-olds has also been reported (Ahadi et al., 1993). A principal axis factor analysis again identified a three-factor solution. The first factor was labeled Surgency (Extraversion) and was defined by loadings for the scales of Positive Anticipation, High Intensity Pleasure, Smiling/Laughter, Activity Level, Impulsivity, and, loading negatively, Shyness. The second factor was labeled Negative Affectivity and was defined by loadings for the scales of Discomfort, Fear, Anger/Frustration, Sadness, and, loading negatively, Soothability. The third factor was labeled Effortful Control and was defined by loadings for the scales of Inhibitory Control, Attentional Focusing, Low Intensity Pleasure, and Perceptual Sensitivity.

Although this structural model was highly consistent with the structure identified in the U.S. samples, there were some important differences. First, whereas Smiling/Laughter loaded primarily on Effortful Control and only moderately on the Extraversion/Surgency factor in the U.S. samples, it loaded only on the Surgency (Extraversion) factor in the Chinese sample. In addition, whereas the Positive Anticipation Scale loaded moderately on both the Extraversion/Surgency



and Negative Affectivity factors in the U.S. samples, it loaded only on the Surgency (Extraversion) factor in the Chinese sample.

Kusanagi (1993) has developed a Japanese translation of the CBQ. She administered the Japanese version of the CBQ to mothers of children attending one of three kindergartens near Sapporo City. Kusanagi presented coefficient  $\alpha$ s for the CBQ scales in the combined sample of 372 Japanese children ranging in age from 3 to 6 years (96% of the sample was 4 to 6 years of age). Coefficient  $\alpha$ s for the CBQ scales ranged from .54 to .93, with a mean internal consistency estimate of .75. Although the scales generally showed adequate reliabilities, it will again be necessary to determine whether translation problems may have attenuated scale reliabilities.

Kusanagi (1993) also reported a matrix of CBQ scale score intercorrelations for the Japanese sample. We submitted this matrix of intercorrelations to a principal axis factor analysis, iterating to communalities, and rotating the extracted factors obliquely. Examination of the three-factor solution revealed a pattern of factor loadings similar to that identified in the other samples. The factor pattern matrix for the three-factor solution is presented in Table 9. The first factor was defined primarily by loadings for the scales of Impulsivity, High Intensity Pleasure, Activity Level, and, loading negatively, Shyness. The scales of Smiling/Laughter, Positive Anticipation, and, loading

negatively, Inhibitory Control, also contributed to this factor. The second factor was defined by loadings for the scales of Anger/Frustration, Discomfort, Sadness, Fear, and, loading negatively, Soothability. Positive Anticipation also loaded strongly on this factor. Finally, the third factor was defined by loadings for the scales of Low Intensity Pleasure, Inhibitory Control, Smiling/Laughter, Perceptual Sensitivity, and Attentional Focusing.

The Japanese and U.S. structures both differed from the Chinese structure, with the primary placement of Positive Anticipation in the Negative Affectivity factor. With respect to the placement of Smiling/Laughter, the Japanese solution again looked more like the U.S. structure than the Chinese solution, with Smiling/Laughter loading primarily on the Effortful Control factor and secondarily on the Surgency (Extraversion) factor.

#### Convergent Validity

*Parental agreement.* To examine the degree of parental agreement on CBQ scales at different ages, correlations between parents' scores for the same child were examined across different samples of children and, for one sample, across age. Table 10 presents correlations between the ratings of mothers and fathers on the 15 CBQ scales. The first column reflects agreement scores from 49 couples who completed the CBQ in a study of temperament in 5-year-old children. Correlations between parent ratings ranged from .28

**Table 9** Factor Pattern of Children's Behavior Questionnaire Scales in a Japanese Sample of 3- to 6-Year-Olds

Scale	Factor		
	1 Negative Affectivity	2 Extraversion/ Surgency	3 Effortful Control
Positive Anticipation	.63 <sup>a</sup>	.37 <sup>a</sup>	.31 <sup>a</sup>
High Intensity Pleasure	-.01	.77 <sup>a</sup>	.04
Smiling/Laughter	.09	.48 <sup>a</sup>	.67 <sup>a</sup>
Activity Level	-.17	.71 <sup>a</sup>	-.05
Impulsivity	.14	.92 <sup>a</sup>	-.11
Shyness	.16	-.68 <sup>a</sup>	-.05
Discomfort	.68 <sup>a</sup>	-.12	-.07
Fear	.53 <sup>a</sup>	-.26 <sup>a</sup>	.15
Anger/Frustration	.78 <sup>a</sup>	.22	-.16
Sadness	.68 <sup>a</sup>	-.09	.16
Soothability	-.57 <sup>a</sup>	.19	.36 <sup>a</sup>
Inhibitory Control	-.26 <sup>a</sup>	-.45 <sup>a</sup>	.68 <sup>a</sup>
Attentional Focusing	-.30 <sup>a</sup>	-.14	.61 <sup>a</sup>
Low Intensity Pleasure	.12	-.03	.77 <sup>a</sup>
Perceptual Sensitivity	.20	.07	.66 <sup>a</sup>

Note:  $N = 419$ .

<sup>a</sup> Loadings greater than or equal to .25.

**Table 10** Correlations between Mother and Father Reports on the Children's Behavior Questionnaire Scales

	Agreement between Parent Ratings		
	$N = 49$	$N = 145$	$N = 114$
Positive Anticipation	.41*	.32*	.37*
High Intensity Pleasure	.62*	.57*	.35*
Smiling/Laughter	.34*	.23*	.17
Activity Level	.60*	.51*	.55*
Impulsivity	.72*	.57*	.53*
Shyness	.79*	.58*	.57*
Discomfort	.28	.46*	.30*
Fear	.44*	.43*	.34*
Anger/Frustration	.43*	.44*	.44*
Sadness	.48*	.40*	.30*
Soothability	.57*	.32*	.20*
Inhibitory Control	.72*	.40*	.47*
Attentional Focusing	.41*	.39*	.40*
Low Intensity Pleasure	.50*	.32*	.24*
Perceptual Sensitivity	.38*	.17*	.32*

Note: See text for further description of the three samples.

\*  $p < .05$ .

to .79, with a mean agreement across scales of .51. The second and third columns reflect agreement scores for couples participating in the OSLC studies. The second column reflects agreement at Time 1, when the children were 5 years of age, and the third column reflects agreement at Time 2, when the children were age 7. Average agreement across scales when the children were 5 years of age was .41, and average agreement when they were 7 years of age was .37. These correlations indicate substantial parental agreement that did not appear to increase across time.

In addition, agreement levels for CBQ scales tended to show consistency across samples. To illustrate, correlations between agreement scores for the CBQ scales were computed. The correlation between agreement scores at 5 years of age and those at 7 years of age was .69 in the OSLC sample. In addition, agreement scores between the independently assessed samples represented in Column 1 of Table 10 correlated with those obtained in the OSLC 5- and 7-year assessments, .61 and .64, respectively. Across the three samples, it appeared not only that levels of agreement varied substantially across scales, but also that the degree of agreement was consistent across samples. Thus, in the OSLC sample, for dimensions on which parents agreed at Time 1, they continued to agree at Time 2. Conversely, for dimensions on which parents showed less agreement at Time 1, they continued to show lower levels of agreement at Time 2.

It is important to note that the same pattern occurred in a third, independent sample. Hence, it appears that parents have consistently higher levels of agreement on some temperament dimensions both across independent samples and across time in the same samples, whereas they show consistently lower levels of agreement for other temperament dimensions. Of course, it is not clear why this occurs. It might be that temperamental characteristics that result in highly overt behavioral manifestations (e.g., shyness) lend themselves to high levels of agreement. It may also be that some traits are more subject to role expectations (e.g., boys may be less prone to show fear around their fathers than around their mothers).

*Stability across time.* Table 11 presents stability estimates from the OSLC sample based on parent ratings when the children were 5 and 7 years of age. Stability of mother ratings across this time period ranged from .50 to .79, with a mean stability estimate of .65 across scales. Stability estimates for father ratings ranged from .48 to .76, with a mean stability estimate of .63 across scales. This demonstrates a very high level of stability in parental ratings. These stability estimates, however, must be interpreted with caution, because it is not yet possible to identify the sources of this stabil-

**Table 11** Stability of Children's Behavior Questionnaire Scales from 5 to 7 Years of Age

	Mothers (N = 114)	Fathers (N = 114)
Positive Anticipation	.61*	.63*
High Intensity Pleasure	.63*	.71*
Smiling/Laughter	.55*	.62*
Activity Level	.71*	.70*
Impulsivity	.73*	.71*
Shyness	.79*	.76*
Discomfort	.50*	.55*
Fear	.63*	.61*
Anger/Frustration	.68*	.58*
Sadness	.65*	.62*
Soothability	.61*	.54*
Inhibitory Control	.73*	.64*
Attentional Focusing	.66*	.56*
Low Intensity Pleasure	.68*	.48*
Perceptual Sensitivity	.57*	.67*

\* $p < .05$ .

ity. As noted previously, because we did not have test-retest data over a short period, we were not able to estimate how much of the change was due to unreliability and how much was due to development. Nevertheless, although test-retest reliabilities (that is, two measurements taken over a short period) have not yet been performed for the CBQ, these findings suggest that such correlations would be substantial. In addition, there are likely parent biases that serve to inflate stability over time, but are unrelated to underlying constructs.

*Predicting social behavior patterns.* Concurrent relations between temperament as assessed via the CBQ and parent-reported socialization-relevant traits were examined (Rothbart, Ahadi, & Hershey, 1994). For a sample of 80 six- to seven-year-olds, mothers were administered the CBQ along with a set of scales designed to assess the traits of aggressiveness, empathy, guilt/shame, help-seeking, and negativity. The CBQ scales were collapsed to form global assessments of Surgency (Extraversion), Negative Affectivity, and Effortful Control. The CBQ scales of Positive Anticipation, High Intensity Pleasure, Smiling/Laughter, Impulsivity, and Shyness (reverse scored) were aggregated to provide an index of Surgency (Extraversion). To assess Negative Affectivity, the scale scores of Discomfort, Fear, Anger/Frustration, Sadness, and Soothability (reverse scored) were aggregated. Scores on the scales of Inhibitory Control, Attentional Focusing, Low Intensity Pleasure, and Perceptual Sensitivity were aggregated to form a measure of Effortful Control.

Regression analyses revealed substantial relations be-

tween the temperament dimensions and socialization-relevant traits. Aggression was positively related to Negative Affectivity and Surgency (Extraversion). Empathy and Guilt/Shame were both strongly positively related to Effortful Control. Guilt/Shame and Empathy were also positively related to Negative Affectivity. Tendencies to seek help and to negatively evaluate and/or oppose new activities were both positively related to Negative Affectivity.

Because Negative Affectivity was related to both pro- and antisocial behavior patterns, these relations were analyzed with respect to the scales underlying Negative Affectivity. Whereas antisocial traits such as aggression were positively related to externalizing aspects of Negative Affectivity such as anger, these antisocial traits were not related to internalizing aspects such as sadness. Conversely, prosocial traits such as guilt/shame were positively related to internalizing aspects of Negative Affectivity such as sadness, but unrelated to the externalizing aspects (Rothbart, Ahadi, & Hershey, 1994).

## DISCUSSION

The CBQ was developed to provide a differential measure of childhood temperament conforming to Rothbart's (1981, 1989; Rothbart & Derryberry, 1981; Rothbart & Posner, 1985) reactive and self-regulative model of temperament. Structural analyses of the CBQ reliably indicated that three major dimensions of temperament underlie the scales assessed in the CBQ for children aged 3 to 7 years. We labeled these dimensions Extraversion (Surgency), Negative Affectivity, and Effortful Control. Very similar factor structures emerged from analyses of child behavior in both China and Japan.

These factors showed considerable conceptual similarity to those extracted from parent report questionnaires for over 1,000 three- to seven-year-old children in the Australian Temperament Study (Sanson et al., 1994). The questionnaire used in that study was Thomas and Chess's (1977) Childhood Temperament Questionnaire. Sanson and colleagues noted similarities between their Sociability or Approach factor and Extraversion, their Persistence factor and Effortful Control, and their Irritability/Inflexibility factor and Negative Affectivity. An additional Activity/Mood factor and small factors for Rhythmicity and Threshold (the latter a highly mixed scale) were also extracted from the Thomas and Chess items (Sanson et al., 1994).

With respect to structural models of personality identified in studies of adults, the CBQ dimensions of Extraversion (Surgency) and Negative Affectivity appear to be quite similar to the major dimensions of Extraversion/Positive Affect and Neuroticism/Nega-

tive Affect found in factor analyses of adult behaviors and personality characteristics (e.g., Eysenck & Eysenck, 1985; Tellegen, 1985). We have suggested elsewhere (Ahadi & Rothbart, 1994) that the third dimension, Effortful Control, may be developmentally related to a third major dimension of adult personality variously labeled Constraint (e.g., Tellegen, 1985), Superego Strength (e.g., Cattell, Eber, & Tatsuoka, 1970), Conscientiousness (e.g., Costa & McCrae, 1985; Goldberg, 1993), and Psychoticism (e.g., Eysenck & Eysenck, 1985), and have now found evidence for these suggested links (Rothbart, Ahadi, & Evans, 2000).

We have also pointed out the relation between these three factors and affective-motivational systems studied by researchers exploring the psychobiology of personality and attention (Rothbart, Ahadi, & Evans, 2000; Rothbart, Derryberry, & Posner, 1994). These systems include a positive affect and approach system, negative affect systems of fear and frustration, and the anterior or executive attentional system (see also the review by Rothbart & Bates, 1998). One surprising finding in our factor analysis, however, was the loading of Positive Anticipation on the Negative Affectivity factor. One possible reason for this finding is a link between the expectancy of a positive outcome and the frustration or sadness experienced when that goal is not satisfied. It is interesting that the Chinese sample did not show this high loading, but more research will be necessary to understand these results. A second surprising finding was the loading of Smiling/Laughter on the Effortful Control factor for both the U.S. and Japanese samples. We have speculated that in cultures that encourage expression of positive affect, Smiling/Laughter may be related to Effortful Control as well as to positive affect (Ahadi et al., 1993). Another possibility is that alerting systems supporting smiling and laughter may also support executive attention processes. Future research is needed to better understand these relations.

The results of the CFAs were in agreement with the results from the EFAs. In terms of both location and magnitude, factor loadings resulting from the restricted factor analysis matched those identified in the exploratory analyses. Such findings were not unexpected, because the principal axis algorithm was designed to reveal simple structure, the most parsimonious structural model underlying covariation among the items.

With respect to reliability, the CBQ scales demonstrated adequate internal consistency. Psychometrically, however, it would be ideal if scale reliabilities were consistently in the .80 region. Examination of internal consistency estimates and analysis of item-total correlations suggests that future revisions of the CBQ

would benefit from greater homogeneity in the item content within certain scales. As noted previously, test-retest reliability data were necessary both for estimating the temporal stability of the CBQ measures and for evaluating possible developmental changes in temperament. With respect to translations of the CBQ, it will also be important to continue to rigorously examine translation accuracy via back translation, as well as to examine the applicability of items to specific cultures.

Parents show substantial levels of agreement in ratings of their children on the CBQ. It is important to ask, however, why even greater parental agreement was not found. Agreement did appear to vary systematically across scales, with some scales showing higher levels of agreement, both within the same sample across time and in independent samples. It is possible that the reports on temperamental shyness, inhibitory control, and impulsivity were more reliable because these characteristics were more easily observed, and that the child's felt discomfort was less observable. Children's smiling and laughter, on the other hand, may differ in response to the eliciting behavior of the two parents, as Goldsmith and Campos (1990) have found with younger children. In their research, parents did not have a strong agreement with each other in their reports of smiling and laughter, but mothers' and fathers' separate reports agreed with the parents' actual (and differing) elicitation of the child's positive affect. Both contextual factors and familiarity are likely to be important determinants of parental agreement in ratings of temperament.

Our belief that contextual factors may have important implications for parental agreement also stems from the stability of parent ratings across time. There were high levels of stability in parents' ratings of children at 5 and 7 years of age. Thus, although parents may not completely agree about what they observe in their children, they appear to be very consistent in their ratings of what they observe. Another contributor to longitudinal stability could be informant bias, but we have recently found that temperament measures assessed in the laboratory for infants predict CBQ parent reports of child temperament at age 7 (Rothbart, Derryberry, & Hershey, 2000).

Two other studies have compared laboratory observations of children's behavior with parent reports of temperament using the CBQ. One set of data was reported by Kochanska and colleagues (Kochanska, Murray, Jacques, Koenig, & Vandergeest, 1996). In the CBQ, Kochanska's Fearfulness was an aggregate score of Fear, Shyness, and Discomfort scale scores. Inhibitory Control was an aggregate of two scales, Inhibitory Control and, reversed, Impulsivity. Moder-

ate correlations were found between CBQ mother ratings and aggregate laboratory scores: for Inhibitory Control,  $r = .30$  at age 3 and  $r = .42$  at age 4; for Fear,  $r = .23$  at age 3 and  $r = .43$  at age 4. Similar agreement between laboratory observations and CBQ parent reports of inhibitory control have been reported by Carlson (1997). Additional convergent validation is offered in our predictions based on laboratory observations of temperament during infancy: the same or similar constructs were predicted using the CBQ when the children were 7 years old (Rothbart, Derryberry, et al., 2000). These findings are important because they offer mutual validation of parent reports and laboratory assessments of temperament.

## CONCLUSION

The CBQ was developed to provide a comprehensive assessment of reactive and self-regulative temperamental behavior patterns in young children. The scales comprising the CBQ show adequate internal consistency and good temporal stability. Structural analyses of the CBQ scales reveal good construct validity. The factor structure remains consistent across age groups and across cultures. Evidence for convergent validity comes from a number of sources, including parental agreement and prediction of social and laboratory behavior patterns. Future research will include revisions of the CBQ to enhance its reliability, and possible development of an accompanying laboratory assessment battery similar to the Laboratory Temperament Assessment Battery (LABTAB; Goldsmith & Rothbart, 1991). LABTAB provides laboratory measurements to accompany parent ratings provided by the IBQ (Rothbart, 1981) and the Toddler Behavior Assessment Questionnaire (TBAQ; Goldsmith, 1996). Additional scales have been developed for the IBQ and TBAQ, to provide more differentiated measures for infants and toddlers (see review by Putnam, Ellis & Rothbart, in press).

## ACKNOWLEDGMENTS

This research was supported by National Institute of Mental Health Grants MH43361 and MH01471 to the first author. The authors are grateful to Hill Goldsmith and Grazyna Kochanska for generously sharing Children's Behavior Questionnaire data with them, and to all the families who participated in these studies.

## ADDRESSES AND AFFILIATIONS

Corresponding author: Mary K. Rothbart, Dept. of Psychology, 1227 University of Oregon, Eugene, OR

97403-1227; e-mail: maryroth@oregon.uoregon.edu. Karen L. Hershey is also at the University of Oregon; Stephan A. Ahadi is at MetriTech, Inc., Champaign, IL; and Phillip Fisher is at the Oregon Social Learning Center, Eugene, OR.

## APPENDIX A

### CHILDREN'S BEHAVIOR QUESTIONNAIRE SCALE DEFINITIONS AND SAMPLE ITEMS

*Activity Level.* Gross motor activity, including rate and extent of locomotion. "Seems always in a big hurry to get from one place to another."

*Anger/Frustration.* Negative affectivity related to interruption of ongoing tasks or goal blocking. "Has temper tantrums when s(he) doesn't get what s(he) wants."

*Attentional Focusing.* Capacity to maintain attentional focus on task-related channels. "When picking up toys or other jobs, usually keeps at the task until it's done."

*Discomfort.* Negative affectivity related to sensory qualities of stimulation, including intensity; rate; or complexities of light, movement, sound, and texture. "Is not very bothered by pain."

*Fear.* Negative affectivity, including unease, worry, or nervousness, which is related to anticipated pain or distress and/or potentially threatening situations. "Is not afraid of large dogs and/or other animals."

*High Intensity Pleasure.* Pleasure or enjoyment related to situations involving high stimulus intensity, rate, complex-

ity, novelty, and incongruity. "Likes going down high slides or other adventurous activities."

*Impulsivity.* Speed of response initiation. "Usually rushes into an activity without thinking about it."

*Inhibitory Control.* Capacity to plan and to suppress inappropriate approach responses under instructions or in novel or uncertain situations. "Can lower his/her voice when asked to do so."

*Low Intensity Pleasure.* Pleasure or enjoyment related to situations involving low stimulus intensity, rate, complexity, novelty, and incongruity. "Rarely enjoys just being talked to."

*Perceptual Sensitivity.* Detection of slight, low-intensity stimuli from the external environment. "Notices the smoothness or roughness of objects s(he) touches."

*Positive Anticipation.* Amount of excitement and anticipation for expected pleasurable activities. "Gets so worked up before an exciting event that s(he) has trouble sitting still."

*Sadness.* Negative affectivity and lowered mood and energy related to exposure to suffering, disappointment, and object loss. "Cries sadly when a favorite toy gets lost or broken."

*Shyness (versus Social Approach).* Slow or inhibited (versus rapid) speed of approach and discomfort (versus comfort) in social situations. "Often prefers to watch rather than join other children playing."

*Smiling/Laughter.* Positive affect in response to changes in stimulus intensity, rate, complexity, and incongruity. "Laughs a lot at jokes and silly happenings."

*Soothability (and Falling Reactivity).* Rate of recovery from peak distress, excitement, or general arousal. "Has a hard time settling down for a nap."

## APPENDIX B

### MEAN CHILDREN'S BEHAVIOR QUESTIONNAIRE ITEM SCALE SCORES FOR THREE AGE GROUPS

	3-Year-Olds <sup>a</sup>				4- and 5-Year Olds <sup>b</sup>				6- and 7-Year Olds <sup>c</sup>			
	M	SD	Minimum	Maximum	M	SD	Minimum	Maximum	M	SD	Minimum	Maximum
Activity Level	5.03	.86	3.08	6.85	4.87	.77	2.38	6.69	4.82	.82	2.85	7.00
Anger/Frustration	4.68	.83	1.85	6.15	4.51	.79	1.42	6.77	4.42	.84	1.38	6.92
Attentional Focusing	4.41	.68	2.93	6.25	4.50	.68	2.50	6.63	4.53	.67	2.36	6.43
Discomfort	4.14	.83	1.36	6.58	4.23	.88	1.67	6.58	4.28	.85	1.67	6.58
Fear	3.75	.82	1.83	5.58	3.93	.94	1.50	6.58	3.81	.92	1.00	6.42
High Intensity Pleasure	4.92	.94	2.46	6.92	4.97	.90	2.54	6.92	5.02	.84	2.50	7.00
Impulsivity	4.68	.76	2.77	6.54	4.53	.73	2.62	6.69	4.53	.76	2.23	6.62
Inhibitory Control	4.26	.73	2.46	6.80	4.75	.83	1.62	6.92	4.91	.80	2.50	6.77
Low Intensity Pleasure	5.59	.78	3.46	7.00	5.64	.73	2.69	6.85	5.53	.64	2.92	6.85
Perceptual Sensitivity	4.85	.74	2.83	6.42	4.98	.65	2.92	6.67	4.99	.72	2.67	7.08
Positive Anticipation	5.15	.69	3.08	7.00	5.20	.71	2.58	6.77	5.13	.69	2.77	6.92
Sadness	4.12	.66	1.30	5.75	4.19	.73	1.50	6.25	4.23	.75	1.91	7.25
Shyness	3.58	1.22	1.15	6.54	3.57	1.21	1.00	6.69	3.26	1.12	1.00	6.54
Smiling/Laughter	5.61	.89	2.33	7.00	5.68	.75	2.92	7.00	5.72	.64	2.54	7.00
Soothability	4.44	.71	2.31	5.85	4.51	.71	1.75	6.38	4.57	.75	1.62	6.92

<sup>a</sup> N = 149.

<sup>b</sup> N = 517.

<sup>c</sup> N = 341.

## REFERENCES

- Ahadi, S. A., & Rothbart, M. K. (1994). Temperament, development and the Big Five. In G. Kohnstamm & C. Halverson (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 189–207). Hillsdale, NJ: Erlbaum.
- Ahadi, S. A., Rothbart, M. K., & Ye, R. M. (1993). Children's temperament in the US and China: Similarities and differences. *European Journal of Personality*, 7, 359–377.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- Bentler, P. M. (1989). *EQS structural equations program manual*. Los Angeles: BMDP Statistical Software.
- Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.
- Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.
- Capaldi, D. M., & Rothbart, M. K. (1992). Development and validation of an early adolescent temperament measure. *Journal of Early Adolescence*, 12, 153–173.
- Carlson, S. M. (1997). *Individual differences in inhibitory control and children's theory of mind*. Unpublished doctoral dissertation, University of Oregon, Eugene.
- Cattell, R. B., Eber, H. J., & Tatsuoka, M. (1970). Handbook for the sixteen personality factor questionnaire (16 PF) in clinical, educational, industrial, and research psychology, for use with all forms of the test. Champaign, IL: Institute for Personality and Ability Testing.
- Chugani, H. T., Phelps, M. E., & Mazziotta, J. C. (1987). Positron emission tomography study of human brain functional development. *Annals of Neurology*, 22, 487–497.
- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, 22, 491–569.
- Depue, R. A., & Iacono, W. G. (1989). Neurobehavioral aspects of affective disorders. In M. R. Rosenzweig & L. Y. Porter (Eds.), *Annual review of psychology* (Vol. 40, pp. 457–492). Palo Alto, CA: Annual Reviews.
- Derryberry, D., & Rothbart, M. K. (1988). Arousal, affective and attentional components of adult temperament. *Journal of Personality and Social Psychology*, 55, 953–966.
- Derryberry, D., & Rothbart, M. K. (1997). Reactive and effortful processes in the organization of temperament. *Development and Psychopathology*, 9, 633–652.
- Diener, E., & Emmons, R. A. (1984). The independence of positive and negative affectivity. *Journal of Personality and Social Psychology*, 47, 1105–1117.
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum.
- Fagot, B. I., & Leve, L. D. (1998). Teacher ratings of externalizing behavior at school entry for boys and girls: Similar early predictors and different correlates. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 39(4), 555–566.
- Fisher, P. A. (1993). Temperament goodness of fit and psychosocial adjustment in children. Unpublished doctoral dissertation. University of Oregon, Eugene.
- Fiske, D. W. (1966). On the coordination of personality concepts and their measurement. *Human Development*, 9, 74–83.
- Fiske, D. W. (1971). *Measuring the concepts of personality*. Chicago: Aldine.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 26–34.
- Goldsmith, H. H. (1996). Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development*, 67, 218–235.
- Goldsmith, H. H., Buss, K. A., & Lemery, K. S. (1997). Toddler and childhood temperament: Expanded content, stronger genetic evidence, new evidence for the importance of environment. *Developmental Psychology*, 33(6), 891–905.
- Goldsmith, H. H., & Campos, J. J. (1990). The structure of temperamental fear and pleasure in infants: A psychometric perspective. *Child Development*, 61, 1944–1964.
- Goldsmith, H. H., & Rothbart, M. K. (1991). Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 249–272). New York: Plenum.
- Gray, J. A. (1982). *The neuropsychology of anxiety*. Oxford, U.K.: Oxford University Press.
- Gray, J. A. (1987). *The psychology of fear and stress*. New York: Cambridge University Press.
- Huttenlocher, P. R. (1990). Morphometric study of human cerebral cortex development. *Neuropsychologia*, 28, 517–527.
- Izard, C. E. (1977). *Human emotions*. New York: Plenum.
- Kagan, J., Reznick, J. S., & Snidman, N. (1988). Biological bases of childhood shyness. *Science*, 240, 167–171.
- Kochanska, G., DeVet, K., Goldman, M., Murray, M., & Putnam, S. (1994). *Maternal reports of conscience, development, and temperament in young children*. *Child Development*, 65, 852–868.
- Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandergest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development*, 67, 490–507.
- Kusanagi, E. (1993). *A psychometric examination of the Children's Behavior Questionnaire* (Annual Report, 1991–1992, No. 15). Sapporo, Japan: Hokkaido University, Faculty of Education, Research and Clinical Center for Child Development.
- Martin, R. P., Wisenbaker, J., & Huttunen, M. (1994). Review of factor analytic studies of temperament measures based on the Thomas–Chess structural model: Implications for the Big Five. In G. Kohnstamm & C. Halverson (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 157–172). Hillsdale, NJ: Erlbaum.
- McDevitt, S. C., & Carey, W. B. (1981). Stability of ratings vs. perceptions of temperament from early infancy to 1–3 years. *American Journal of Orthopsychiatry*, 51, 342–345.

- Norusis, M. J. (1994). *SPSS advanced statistics (Version 6.1)*. Chicago: SPSS.
- Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behavioral and Brain Sciences*, *5*, 407–467.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
- Posner, M. I., & Raichle, M. E. (1994). *Images of the mind*. New York: Scientific American Library.
- Posner, M. I., & Rothbart, M. K. (1998). Attention, self-regulation and consciousness. *Philosophical Transactions of the Royal Society of London: Biological Sciences*, *353*, 1915–1927.
- Putnam, S. P., Ellis, L. K., & Rothbart, M. K. (in press). The structure of temperament from infancy through adolescence. In A. Elias & A. Angleitner (Eds.), *Advances/proceedings in research on temperament*. Germany: Pabst Scientific.
- Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, *52*, 569–578.
- Rothbart, M. K. (1988). Temperament and the development of inhibited approach. *Child Development*, *59*, 1241–1250.
- Rothbart, M. K. (1989). Temperament and development. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 187–247). New York: Wiley.
- Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology*, *78*, 122–135.
- 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., & Bates, J. E. (1998). Temperament. In N. Eisenberg (Ed.), W. Damon (Series Ed.), *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (5th ed., pp. 105–176). New York: Wiley.
- Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (Vol. 1, pp. 37–86). Hillsdale, NJ: Erlbaum.
- Rothbart, M. K., Derryberry, D., & Hershey, K. (2000). Stability of temperament in childhood: Laboratory infant assessment to parent report at seven years. In V. J. Molfese & D. L. Molfese (Eds.), *Temperament and personality development across the life span* (pp. 85–119). Hillsdale, NJ: Erlbaum.
- Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychobiological approach to the development of temperament. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior* (pp. 83–116). Washington, DC: American Psychological Association.
- Rothbart, M. K., & Mauro, J. A. (1990). Questionnaire measures of infant temperament. In J. W. Fagen & J. Colombo (Eds.), *Individual differences in infancy: Reliability, stability and prediction* (pp. 441–429). Hillsdale, NJ: Erlbaum.
- Rothbart, M. K., & Posner, M. I. (1985). Temperament and the development of self-regulation. In L. C. Hartlage & C. F. Telzrow (Eds.), *The neuropsychology of individual differences: A developmental perspective* (pp. 93–123). New York: Plenum.
- Rowe, D. C., & Plomin, R. (1977). Temperament in early childhood. *Journal of Personality Assessment*, *41*, 150–156.
- Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development: Themes and variations*. New York: Oxford University Press.
- Sanson, A. V., Smart, D. F., Prior, M., Oberklaid, F., & Pedlow, R. (1994). The structure of temperament from age 3 to 7 years: Age, sex, and sociodemographic influence. *Merrill-Palmer Quarterly*, *40*, 233–252.
- Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum.
- Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.
- Thomas, A., Chess, S., & Birch, H. (1968). *Temperament and behavior: Disorders in children*. New York: New York University Press.
- Thomas, A., Chess, S., Birch, H. G., Herzig, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood*. New York: New York University Press.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affectivity: The PANAS Scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.
- Zuckerman, M. (1979). *Sensation seeking: Beyond the optimal level of arousal*. Hillsdale, NJ: Erlbaum.
- Zuckerman, M. (1990). The psychophysiology of sensation seeking. *Journal of Personality*, *58*, 313–345.