

Studying infant temperament via the Revised Infant Behavior Questionnaire

Maria A. Gartstein^{a,*}, Mary K. Rothbart^b

^a *Department of Psychology, Washington State University, P.O. Box 644820, Pullman, WA 99164-4820, USA*

^b *Department of Psychology, 1227 University of Oregon, Eugene, OR 97403-1227, USA*

Received 2 July 2002; received in revised form 27 September 2002; accepted 2 October 2002

Abstract

This study describes a revision of a widely used parent-report measure of infant temperament, the Infant Behavior Questionnaire (IBQ; Rothbart, 1981). A rationally derived instrument was developed that included nine new scales and minor modifications of the seven scales of the IBQ. Parents of 360 infants, equally distributed over three age groups: 3–6 months; 6–9 months; and 9–12 months of age, participated. Conceptual and item analyses provided support for 14 of the 16 proposed scales, demonstrating satisfactory internal consistency. Inter-rater reliability was evaluated, with evidence of moderate agreement between primary and secondary caregivers. Monomethod discriminant validity was demonstrated through an examination of correlations among the Infant Behavior Questionnaire—Revised (IBQ-R) scale scores. Results of the factor analytic procedure were consistent with three broad dimensions of Surgency/Extraversion, Negative Affectivity, and Orienting/Regulation. Developmental and gender differences were also noted for a number of the IBQ-R scales. Specifically, older infants received higher scores on Approach, Vocal Reactivity, High Intensity Pleasure, Activity, Perceptual Sensitivity, Distress to Limitations, and Fear, whereas younger infants' scores were higher for Low Intensity Pleasure, Cuddliness/Affiliation, and Duration of Orienting. Male infants obtained higher scores on Activity and High Intensity Pleasure, and female infants were rated higher on the Fear scale.

© 2002 Elsevier Science Inc. All rights reserved.

Keywords: Infancy; Temperament; Measurement; Parent-report

* Corresponding author. Tel.: +1-509-335-6154; fax: +1-509-335-5043.

E-mail addresses: gartstma@wsu.edu (M.A. Gartstein), maryroth@oregon.uoregon.edu (M.K. Rothbart).

1. Introduction

Interest in the study of infant temperament has increased steadily during recent years, as early individual differences have been increasingly related to later personality and social development (Kagan, 1998; Rothbart & Bates, 1998). During this period, developing techniques for assessment of infant temperament has become an important area of study, with a variety of laboratory and parent-report questionnaire approaches now in use (Rothbart, Chew, & Gartstein, 2001). Since its introduction in 1981, the Infant Behavior Questionnaire (IBQ; Rothbart, 1981) has been one of the most widely used of these parent-report measures. It has been used in a variety of investigations that have provided further evidence for its reliability and validity (e.g., Clark, Hyde, Essex, & Klein, 1997; Fagen, Ohr, Singer, & Fleckenstein, 1987; Worobey, 1997). This study describes a revision of the IBQ, designed to reflect what has been learned about temperament in the two decades since the instrument was introduced in 1981. In this study, nine additional scales were created, and seven earlier IBQ scales modified to incorporate new content, while maintaining discriminant validity.

Development of the IBQ-R was based on a fine-grained approach to the assessment of infant temperament, allowing investigation of relations among dimensions of temperament that may be obscured in the study of more global constructs (Derryberry & Rothbart, 1988). These relations are important in that one dimension of temperament may influence the expression of another. For example, in preschool and school-age children, both attentional effortful control and fear appear to moderate the expression of approach and aggressive tendencies, and to promote the development of conscience (Kochanska, 1993; Rothbart, Ahadi, & Evans, 2000; Rothbart, Ahadi, & Hershey, 1994; Rothbart & Bates, 1998).

1.1. *The Infant Behavior Questionnaire*

Caregiver or parent report instruments take advantage of caregivers' extensive opportunities to observe young children across a broad array of contexts. Cautions regarding parental report have been raised, however, with concerns about potential error (Kagan, 1994; Rothbart & Bates, 1998; Rothbart & Goldsmith, 1985). For example, error may occur due to the influence of social desirability, the caregivers' limited knowledge of the infant's behavior and its meaning, unfamiliarity with the behavior of other infants with whom the child is compared, and/or limited accuracy in caregivers' memories of events involving the infant.

These concerns can be at least partially addressed by careful construction and presentation of items, asking about only recently occurring events, and inquiring about concrete infant behaviors rather than asking the parents to make abstract or comparative judgments (Rothbart & Goldsmith, 1985). The original IBQ attempted to address these concerns (Rothbart, 1981). Caregivers were asked to report, on a 7-point scale, the relative frequency of occurrence of specified infant reactions in concrete situations during the previous week, or for some items, 2 weeks. This format is likely to minimize problems associated with recall, and to limit biases associated with more global questions that would require respondents to aggregate information across contexts or situations prior to answering.

The IBQ also does not require caregivers to make comparative judgments that would be difficult if they lacked familiarity with other infants. Finally, the response format of the IBQ

presents sets of items based on the context or situation eliciting the infant's reactions (e.g., bathing and dressing), which may serve to enhance specific recall and limit social desirability.

Rothbart and Derryberry (1981) provided the theoretical basis for the original IBQ (Rothbart, 1981). In their approach, temperament is defined as constitutionally based individual differences in reactivity and self-regulation, with constitutional referring to the relatively enduring biological make-up of the individual, influenced by heredity, maturation, and experience. Reactivity refers to the arousability of emotional, motor, and attentional responses, assessed by threshold, latency, intensity, time to peak intensity, and recovery time of reactions. Self-regulation refers to processes such as attention, that can serve to modulate reactivity (Rothbart & Derryberry, 1981).

Constructs or dimensions of temperament addressed by the original IBQ were based on the work of Thomas and colleagues (Thomas, Chess, & Birch, 1968; Thomas, Chess, Birch, Hertzig, & Kom, 1963), Escalona (1968), Shirley (1933) and others (Diamond, 1957). IBQ items were rationally derived from the constructs to be assessed, with scale scores formed by aggregating individual items across a range of contexts. Scales were developed to assess: Activity Level, Smiling and Laughter, Fear, Distress to Limitations, Duration of Orienting, and Soothability. The Vocal Reactivity scale was not part of the original questionnaire, but was added later as an element of an IBQ validation study (Rothbart, 1986).

A number of studies have provided validation support for the original IBQ (Bridges, Palmer, Morales, Hurtado, & Tsai, 1993; Crockenberg & Acredolo, 1983; Goldsmith & Rothbart, 1991; Reznick, Gibbons, Johnson, & McDonough, 1989; Rothbart, 1986). Reliability, convergent validity, and relative stability have also been demonstrated for the IBQ with infants as young as 2 weeks of age (Worobey, 1986; Worobey & Blajda, 1989).

The IBQ has been used to achieve at least three major goals. The first has been to measure individual differences in reactivity and regulation, often in the context of structured laboratory tasks and/or physiological measures (Calkins, Fox, & Marshall, 1996; Goldsmith & Campos, 1990; Huffman et al., 1998; Kochanska, Coy, Tjebkes, & Husarek, 1998; Rothbart, Derryberry, & Hershey, 2000).

The IBQ has also been employed to identify the structure of infant temperament. Factor analytic work with the IBQ has generally yielded dimensions related to Positive and Negative Affectivity (Goldsmith & Campos, 1990; Kochanska et al., 1998; Rothbart, 1986). These factors were extracted for mothers and fathers, with Smiling and Laughter, Duration of Orienting, Soothability, and Activity loading onto the Positive Affectivity factor; Distress to Limitations and Fear comprising the Negative Affectivity factor (Kochanska et al., 1998). Goldsmith and Campos (1990) performed exploratory factor analyses with the IBQ, considering one to three-factor solutions, and different rotations. Although two factors were extracted, the authors chose to utilize a single factor (combining Positive and Negative Emotionality into a single dimension) on the basis of a scree test, indicating that all except the first factor may be due to random variation (Goldsmith & Campos, 1990). The addition of a number of new scales to the IBQ-R in this revision will facilitate efforts to further study the structure of temperament in infancy via parent-report questionnaires, by allowing researchers to detect a more complex structure.

A third goal has been to evaluate relations between temperament, socialization, and parental and family functioning (Belsky, 1996; Clark et al., 1997; Seifer, Schiller, Sameroff, Resnick, & Riordan, 1996; Ventura & Stevenson, 1986). Seifer et al. (1996) demonstrated that temperament

measures, including IBQ Distress to Limitations, accounted for unique portions of the variance associated with attachment security, measured via Q-sort, whereas measures of maternal sensitivity did not. In another study, a higher level of infant Distress to Limitations was associated with maternal physical symptoms, feelings of incompetence, negative reinforcement from the infant, and negatively associated with spousal emotional support, controlling for variables such as maternal education, employment, and bottle/breast feeding (Clark et al., 1997). Ventura and Stevenson (1986) reported associations between IBQ scales and measures of parental psychopathology, including depression and somatic complaints, as well as socio-demographic factors (interestingly, infants from higher socio-economic status backgrounds were perceived as exhibiting lower levels of Soothability and higher levels of Distress to Limitations). These studies have also provided further evidence for reliability and validity of the IBQ.

1.2. Proposed revision of the instrument

Significant advances in research addressing the development of temperament have also occurred since the introduction of the IBQ. A number of additional temperament measures have also been introduced, designed to evaluate temperament throughout the lifespan, and using laboratory (Goldsmith & Riley, 1994; Goldsmith & Rothbart, 1996) and questionnaire methodologies (Putnam, Ellis, & Rothbart, 2001; Rothbart, Ahadi, et al., 2000; Rothbart, Ahadi, Hershey, & Fisher, 2001). Research involving the Children's Behavior Questionnaire (CBQ) is of particular relevance for development of the IBQ-R. The CBQ is a highly differentiated assessment of temperament for children between 3 and 7 years of age (Rothbart, Ahadi, et al., 2001), including 15 scales that have consistently clustered into three broad factors (Ahadi, Rothbart, & Ye, 1993; Kochanska, DeVet, Goldman, Murray, & Putnam, 1994). The first factor, Surgency/Extraversion, is defined by scale scores of Impulsivity, High Intensity Pleasure, Activity Level, and negatively loading Shyness, with significant loadings for Positive Anticipation and Smiling and Laughter. The second factor, Negative Affectivity, is defined by scales of Sadness, Discomfort, Anger/Frustration, Fear, and loading negatively, Falling Reactivity/Soothability. The third factor, Effortful Control, is defined by scales addressing Low Intensity Pleasure, Inhibitory Control, Attentional Focusing, and Perceptual Sensitivity. Smiling and Laughter has also demonstrated significant loadings on this factor in U.S. and Japanese, but not Chinese, samples (Ahadi et al., 1993).

Consistency between parent report on the CBQ and indicators of temperament based on home and laboratory observations has been demonstrated (Carlson, 1997; Gerardi-Caulton, in press; Kochanska, Murray, Jacques, Koenig, & Vandegest, 1996; Schaughency & Fagot, 1993). The success of the CBQ has suggested that the "downward extension" of a number of the CBQ scales would provide more detailed and meaningful information regarding individual differences in temperament in the first year of life. Thus, new scales for the IBQ-R test version were primarily downward extensions from the CBQ, including Positive Anticipation (Approach), Falling Reactivity, High and Low Intensity Pleasure, Perceptual Sensitivity, and Sadness.

Approach was defined as positive excitement and rapid approach toward pleasurable activities. *Falling Reactivity* was defined as rate of recovery from peak distress, excitement, or general arousal, and reflects the infant's ability to regulate his/her own state. *High and Low Intensity Pleasure* both refer to enjoyment related to stimulus characteristics (i.e., high or low intensity,

rate, complexity, novelty, and incongruity). The intensity distinction is consistent with optimal level theories (e.g., Bell, 1974; Strelau, 1983) stressing variability in preferences for stimulation at varying levels. The high intensity pleasure dimension is also related to the construct of sensation seeking (Zuckerman, 1990). *Perceptual Sensitivity* refers to the detection of slight, low intensity environmental stimuli, corresponding to the Thomas et al. (1963) temperament dimension of Threshold. *Sadness* is defined as general low mood, or lowered mood and activity related to personal suffering, physical state, object loss, or inability to perform a desired action.

Cuddliness/Affiliation is not included in the CBQ, but was included in the IBQ-R based on contemporary theory and research addressing the affiliativeness dimension of temperament (Panksepp, 1998). Early studies of social development have identified individual differences in cuddliness (Schaffer & Emerson, 1964a, 1964b), and maternal reports of infant cuddliness have been negatively associated with indicators of distress proneness and later behavior problems (Stevenson, Thompson, & Sonuga-Barke, 1996). *Social Fear*, defined as slow/inhibited approach and/or distress in social situations that involve novelty or uncertainty, was developed on the basis of the evidence pointing to shyness as an important domain of individual differences (Kagan, Reznick, & Snidman, 1988). Shyness is assessed by the CBQ as slow or inhibited approach in situations involving novelty or uncertainty. *Attentional Shifting* was also assessed, defined as ease of transferring attentional focus from one activity/task to another.

1.3. Development of temperament

Manifestations of temperament change over time as a child develops, with rapid development during infancy (Rothbart, 1989). Positive emotionality (e.g., smiling, etc.) is rarely expressed during the newborn period, but is observed more reliably between 2 and 3 months of age, increasing in expression throughout the first year of life (Rothbart, 1989). Levels of activity, approach, distress to limitations, and fear have been found to increase throughout the first year of life (Carranza, Perez-Lopez, Gonzalez, & Martinez-Fuentes, 2000; Rothbart, 1986, 1988).

Activity level has been found to increase throughout infancy (Buss & Plomin, 1975). Researchers have reported a U-shaped tendency for the developmental trajectory of anger reactions in infancy (Carranza et al., 2000; Rothbart, 1981). The decrease in anger responses occurring between 2 and 6 months of age has been linked to greater flexibility in attention shifting (Johnson, Posner, & Rothbart, 1991). In the second half of the first year, infants are likely to respond with anger when unable to grasp an attractive stimulus that has been placed out of reach, or when a caregiver has removed a forbidden object. Fear generally increases throughout the second half of the first year of life (Carranza et al., 2000; Rothbart, 1986, 1988), with development of inhibition of approach toward novel and/or intense stimuli (Rothbart, 1988, 1994).

The developmental course of attentional orienting has been described as U-shaped in the first year of life (Ruff & Rothbart, 1996). Carranza et al. (2000), for example, have noted decreases in Duration of Orienting between 6 and 9 months, followed by an increase between 9 and 12 months of age. Toward the end of the first year, skills associated with development of the executive attention system may come to contribute to the flexibility of orienting reactions (Posner & Rothbart, 1991). The increasing maturation of this attention system over the preschool years has been linked to goal-oriented and planned behavior, allowing the child to engage and persist longer in interactions with objects. Infants also gain communication skills

rapidly during the first year of life (Worobey, 1989), and thus might be expected to exhibit greater vocal reactivity over time.

1.4. Gender differences in temperament

Although a number of gender differences have been reported for older children and adults, few have been found for children younger than one year of age (Bates, 1987; Rothbart, 1989). Differences in infancy have been limited to activity level and fear/behavioral inhibition. Higher activity level and approach have been reported for boys (Campbell & Eaton, 1999; Maziade, Boudreault, Thivierge, Caperaa, & Cote, 1984), with girls exhibiting greater hesitation in approaching novel objects (Martin, Wisenbaker, Baker, & Huttunen, 1997; Rothbart, 1988). Campbell and Eaton applied meta-analytic procedures to summarize 46 studies addressing activity level in infancy, estimating the size of the gender difference at 0.2 standard deviations. Gender differences in approach-withdrawal have been reported for cross-cultural samples (Carey & McDevitt, 1978; Hsu, Soong, Stigler, Hong, & Liang, 1981; Maziade et al., 1984), with parents rating males higher in their levels of approach. Martin et al. (1997) reported a large and significant gender difference for the Distress to Novelty dimension of temperament, with 6-month-old girls receiving higher scores than boys.

In this paper, we describe IBQ-R instrument development and studies of gender and age differences by reporting: (1) item development and analysis; (2) internal consistency and inter-rater analyses; (3) a preliminary evaluation of validity, focusing on “monomethod” discriminant validity (Campbell & Fiske, 1959; Goldsmith, 1996); (4) examination of the structure of infant temperament through factor analysis; and (5) evaluation of age and gender differences for the IBQ-R. This approach provides the opportunity to address discriminant validity, with low to moderate intercorrelations expected among scales. If consistently high intercorrelations were obtained for the majority of the scales, lack of discriminant validity for the IBQ-R, and a potential susceptibility to reporting biases would be indicated (Rothbart, Chew, et al., 2001). Some low to moderate correlations were expected, because the IBQ-R scales were designed to measure regulatory as well as reactive dimensions (Rothbart et al., 1994).

Evaluation of the structure of infant temperament in this study was largely exploratory, and specific hypotheses regarding the outcome of factor analysis were not put forward. Age and gender differences, however, consistent with those reported in previous studies, were hypothesized: (1) levels of activity, approach, vocal reactivity, distress to limitations (anger), and fear were expected to increase over the first year and (2) gender differences in activity level, with boys exhibiting higher levels, and fear, with girls exhibiting higher levels, were anticipated.

Specific age-by-gender interactions were not hypothesized, although exploratory analyses of such interactions were undertaken.

2. Method

2.1. Participants

A community sample of 360 primary caregivers of infants between 3 and 12 months of age was recruited for the study, with caregivers completing the IBQ-R. The infants were equally

distributed across three age groups: 3–6, 6–9, and 9–12 months. Approximately half of the infants whose parents participated in this study were male ($n = 179$), half female ($n = 181$). Data were also collected from secondary caregivers of 26 infants, who were selected to produce a broad age range ($M = 7.5$ months; $SD = 2.5$ months, range = 4–12 months of age), and a relatively balanced gender distribution (15 males and 11 females) for infants evaluated. Participants were recruited by telephone on the basis of birth announcements published in the local newspaper for Eugene-Springfield, Oregon. Birth announcements are typically published for all the infants born in the local hospitals, unless a family specifically requests that the announcement not be made.

Four hundred and seventy-six parents were initially contacted, of whom 46 (9.7%) declined participation. Nineteen additional caregivers (4%) declined participation after initially agreeing to take part. In addition, 50 IBQ-R forms (10.5%) were either not completed or were lost in the mail. Completion of the questionnaires was encouraged by up to three follow-up phone calls to the participating parents. Thus, 76% of the contacted families participated in this study ($N = 361$), and data for one infant was later randomly chosen and excluded to maintain equal sized age groups. Respondents were mostly mothers (90%), and the secondary caregivers who completed the IBQ-R were mostly fathers (58%). Parents included in the sample were primarily Caucasian, consistent with the population of the Eugene-Springfield area. The majority of participants were married (96%), with 2.3% of the sample unmarried, and 1.7% reporting separation and/or divorce. Evaluation of the SES data for this sample indicated that the primary caregivers were mostly employed in service oriented professions. Additional demographic information is provided in [Table 1](#).

2.2. Procedure

2.2.1. Item/scale construction

A multi-phase rational scale construction method was employed, as in the development of the original IBQ ([Rothbart, 1981](#)), beginning by generating precise operational definitions for each dimension of temperament. Items, consistent with the developed definitions, were constructed in a 7-point Likert format for each of the scales. Content of items was then re-evaluated

Table 1
Demographic features of families participating in the study ($N = 360$)

Variable	Mean	SD
Family SES ^a	35.00	24.10
Family SES ^b	54.12	19.20
Gross family income ^c	41,798.30	19,1545.00
Age of primary caregiver	31.10	5.30
Years of education of primary caregiver	14.51	2.40

^a Revised Duncan Sociometric Index (TSEI2; [Stevens & Featherman, 1981](#)), a widely used indicator of occupation ranking.

^b Revised Duncan Sociometric Index computed on a subsample, excluding parents without current employment ($N = 189$).

^c In thousands of dollars.

for correspondence to operational definitions by the authors and their research group. Subsequently, items and scales that contributed to conceptual overlap between the dimensions of temperament were eliminated. Sixteen scales were initially developed (Approach, Vocal Reactivity, High Intensity Pleasure, Smiling and Laughter, Activity Level, Perceptual Sensitivity, Sadness, Distress to Limitations, Fear, Falling Reactivity, Low Intensity Pleasure, Cuddliness/Affiliation, Duration of Orienting, Soothability, Social Fear, Attention Shifting), and 14 were supported in the course of the empirical evaluation (see [Table 2](#)).

Following data collection, item analysis was performed separately for each of the three age groups in the sample (Group 1: 3–6 months, Group 2: 6–9 months, and Group 3: 9–12-month-old infants). Prior to 6 months of age, infants were assigned to Group 1, at 6 months of age they were considered to be part of Group 2, and at 9 months of age infants were included in Group 3. The sample was split by age because in the original IBQ, some items were found to be appropriate for one age group, but not others, due to rapid development during this period. Items with a large number (more than 50% of the cases) of missing responses were eliminated, as were items that failed to contribute to the internal consistency of their respective scales (i.e., item-total correlations below .3).

The second phase of conceptual analysis focused on verifying consistency between operational definitions of the scales and the items, given changes resulting from the item analysis. An attempt was again made to minimize conceptual overlap between the scales, ensuring their independence. This analytic step led to an elimination of a number of items that had contributed to non-independence between several scales, as well as to elimination of one of the scales (Attention Shifting), for which an operational definition without conceptual overlap with other scales could not be established. Finally, two of the scales (Fear and Social Fear) were combined because of their similarity in operational definitions and item content. Definitions of temperament dimensions after completion of conceptual and item analyses are listed in [Table 2](#).

2.3. Data collection

Parents of infants between 3 and 12 months of age whose births were reported in the Eugene-Springfield newspaper, were contacted by phone until 120 questionnaires were obtained for each of the three age groups evaluated in this study (Group 1: 3–6 months; Group 2: 6–9 months; and Group 3: 9–12 months). Parents were told that their participation would involve completing the IBQ-R and a demographic form, which would take about 1 hr. Parents were also informed they would be receiving a check for \$5.00, enclosed with the materials, in an effort to compensate them for their time. A subgroup of families ($n = 40$) was asked about possible participation of a secondary caregiver. Of these, three parents reported that a secondary caregiver was not available (7.5%); six declined participation of the secondary caregiver (15%; citing primarily busyness as a reason for declining participation); two parents initially agreed to have a secondary caregiver participate, but later declined (5%); and materials were never received from three additional participating secondary caregivers (7.5%). Thus, of 37 available secondary caregivers, completed questionnaires were received from 26 (70%). Each respondent was mailed a consent form with the questionnaire materials, which s/he signed and returned with the other information.

Table 2
Scale definitions: Infant Behavior Questionnaire—Revised

Approach	Rapid approach, excitement, and positive anticipation of pleasurable activities. (“When given a new toy, how often did the baby get very excited about getting it?”)
Vocal Reactivity	Amount of vocalization exhibited by the baby in daily activities. (“When being dressed undressed during the last week, how often did the baby coo or vocalize?”)
High Intensity Pleasure	Pleasure or enjoyment related to high stimulus intensity, rate, complexity, novelty, and incongruity. (“During a peek-a-boo game, how often did the baby smile?”)
Smile and Laughter	Smiling or laughter during general caretaking and play. (“How often during the last week did the baby smile or laugh when given a toy?”; shorter and different in content from the original IBQ scale)
Activity Level	Gross motor activity, including movement of arms and legs, squirming and locomotor activity. (“When put into the bath water, how often did the baby splash or kick?” similar in length/content to the original IBQ scale)
Perceptual Sensitivity	Detection of slight, low intensity stimuli from the external environment. (“How often did the baby notice fabrics with scratchy texture (e.g., wool)?”)
Sadness	Lowered mood and activity related to personal suffering, physical state, object loss, or inability to perform a desired action; general low mood. (“Did the baby seem sad when the caregiver was gone for an unusually long period of time?”)
Distress to Limitations	Fussing, crying or showing distress while (a) in a confining place or position; (b) in caretaking activities; (c) unable to perform a desired action. (“When placed on his/her back, how often did the baby fuss or protest?”; shorter, but similar in content to the original IBQ scale)
Fear	Startle or distress to sudden changes in stimulation, novel physical objects or social stimuli; inhibited approach to novelty. (“How often during the last week did the baby startle to a sudden or loud noise?”; different in content from the original IBQ)
Falling Reactivity/rate of recovery from distress	Rate of recovery from peak distress, excitement, or general arousal; ease of falling asleep. (“When frustrated with something, how often did the baby calm down within 5 min?”)
Low Intensity Pleasure	Amount of pleasure or enjoyment related to low stimulus intensity, rate, complexity, novelty and incongruity. (“When playing quietly with one of his/her favorite toys, how often did the baby show pleasure?”)
Cuddliness	Expression of enjoyment and molding of the body to being held by a caregiver. (“When rocked or hugged, during the last week, how often did the baby seem to enjoy him/herself?”)
Duration of Orienting	Attention to and/or interaction with a single object for extended periods of time. (“How often during the last week did the baby stare at a mobile, crib bumper or picture for 5 min or longer?”; similar in length/content to the original IBQ scale)
Soothability	Reduction of fussing, crying, or distress when soothing techniques are used by the caregiver. (“When patting or gently rubbing some part of the baby’s body, how often did s/he soothe immediately?”; similar in length/content to the original IBQ scale)
Social Fear ^a	Slow/inhibited approach and/or distress in social situations that involve novelty/uncertainty
Attentional Shifting	The ability to transfer attentional focus from one activity/task to another

^a This scale was eliminated following conceptual and item analyses described in [Section 3](#).

In completing the IBQ-R, parents were asked to read each description of the baby’s behavior, indicating how often the infant engaged in the various behaviors during the last week (i.e., the past 7 days). Parents responded by using a 7-point, Likert-type scale (never, very rarely, less than half the time, half the time, more than half the time, almost always, always).

Table 3

IBQ-R internal consistency: Cronbach's alphas for Group 1 (3–6 months)^a, Group 2 (6–9 months), and Group 3: (9–12 months of age)

Scale	Alpha			Initial number of items	Final number of items
	Group 1	Group 2	Group 3		
Approach	0.87	0.78	0.81	11	12
Vocal Reactivity	0.82	0.73	0.78	13	12
High Pleasure	0.77	0.76	0.80	12	11
Smile and Laughter	0.85	0.72	0.82	17	10
Activity Level	0.82	0.70	0.78	16	15
Perceptual Sensitivity	0.81	0.84	0.83	14	12
Sadness	0.85	0.85	0.71	17	14
Distress to Limitations	0.81	0.83	0.82	23	16
Fear	0.90	0.89	0.87	16	16
Falling Reactivity	0.84	0.79	0.83	18	13
Low Pleasure	0.81	0.87	0.80	14	13
Cuddliness	0.84	0.85	0.87	20	17
Duration of Orienting	0.81	0.82	0.77	15	12
Soothability	0.81	0.81	0.83	11	11

Note: The Social Fear scale initially contained 22 items, and the Attentional Shifting scale included 12 items.

^a $N = 120$ for each of the three age groups.

3. Results

3.1. Item analysis and evaluation of operational definitions

Iterative conceptual and item analyses led to the elimination of 67 of the original 251 items, producing an instrument with 184 items (see Table 3 for number of items in each scale, prior to and after conceptual/item analyses). One item belonging to the High Intensity Pleasure scale and one from the Attentional Shifting scale were eliminated because of the large number of non-respondents (at least 40% in each of the three age groups). Fifty-four items were eliminated from the following scales: Activity Level, Low Intensity Pleasure, Approach, Cuddliness/Affiliation, Perceptual Sensitivity, Duration of Orienting, Sadness, Attentional Shifting, Social Fear, Falling Reactivity, Distress to Limitations, Smiling/Laughter, and Fear, because they did not correlate at least .3 with the scale total, and/or because their removal contributed to the improvement of internal consistency for the scale. In addition, 11 items were eliminated, and two moved, due to conceptual and content considerations. After changes dictated by the item analysis were carried out, the Attention Shifting scale addressed essentially distractibility. The remaining six items for the Attentional Shifting scale were not retained as part of the IBQ-R, because an operational definition for this scale could not be established independent of Duration of Orienting, a previously validated scale.

Separate scales were initially proposed to assess fear and social fear. A number of items were eliminated based on the results of the item analysis, leaving the Fear scale with seven, and the Social Fear scale with nine items. Subsequent reevaluation of operational definitions and individual items led us to consider combining these scales which were correlated ($r(358) = .46$), to form a single dimension addressing fear, similar to the previous Fear scale for the

IBQ. The seven remaining items addressing non-social fearfulness, combined with nine items from the Social Fear scale, formed a coherent, internally consistent scale, with all item-total correlations of at least .3, forming an overall Fear scale.

3.2. *Internal consistency/factor structure of the IBQ-R*

Conceptual and item analyses led to the development of 14 scales, internally consistent across the three age groups (Table 3). Initially, we had considered including different sets of items in scales for infants in different age groups, because this approach was used in the development of the original IBQ. However, in this study, the same sets of items demonstrated reliability across the three age groups (3–6, 6–9, and 9–12 months). In addition, no differences were observed in the patterns of missing data (i.e., items not endorsed by parents because they were considered to be non-applicable) across the three age groups. Thus, a decision was made to retain uniform item sets across the evaluated age groups.

A number of moderate and significant correlations among the IBQ-R scales were observed across the three age groups, as well as for the entire sample (see Table 4). Factor analysis was conducted in order to evaluate higher-order relationships between the different scales of the IBQ-R, following an analytic strategy parallel to the one employed in the development of the CBQ (Rothbart, Ahadi et al., 2001; Rothbart, Chew et al., 2001). Principal Axis extraction was utilized, iterating to communalities, and the extracted factors were subsequently obliquely rotated via the Oblimin algorithm (Norusis, 1994). The three-factor solution derived for the sample of 360 infants is presented in Table 5.

The first factor contained loadings for Approach, Vocal Reactivity, High Intensity Pleasure, Smiling and Laughter, Activity Level, and Perceptual Sensitivity, and appeared to represent a broad dimension of Surgency/Extraversion. The second factor included loadings from Sadness, Distress to Limitations, Fear, and loading negatively, Falling Reactivity. The Fear scale was associated with the smallest loading (0.31) on this factor. The second factor was consistent with a higher-order dimension of Negative Affectivity. The third factor included loadings for Low Intensity Pleasure, Cuddliness/Affiliation, Duration of Orienting, and Soothability, with a secondary loading above 0.3 for Smiling and Laughter. This third factor can be thought of as including regulatory functioning, and was labeled Orienting/Regulation. Low correlations were generally observed among the factors: $r = .16$ for the first and second factors (Surgency/Extraversion and Negative Emotionality), $r = .25$ for the first and third factors (Surgency/Extraversion and Orienting/Regulation), $r = -.30$ for the second and third factors (Negative Emotionality and Orienting/Regulation). Internal consistency of the three broad factors was also evaluated. Cronbach's alpha for Surgency/Extraversion was 0.92, with estimates for the Negative Affectivity and Orienting/Regulation factors both equaling 0.91.

3.3. *Inter-rater reliability*

Inter-rater agreement was evaluated for a subset of participating families ($n = 26$). Reports of primary and secondary caregivers were generally moderately correlated (Table 6). A number of these correlations did not reach statistical significance, however, because of the small sample size. Correlations for all scales but one (Soothability) were above .25, and ranged from .30

Table 4
Correlations among the 14 IBQ-R scales

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Approach	–	53**	52**	38**	24**	43**	–02	13	19**	05	15**	–13	12	17*
Vocal Reactivity	50*	–	51**	63**	23**	30**	–05	–01	17*	14	27**	06	37**	25**
High Pleasure	49**	52**	–	45**	13	31**	–15*	07	10	06	20**	–06	09	17*
Smile/Laughter	30**	61**	46**	–	09	24**	–05	–21**	15*	21**	41**	24**	33**	29**
Activity	48**	34**	30**	18*	–	12	21**	29**	20**	–09	–07	–27**	12	–07
Perceptual Sensitivity	40**	43**	25**	32**	32**	–	13	16*	10	01	18*	–09	15*	14
Sadness	15*	01	–05	–13	21**	01	–	55**	25**	–42**	–16*	–08	–01	–18**
Distress to Limitations	27**	09	05	–15*	33**	13	56**	–	13	–68**	–40**	–25**	–08	–25**
Fear	24**	31**	08	10	16*	30**	27**	39**	–	–02	–05	–05	13	07
Falling Reactivity	–05	07	04	13	–13	–03	–44**	–49**	–12	–	34**	19*	01	32**
Low Pleasure	15*	22**	20**	35**	–01	21**	–13	–37**	–07	20**	–	33**	31**	43**
Cuddliness	–09	–06	–04	11	–33**	–08	–27**	–36**	–11	21**	42**	–	14	31**
Duration of Orienting	12	17*	22**	28**	–04	18*	–01	–13	05	08	44**	12	–	22**
Soothability	28**	41**	24**	42**	21**	23**	–11	–11	18*	18*	42**	21**	27**	–

Note. Values below the diagonal represent correlation coefficients for male infants; correlations for females are presented above the diagonal.

* $p < .05$.

** $p < .01$.

Table 5
IBQ-R factor loadings: three-factor solution ($N = 360$)

IBQ-R scale	Factor 1	Factor 2	Factor 3
Approach	0.74		
Vocal Reactivity	0.74		
High Intensity Pleasure	0.69	−0.25	
Smiling and Laughter	0.55		0.33
Activity Level	0.49		−0.21
Perceptual Sensitivity	0.45		
Sadness		0.79	
Distress to Limitations		0.69	−0.28
Fear		0.31	
Falling Reactivity		−0.56	
Low Intensity Pleasure			0.70
Cuddliness	−0.24		0.56
Duration of Orienting			0.43
Soothability	0.27		0.43

Note: Principle axis factor analysis/oblique rotation; loadings $<.20$ not included.

to .71. Inter-rater agreement for the primary and secondary caregivers was also demonstrated with the three IBQ-R broad factors: Surgency/Extraversion (0.49), Negative Affectivity (0.70), and Orienting/Regulation (0.31).

3.4. Age and gender differences

Age and gender differences in temperament between the three age groups were then examined. Analysis of variance was selected as the analytic strategy because it allowed for an evaluation of age-by-gender interactions. Significant age effects emerged in ANOVAs for

Table 6
Correlations for the IBQ-R subscale scores between the primary and the secondary caregivers ($N = 26$)

Temperament variable	Correlation coefficient
Approach	.49**
Vocal Reactivity	.35
High Intensity Pleasure	.30
Smiling and Laughter	.30
Activity	.63**
Perceptual Sensitivity	.47*
Sadness	.27
Distress to Limitations	.57**
Fear	.75**
Falling Reactivity	.69**
Low Intensity Pleasure	.37
Cuddliness/Affiliation	.38
Duration of Orienting	.19
Soothability	.06

* $p < .05$, all one-tailed tests.

** $p < .01$, all one-tailed tests.

Table 7

Significant age comparisons: Group 1: 3–6 months of age ($N = 120$); Group 2: 6–9 months ($N = 120$), and Group 3: 9–12 months of age ($N = 120$)

Temperament variable	Group 1	Group 2	Group 3	<i>F</i>
Approach				
<i>M</i>	4.49	5.35 ^a	5.52 ^a	44.31**
SD	1.20	0.81	0.69	
Vocal Reactivity				
<i>M</i>	4.33	4.67	5.08	23.25**
SD	0.94	0.80	0.81	
High Intensity Pleasure				
<i>M</i>	5.67	6.03 ^a	6.15 ^a	19.55**
SD	0.79	0.59	0.57	
Activity Level				
<i>M</i>	4.00	4.37 ^a	4.47 ^a	12.08**
SD	0.95	0.73	0.82	
Perceptual Sensitivity				
<i>M</i>	3.64	4.14 ^a	4.21 ^a	11.38**
SD	1.11	1.05	1.01	
Distress to Limitations				
<i>M</i>	3.21	3.56	3.98	26.66**
SD	0.81	0.87	0.88	
Fear				
<i>M</i>	2.26	2.46	2.80	10.49**
SD	1.03	0.97	0.87	
Low Intensity Pleasure				
<i>M</i>	5.32 ^a	5.07 ^{a,b}	4.84 ^b	9.57**
SD	0.85	0.95	0.88	
Cuddliness/Affiliation				
<i>M</i>	5.97	5.72	5.41	23.84**
SD	0.58	0.63	0.76	
Duration of Orienting				
<i>M</i>	3.92	3.60 ^a	3.49 ^a	6.05**
SD	1.04	1.13	0.92	

Note: Degrees of freedom were (1, 357) for all of the comparisons; means sharing common superscripts (^{a,b}) are not significantly different.

** $p < .01$, all two-tailed tests.

Approach, Vocal Reactivity, High Intensity Pleasure, Activity Level, Perceptual Sensitivity, Distress to Limitations, Fear, Low Intensity Pleasure, Cuddliness/Affiliation, and Duration of Orienting (Table 7). Older infants obtained higher scores on Approach, Vocal Reactivity, High Intensity Pleasure, Activity, Perceptual Sensitivity, Distress to Limitations, and Fear, whereas younger infants' scores were higher for Low Intensity Pleasure, Cuddliness/Affiliation, and Duration of Orienting.

Potential gender effects were evaluated with ANOVAs, which indicated significant gender differences for three of the scales, with male infants receiving higher scores on Activity and High Intensity Pleasure, and female infants obtaining higher scores on the Fear scale (Table 8).

Table 8
Significant gender comparisons: males ($N = 179$), females ($N = 181$)

Temperament variable	Males	Females	F
High Intensity Pleasure			
M	6.03	5.86	7.45**
SD	0.69	0.67	
Activity Level			
M	4.36	4.20	4.61*
SD	0.88	0.84	
Fear			
M	2.39	2.62	4.60*
SD	0.93	1.03	

Note: Degrees of freedom were (1, 358) for all of the comparisons.

* $p < .05$, all two-tailed tests.

** $p < .01$, all two-tailed tests.

Interactions between age and gender were evaluated despite the lack of specific hypotheses, but they did not reach statistical significance.

Developmental and gender differences were also evaluated for the three broad temperament factors. Significant age differences emerged for the Surgency/Extraversion [$F(2, 357) = 11.42$; $p < .001$], Negative Affectivity [$F(2, 357) = 9$; $p < .001$], and Orienting/Regulation [$F(2, 357) = 4.79$; $p < .01$] factors. Older infants demonstrated higher levels of Surgency/Extraversion (Group 1: $M = 3.97$, $SD = 0.81$; Group 2: $M = 4.53$, $SD = 0.57$; Group 3: $M = 4.78$; $SD = 0.55$), and Negative Emotionality (Group 1: $M = 2.55$, $SD = 0.67$; Group 2: $M = 2.77$, $SD = 0.63$; Group 3: $M = 3.04$; $SD = 0.63$), across the first year, with significant differences emerging between all three age groups. Parents reported lower levels of Orienting/Regulation for infants in the oldest age group ($M = 4.39$, $SD = 0.59$) in comparison with the younger, but not the middle age group ($M = 4.61$, $SD = 0.63$). No significant gender differences or age-by-gender interactions were found for the three broad IBQ-R factors.

4. Discussion

The IBQ-R was developed to provide a more differentiated measure of infant temperament, consistent with Rothbart's psychobiologically oriented approach to temperament emphasizing both reactive and regulatory capacities (Rothbart, 1981, 1989; Rothbart & Derryberry, 1981; Rothbart & Posner, 1985). This study provided initial support for the reliability and validity of the IBQ-R, and demonstrated a factor structure generally consistent with the broad dimensions reported for older children and adults (Eysenck & Eysenck, 1985; Rothbart, Ahadi, et al., 2000, 2001; Sanson, Smart, Prior, Oberklaid, & Pedlow, 1994; Tellegen, 1985), but with some important differences.

Fourteen of the 16 proposed scales demonstrated satisfactory internal consistency, and were retained following conceptual analyses. Attentional Shifting was eliminated because its internal consistency could not be improved without compromising the conceptual integrity of

the scale. Fear and Social Fear scales were combined into a single indicator due to similarities in operational definitions, a substantial correlation between the scales, and the small number of items remaining in one scale following conceptual and item analyses. Subsequently, 14 internally consistent and conceptually independent scales were retained for further analyses.

The failure to develop distinct Fear and Social Fear scales may reflect a general difficulty in defining the construct of social fear independently in infancy, although Social Fear has been successfully evaluated with toddlers and older children (Goldsmith, 1996; Kochanska, 1991; Rothbart, Ahadi, et al., 2001). Our difficulty in creating an internally consistent and conceptually independent scale of attentional shifting may be developmentally based, related to problems in differentiating more voluntary shifts in attention from those driven by distracting stimuli, which are likely to be captured in our Perceptual Sensitivity scale.

Inter-rater reliability demonstrated moderate agreement between primary and secondary caregivers. A number of these coefficients did not reach statistical significance because of the small sample size ($n = 26$). Nevertheless, only one of the correlations was below .25 (Soothability). Thus, caregivers demonstrated some agreement regarding the infants' temperament attributes assessed using the IBQ-R, despite the fact that the two adults may interact with the infant at different times, provide different elicitations of behavior, and may differentially interpret infant behavior. Discrepancies in caregivers' reports of child temperament have been described in previous research (Goldsmith & Campos, 1990), with mothers' and fathers' individual reports related to their actual, and different, elicitation behaviors. The lowest index of inter-rater reliability was demonstrated for the Soothability scale. This low level of agreement is not surprising, given the variability of the infants' reactions to the soothing attempts of the different caregivers, who likely differ in their use of the soothing techniques, and their effectiveness.

Monomethod discriminant validity received some support in examination of the correlations between the temperament constructs evaluated with the IBQ-R. A number of intercorrelations between the scales were in the low to moderate range. Some moderately high coefficients were also found, indicating that the scales were not completely independent, yet not redundant in terms of the characteristics evaluated. Non-independence was anticipated, in that some of these scales were designed to evaluate systems that were expected to be related or to modulate each other (Rothbart et al., 1994). The IBQ-R scales also demonstrated patterns of intercorrelations consistent with three overarching dimensions of Surgency/Extraversion, Negative Affectivity, and characteristics associated with regulatory functions. Two factors in this three-factor structure are consistent with considerable theoretical and empirical evidence (Putnam et al., 2001; Rothbart & Bates, 1998; Tellegen, 1985). However, there are conceptual differences between the factor of Effortful Control and the infant Regulatory/Orienting factor.

Considerable similarity was noted between the Negative Affectivity factor extracted in this investigation, and a Negative Affectivity factor found in studies of temperament later in childhood (Rothbart, Ahadi, et al., 2001; Sanson et al., 1994). The differences between the finding for the IBQ-R and the CBQ involved minor changes in factorial composition. Unlike the CBQ, the IBQ-R did not include a Discomfort scale. The IBQ-R Negative Affectivity factor was associated with a negative loading for Falling Reactivity, whereas for the CBQ, the Negative Affectivity factor was negatively associated with Soothability. Falling Reactivity

can be thought of as an infant counterpart of the children's Soothability scale, focusing on the infant's own abilities to change her state. The Soothability scale in the IBQ-R, however, addresses the parent's capacity to lower levels of infant distress/arousal. Negative Affectivity also appears consistent with the Big Five personality factor of Neuroticism (Digman & Inouye, 1986; Goldberg, 1993; Rothbart, Ahadi, et al., 2000).

The Positive Emotionality/Surgency factor in the IBQ-R was also generally consistent with the Extraversion/Surgency factor reported for the CBQ. Again, some differences in factorial composition were observed, with the CBQ factor including the Impulsivity scale, a scale that is not appropriate for the period of infancy. Instead, the conceptually similar Approach scale on the IBQ-R loads on the Positive Emotionality/Surgency factor. The IBQ-R also included Vocal Reactivity, not appropriate for older children. The CBQ factor included a negatively loading Shyness scale, which was not available for the IBQ-R, although we had attempted to develop a scale addressing social fear. Overall, the Positive Emotionality/Surgency factors demonstrated in research with children are consistent with the personality dimension labeled as Extraversion (Eysenck & Eysenck, 1985; Rothbart, Ahadi, et al., 2000; Rothbart & Bates, 1998; Tellegen, 1985).

The third, more regulatory oriented factor emerging in the present study contained scales associated with attentional Duration of Orienting, positive affect, and soothing. The IBQ-R Orienting/Regulation factor also included the Low Intensity Pleasure scale, and was associated with a significant secondary factor loading for Smiling/Laughter. The association between regulatory functions and low intensity pleasure observed for infants in this study is consistent with the factor analytic findings for older children, based on the CBQ (Rothbart, Ahadi, et al., 2001), where Low Intensity Pleasure consistently loaded on the Effortful Control factor, along with attentional control scales, such as Inhibitory Control and Attention Shifting.

There are considerable differences in the characteristics related to self-regulation at the younger and older ages, however, reflected in the scales we were able to generate for the IBQ-R. For example, an Inhibitory Control scale was not appropriate for the IBQ-R because infants are not yet able to engage in delay following instructions. Perceptual Sensitivity in infancy also loaded on the Positive Emotionality/Surgency factor, rather than the Orienting/Regulation factor. Later in childhood, Perceptual Sensitivity loads on the Effortful Control factor, which has in turn been empirically linked to the Control/Constraint/Conscientiousness factor found in personality studies (Rothbart, Ahadi, et al., 2001).

These differences in the content of regulatory factors appear to reflect significant developmental shifts in self-regulation and related processes, observed between infancy and the preschool period (Posner & Rothbart, 1998). For example, higher-order attentional skills addressed in the Inhibitory Control scale have been linked to the development of the executive attention system, which first appears toward the end of the first year of life, and continues to develop during childhood (Posner & Rothbart, 1998; Rothbart & Bates, 1998). The immaturity of inhibitory and attentional control, and related capacities in infancy, often necessitates external (i.e., caregiver) involvement in regulatory functions. Caregiver involvement in infants' regulatory functions is likely reflected in the IBQ-R scales of Soothability and Cuddliness, which also load on the Orienting/Regulatory factor in infancy. As children develop self-regulatory skills, adult interventions aimed at regulating behavioral and emotional reactions become less prominent. Longitudinal investigations relating early Orienting/Regulation to later Effortful

Control will be of interest in determining whether early emotion regulation is related to later attention regulation, as [Posner and Rothbart \(1998\)](#) have suggested.

Our evaluation of age differences produced a number of significant results that are consistent with previous research. Higher scores on Approach, Activity Level, Distress to Limitations, and Fear for older infants found in the present study have been previously reported ([Carranza et al., 2000](#); [Rothbart, 1988, 1994](#)). Higher scores on Activity Level and Approach scales for older children may stem, at least in part, from increased capacities for locomotion.

Increases in Distress to Limitations may also be related to the development of locomotion, as well as to increased parental limitations on exploratory activity. Increases in Distress to Limitations might also be related to emerging cognitive skills, including goal directed thinking and long-term memory, allowing goals to be kept in mind, and creating greater potential for frustration. Infants developing these capacities are more likely to show distress when unable to grasp desired objects, or when a caregiver removes a desired object ([Carranza et al., 2000](#)). Significant increases in Fear between 3 and 12 months are consistent with prior research on development of inhibition of approach toward novel and/or intense stimuli in the first year of life ([Rothbart, 1988, 1994](#)). Higher scores on the Vocal Reactivity scale for older infants may be related to increases in communicative skill in the first year of life ([Worobey, 1989](#)). In the future, it will be interesting to examine whether this early vocal reactivity is predictive of aspects of later language development. Consistent with findings for the fine-grained subscales described above, significant age differences emerged for the broad IBQ-R factors. Specifically, older infants demonstrated higher levels of Surgency/Extraversion and Negative Emotionality across the first year of life.

Observed lower levels of Duration of Orienting for older infants may reflect of maturational changes in the posterior attention network that allow the infant to disengage from visual stimuli more efficiently ([Ruff & Rothbart, 1996](#)). This decrease in Duration of Orienting is also consistent with the literature suggesting that older infants habituate to objects more rapidly ([Colombo, 2001](#)). More rapid habituation may also be responsible for higher levels of Perceptual Sensitivity demonstrated for older infants in this study ([Ruff & Rothbart, 1996](#)). Perceptual Sensitivity has been linked with executive attention and regulatory functioning in the Effortful Control factor extracted for school-age children. However, as noted above, the Perceptual Sensitivity scale was not part of the regulatory capacity factor in infancy; rather, it was associated with Surgency/Extraversion.

Several additional age differences emerged in this study that had not been predicted. Older infants were described as demonstrating higher levels of High Intensity Pleasure, and lower levels of Low Intensity Pleasure and Cuddliness. These differences could be related to increases in mobility that make children less likely to enjoy being held closely by a caregiver, or being involved in quiet activities, without moving around. Preferences for high intensity stimulation may also be related to habituation of earlier preferences for low intensity stimuli. These age differences should be interpreted with caution, however, because they were not predicted prior to data analysis. Observed decreases in the levels of Duration of Orienting, Cuddliness, and Low Intensity Pleasure over the first year of life were likely contributors to the finding of lower Orienting/Regulation for older infants.

Overall, age differences demonstrated in this study provide further support for overall changes in the expression of temperament during infancy ([Riese, 1987](#); [Rothbart, 1989](#)). The

cross-sectional nature of these data does not allow for an evaluation of stability and instability in temperamental characteristics; nevertheless, several developmental trends emerged in this study.

Analyses addressing gender differences yielded far fewer significant results, with males receiving higher scores, as predicted, on Activity Level, and females on Fear. The gender difference in activity is consistent with meta-analysis of activity level in infancy (Campbell & Eaton, 1999). Higher scores on the Fear scale for females are also consistent with previous research, with a greater tendency toward withdrawal for females as assessed by parent report on Approach-Withdrawal scales across culturally diverse samples (Carey & McDevitt, 1978; Hsu et al., 1981; Maziade et al., 1984). Male infants were also described as exhibiting higher levels of High Intensity Pleasure, indicating greater involvement in activities characterized by high stimulus intensity, rate, complexity, novelty, and/or incongruity. The latter finding, however, was not predicted, and thus should be interpreted with caution. The influence of socialization on these early appearing gender differences cannot be ruled out, although these differences may be indicative of more biologically based sex differences.

4.1. Summary and future directions

This initial evaluation of the IBQ-R psychometric properties provides evidence of the instrument's reliability and validity. The IBQ-R is representative of a fine-grained, psychobiologically oriented approach to the measurement of temperament, allowing researchers to investigate interactions between different domains of temperament, as well as their relationships with environmental influences (e.g., parenting). The IBQ-R provides an efficient means for evaluating a wide range of temperament variables because it is completed by parents, who have a unique opportunity to observe their infants in a variety of situations. Future research will need to demonstrate whether or not the IBQ-R is as successful at overcoming limitations generally associated with parent report as its predecessor, the original IBQ, has been. The IBQ-R could also be adapted for clinical applications, addressing another important goal of temperament assessment (Bates, 1989). Clinical adaptations to date have included the informal use of questionnaires or observations in clinical diagnosis and treatment, along with encouragement of parents' observing their children's behavior patterns more closely (Parker & Zuckerman, 1990).

The scope of this study did not allow for an evaluation of all important psychometric properties of this instrument. Specifically, establishing convergent and predictive validity of the IBQ-R will require further research. It will be especially crucial to demonstrate convergence between the IBQ-R and other established measures of temperament, as well as indicators of temperament based on standardized laboratory assessments. Cross-sectional findings reported in this study should be replicated with a longitudinal design. Future studies would need to address developmental trajectories of the different temperament variables, given the preliminary findings of age differences reported here. A replication of the factor analytic findings presented in this study should also be undertaken in the future with a larger and a more diverse sample, allowing for examination of possible developmental changes in the structure of temperament within the first year of life.

Acknowledgments

The authors gratefully acknowledge the contributions of student research assistants, especially Jesse Chasman and Jennifer Greenberg. We are also extremely grateful to Sam Putnam, University of Oregon, and Carmen Gonzalez, University of Murcia, for their helpful comments on previous versions of this paper.

References

- Ahadi, S. A., Rothbart, M. K., & Ye, R. (1993). Children's temperament in the US and China: Similarities and differences. *European Journal of Personality*, *7*, 359–378.
- Bates, J. E. (1987). Temperament in infancy. In J. D. Osofsky (Ed.), *Handbook of infant development*. New York: Wiley.
- Bates, J. E. (1989). Concepts and measures of temperament. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 3–26). New York: Wiley.
- Bell, R. Q. (1974). Contributions of human infants to caregiving and social interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 1–20). New York: Wiley.
- Belsky, J. (1996). Parent, infant, and social-contextual antecedents of father–son attachment security. *Developmental Psychology*, *32*, 905–913.
- Bridges, L. J., Palmer, S. A., Morales, M., Hurtado, M., & Tsai, D. (1993). Agreement between affectively based observational and parent-report measures of temperament at infant age 6 months. *Infant Behavior and Development*, *16*, 501–506.
- Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.
- Calkins, S. D., Fox, N. A., & Marshall, T. R. (1996). Behavioral and physiological antecedents of inhibited and uninhibited behavior. *Child Development*, *67*, 523–540.
- Campbell, D. W., & Eaton, W. O. (1999). Sex differences in the activity level of infants. *Infant and Child Development*, *8*, 1–17.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.
- Carey, W. B., & McDevitt, S. C. (1978). Revisions of the infant temperament questionnaire. *Pediatrics*, *61*, 735–739.
- Carlson, S. M. (1997). *Individual differences in inhibitory control and children's theory of mind*. Unpublished doctoral dissertation, University of Oregon, Eugene.
- Carranza, J. A., Perez-Lopez, J., Gonzalez, C., & Martinez-Fuentes, M. T. (2000). A longitudinal study of temperament in infancy: Stability and convergence of measures. *European Journal of Personality*, *14*, 21–37.
- Clark, R., Hyde, J. S., Essex, M. J., & Klein, M. H. (1997). Length of maternity leave and quality of mother–infant interactions. *Child Development*, *68*, 364–383.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367.
- Crockenberg, S., & Acredolo, C. (1983). Infant temperament ratings: A function of infants of mothers, or both? *Infant Behavior and Development*, *6*, 61–72.
- Derryberry, D., & Rothbart, M. K. (1988). Arousal, affect, and attention as components of temperament. *Journal of Personality and Social Psychology*, *55*, 958–966.
- Diamond, S. (1957). *Personality and temperament*. New York: Harper.
- Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, *50*(1), 116–123.
- Escalona, S. K. (1968). *The roots of individuality: Normal patterns of development in infancy*. Chicago: Aldine.
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum.
- Fagen, J. W., Ohr, P. S., Singer, J. M., & Fleckenstein, L. K. (1987). Infant temperament and performance on the Bayley scales of infant development at 4, 8, and 12 months of age. *Infant Behavior and Development*, *10*(4), 505–512.

- Gerardi-Caulton, G. (in press). *Measuring executive attention in the third year of life: A new task and its relation to the development of self-regulation.*
- 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., & Campos, J. J. (1990). The structure of temperamental fear and pleasure in infants: A psychometric perspective. *Child Development*, *61*, 1944–1964.
- Goldsmith, H. H., & Riley, J. (1994). *The Laboratory Temperament Assessment Battery: Prelocomotor Version*, V.0.5. Unpublished manuscript.
- Goldsmith, H. H., & Rothbart, M. K. (1991). Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In A. Angleitner & J. Strelau (Eds.), *Explorations in temperament: International perspectives on theory and measurement*. New York: Plenum.
- Goldsmith, H. H., & Rothbart, M. K. (1996). *The Laboratory Temperament Assessment Battery. Prelocomotor Version*, V.3.00. Unpublished manuscript.
- Hsu, C., Soong, W., Stigler, J. W., Hong, C., & Liang, C. (1981). The temperamental characteristics of Chinese babies. *Child Development*, *52*, 1337–1340.
- Huffman, L. C., Bryan, Y. E., del Carmen, R., Pedersen, F. A., Doussard-Roosevelt, J. A., & Porges, S. W. (1998). Infant temperament and cardiac vagal tone: Assessments at twelve weeks of age. *Child Development*, *69*, 624–635.
- Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning anticipatory looking and disengaging. *Journal of Cognitive Neuroscience*, *3*, 335–344.
- Kagan, J. (1994). *Galen's prophecy: Temperament in human nature*. New York: Basic Books.
- Kagan, J. (1998). Biology and the child. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (5th ed.). New York: Wiley.
- Kagan, J., Reznick, J. S., & Snidman, N. (1988). Biological bases of childhood shyness. *Science*, *240*, 167–171.
- Kochanska, G. (1991). Patterns of inhibition to the unfamiliar in children of normal and affectively ill mothers. *Child Development*, *62*, 250–263.
- Kochanska, G. (1993). Toward a synthesis of parental socialization and child temperament in early development of conscience. *Child Development*, *64*, 325–347.
- Kochanska, G., Coy, K. C., Tjebkes, T. L., & Husarek, S. J. (1998). Individual differences in emotionality in infancy. *Child Development*, *64*, 375–390.
- Kochanska, G., DeVet, K., Goldman, M., Murray, K., & Putnam, S. P. (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., & Vandegest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development*, *67*, 490–507.
- Martin, R. P., Wisenbaker, J., Baker, J., & Huttunen, M. O. (1997). Gender differences in temperament at six months and five years. *Infant Behavior and Development*, *20*, 339–347.
- Maziade, M., Boudreault, M., Thivierge, J., Caperaa, P., & Cote, R. (1984). Infant temperament: SES and gender differences and reliability of measurement in a large Quebec sample. *Merrill-Palmer Quarterly*, *30*, 213–226.
- Norusis, M. J. (1994). *SPSS professional statistics 6.1*. Chicago: SPSS, Inc.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
- Parker, S. J., & Zuckerman, B. S. (1990). Therapeutic aspects of the assessment process. In S. J. Meisels & J. P. Shonkoff (Eds.), *Handbook of early childhood intervention*. New York: Cambridge University Press.
- Posner, M. I., & Rothbart, M. K. (1991). Attention mechanisms and conscious experience. In M. Rugg & A. D. Milner (Eds.), *The neuropsychology of consciousness* (pp. 91–112). London: Academic Press.
- Posner, M. I., & Rothbart, M. K. (1998). Attention, self-regulation and consciousness. *Philosophical Transactions of the Royal Society of London B*, *353*, 1915–1927.
- Putnam, S. P., Ellis, L. K., & Rothbart, M. K. (2001). The structure of temperament from infancy through adolescence. In A. Elias & A. Angleitner (Eds.), *Advances in research on temperament* (pp. 165–182). Germany: Pabst Science Publishers.

- Reznick, J. S., Gibbons, J. L., Johnson, M. O., & McDonough, P. M. (1989). Behavioral inhibition in a normative sample. In J. S. Reznick (Ed.), *Perspectives on behavioral inhibition. The John D. and Catherine T. MacArthur Foundation series on mental health and development* (pp. 25–49). Chicago, IL: The University of Chicago Press.
- Riese, M. L. (1987). Temperament stability between the neonatal period and 24 months. *Developmental Psychology*, 23, 216–222.
- Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, 52, 569–578.
- Rothbart, M. K. (1986). Longitudinal observation of infant temperament. *Developmental Psychology*, 22, 356–365.
- Rothbart, M. K. (1988). Temperament and the development of the 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. 187–248). New York: Wiley.
- Rothbart, M. K. (1994). Emotional development: Changes in reactivity and self-regulation. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 369–372). New York: Oxford University Press.
- 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., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development*, 72(5), 1394–1408.
- Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3, social, emotional, and personality development* (5th ed., pp. 105–176). New York: Wiley.
- Rothbart, M. K., Chew, K. H., & Gartstein, M. A. (2001). Assessment of temperament in early development. In L. Singer & P. S. Zeskind (Eds.), *Biobehavioral assessment of the infant* (pp. 190–208). New York: Guilford.
- 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 over the life span* (pp. 85–119). Hillsdale, NJ: Erlbaum.
- Rothbart, M. K., & Goldsmith, H. H. (1985). Three approaches to the study of infant temperament. *Developmental Review*, 5, 237–260.
- 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.
- 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.
- Schaffer, H. R., & Emerson, P. E. (1964a). Patterns of response to physical contact in early human development. *Journal of Child Psychology and Psychiatry*, 5, 1–13.
- Schaffer, H. R., & Emerson, P. E. (1964b). The development of social attachment in infancy. *Monographs of the Society for Research in Child Development*, 94, 1–77.
- Schaughency, E. A., & Fagot, B. I. (1993). The prediction of adjustment at age 7 from activity level at age 5. *Journal of Abnormal Child Psychology*, 21, 29–49.
- Seifer, R., Schiller, M., Sameroff, A. J., Resnick, S., & Riordan, K. (1996). Attachment, maternal sensitivity, and temperament during the first year of life. *Developmental Psychology*, 32, 3–11.
- Shirley, M. M. (1933). *The first two years: A study of 2.5 babies*. Minneapolis: University of Minnesota Press.
- Stevens, G., & Featherman, D. L. (1981). A revised socioeconomic index of occupational status. *Social Science Research*, 10, 365–395.
- Stevenson, J., Thompson, M. J. J., & Sonuga-Barke, E. (1996). Mental health of preschool children and their mothers in a mixed urban/rural population. *British Journal of Psychiatry*, 168, 26–32.

- Strelau, J. (1983). *Temperament personality activity*. New York: Academic Press.
- 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., & Birch, H. G. (1968). *Temperament and behavior disorders in children*. New York: New York University Press.
- Thomas, A., Chess, S., Birch, H. G., Hertzog, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood*. New York: New York University Press.
- Ventura, J. N., & Stevenson, M. B. (1986). Relations of mothers' and fathers' reports of infant temperament parents' psychological functioning and family characteristics. *Merrill-Palmer Quarterly*, 32, 275–289.
- Worobey, J. (1986). Convergence among assessments of temperament in the first month. *Child Development*, 57, 47–55.
- Worobey, J. (1989). Mother–infant interaction: Protocommunication in the developing dyad. In J. F. Nussbaum (Ed.), *Life-span communication* (pp. 7–25). Hillsdale, NJ: Erlbaum.
- Worobey, J. (1997). Convergence between temperament ratings in infancy. *Developmental and Behavioral Pediatrics*, 18, 260–263.
- Worobey, J., & Blajda, V. M. (1989). Temperament ratings at 2 weeks, 2 months, and 1 year: Differential stability of activity and emotionality. *Developmental Psychology*, 25, 257–263.
- Zuckerman, M. (1990). The psychophysiology of sensation seeking. *Journal of Personality*, 58, 313–345.