



Published in final edited form as:

Infant Behav Dev. 2008 April ; 31(2): 227–238.

FEAR AND POSITIVE AFFECTIVITY IN INFANCY: CONVERGENCE/DISCREPANCY BETWEEN PARENT-REPORT AND LABORATORY-BASED INDICATORS

Maria A. Gartstein and Julia Marmion

Abstract

This study examined convergence between indicators of infant temperament derived via parent-report and those obtained in the context of structured laboratory observations. Discrepancies between scores resulting from these methodological approaches were examined in an attempt to explain these differences by considering multiple reporter (i.e., parent) characteristics. Convergence between the two sources of information was hypothesized; however, discrepancies were also expected. This study was aimed at examining whether increased maternal depression and low parenting self-efficacy were related to higher levels of infant fear and decreased positive affectivity, as reported by mothers, relative to the scores derived from the laboratory procedure. Results indicated that the fear scores based on parent-report and structured observations, respectively, were significantly correlated; however, the correlation for smiling and laughter scores did not reach statistical significance. Furthermore, parents higher in negative affect reported a higher level of fear for their infants, relative to the results of the laboratory observation.

Introduction

Although the construct of temperament has a long-standing history, the study of temperament in children, focusing on its developmental course, represents a relatively recent endeavor. Multiple approaches to understanding temperament exist, with some variability in the views of what temperament truly represents, and what traits constitute temperament. Rothbart and Derryberry (1981) define temperament as “constitutional differences in reactivity and self-regulation” (p. 37). Their view encompasses the behavioral, as well as the biological elements of temperament (Rothbart & Derryberry, 1981). Reactivity, in this definition, is reflected by an individual’s autonomic nervous system reactions to modifications in the environment. Self-regulation is defined as the “processes functioning to modulate this reactivity, e.g., attentional and behavioral patterns of approach and avoidance” (p. 37). The notion of “constitutional basis” of temperament refers to the relatively enduring biological make-up of the individual, influenced by heredity, maturation, and experience (Rothbart & Derryberry, 1981).

To assess the different aspects of temperament in infancy, several methods have been widely utilized, including parent-report questionnaires and laboratory observations. Parent-report is beneficial because parents are able to observe their children in a variety of situations and different times of the day. Hence, parents can provide a vast amount of information regarding their child’s behavior. Laboratory observations, on the other hand, allow researchers to control

Julia Marmion, MS, 250 E. 5th Street, Apt. 303, St. Paul, MN 55101, (651) 290-2999, juliamarmion@hotmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

the environment, allowing them to compare children's behavior in similar situations, and considerably reducing potential external influences on child behavior (Rothbart & Bates, 2006).

However, each of these approaches is associated with certain limitations. Kagan (1998), for example, criticized the use of parent-report, noting that parents tend to form a consistent disposition toward their child, although parents' prior experiences with the child may not have been that consistent. Another limiting factor of parent-report, according to Kagan (1998), is that parents unwittingly make comparisons when judging characteristics of their child, which are subject to the influence of prior experiences not related to their child's temperament. Fortunately, a number of concerns raised by Kagan (1998) can be 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). Rothbart (1981) also cautioned that parent-report measures of temperament were not independent of the home environment of the child. Although methodological concerns related to potential sources of error have been often raised in relation to caregiver report questionnaires (e.g., Kagan, 1994, 1998), these also extend to the laboratory observation measures (Goldsmith & Rothbart, 1991; Rothbart & Bates, 2006). Sources of observational error include those related to the characteristics of the rater, effects of the measure on child behaviors, and interactions between rater characteristics and child behavior (Rothbart & Goldsmith, 1985).

Laboratory observations could also be problematic because only a limited set of behaviors may be evoked in this artificial setting, primarily due to the necessarily short time frame of the assessment. Additionally, carryover effects represent a significant threat when repeated testing is required (Rothbart & Bates, 2006). Carryover effects could be reduced by including alternative methods for assessing temperament constructs, and combining scores derived from these alternative techniques in multi-method composites. The novelty of the new environment could also affect the infant's behavior because the novelty may make the child wary/fearful, making the behavior exhibited in the lab not representative of the behavior in the home environment (Rothbart & Goldsmith, 1985). Thus, the child could be less likely to display positive affect in the laboratory relative to more familiar settings (e.g., home). In addition, thus far no single laboratory observation protocol has been established as the "gold standard" on the basis of research addressing reliability and validity of this methodological approach. Rather, multiple structured observation procedures have been developed, and their use has varied widely across different laboratories (Rothbart & Bates, 2006).

Upon considering these methodological issues, it has become apparent that neither parental report, nor laboratory observations alone represent infallible assessment tools for the evaluation of infant temperament. However, a more comprehensive and representative picture of the infant's behavior can be obtained by utilizing both methodological approaches in tandem, having demonstrated their convergence (Rothbart & Bates, 2006). This construct-building approach has enabled researchers to study convergence between different sources of information, presumably reflecting temperament "true scores." Inclusion of the structured laboratory observations and the parent-report measures has also enabled researchers to evaluate temperament from a multitrait-multimethod matrix perspective (Campbell & Fiske, 1959). This framework could be utilized to ensure that consistent indicators of each trait are provided by the different methodological approaches, and/or discern that a given temperament score is sufficiently different from the other attribute scores. More recently, a number of important investigations addressing temperament have relied on multi-method temperament assessments (e.g., Kochanska, Murray, & Coy, 1997; Goldsmith & Campos, 1990), having demonstrated convergence between caregiver report and laboratory observations. For example, Kochanska et al. (1997) found significant correlations ($r = .35$) between parent report and laboratory

observations of inhibitory control, a regulation-related domain of temperament. However, relatively little attention has been devoted to understanding discrepancies that regularly arise between indices reflecting parents' perceptions and laboratory-based measures.

Convergence between these two sets of measures could be at least in part explained by parental depression. Results of multiple studies have indicated that higher levels of depressive symptoms may lead to parental over-reporting of child behavior problems and subsequent increase in disagreement with other sources of information (Fergusson, Lynskey, & Horwood, 1993; Richters & Pellegrini, 1989). Additionally, research has shown that maternal perception of the child as temperamentally "difficult" was related to maternal depression (Whiffen, 1990), with more depressed mothers reporting greater levels of "difficult" temperament. Whiffen (1990) also found that parental disagreement concerning their child's temperament increased when the mother reported higher levels of depression. Other researchers found that both depressed mothers as well as their partners reported their child as having a more "difficult" temperament (Edhborg, Seimyr, Lundh, & Widström, 2000). Neither of these studies used measures of child temperament that did not rely on parental report, which is noteworthy because it is not possible to determine whether the children of depressed mothers truly possess more "difficult" temperament characteristics, or whether there is maternal bias associated with elevated levels of depression, leading to over-reporting of the challenging child temperament attributes. The author of a meta-analysis of 17 studies examining the relationship between infant temperament and postpartum depression reported a moderate relationship between these two factors (Beck, 1996). However, Beck (1996) also warned readers that the results should be interpreted with caution given that most studies only used maternal reports of infant temperament, which is problematic because these reports may have been distorted due to the mothers' depressive symptoms, and additional sources of information (e.g., observations) are necessary in order to distinguish potential "distortion" effects of parental depression (Beck, 1996).

To our knowledge, Leerkes and Crockenberg (2003) conducted the only study examining the influence of maternal depression on the divergence between laboratory-based and parent-report measures of infant temperament. Specifically, the effect of depressive symptoms on caregiver-reported and observed infant negative emotionality was studied. The participating mothers completed a depression rating scale, and infant negative emotionality was measured utilizing two subscales of the Infant Behavior Questionnaire (IBQ; Rothbart, 1981): Distress to Limitations and the Distress and Latency to Approach Sudden or Novel Stimuli. Infant negative emotionality was also measured in the laboratory via novelty and limitation tasks (Leerkes & Crockenberg, 2003). It was found that the report of highly depressed mothers was less concordant with laboratory-based observations of distress to novelty, compared to the information provided of less depressed caregivers.

The study by Leerkes and Crockenberg (2003) represents an important step in explaining factors contributing to the discrepancy between different sources of information regarding infant temperament, but a number of questions remain. First, parent and parent-child interaction characteristics other than those addressed in the study (e.g., parent temperament and self-efficacy) may contribute to these disagreements. Second, discrepancies between parent-report and observation-based indicators related to infant temperament characteristics associated with positive emotionality represent another important area for research to examine.

For example, it is of interest to study whether maternal temperament, particularly negative emotionality/neuroticism, is related to over- or under-reporting of infant temperament characteristics, relative to observational indicators, given that higher maternal prenatal depression has been linked to increased incongruence between caregiver-report and observational measures of temperament (Leerkes & Crockenberg, 2003). Maternal negative

emotionality was related to child problem behaviors, measured through observational techniques and maternal report (Kochanska, Clark, & Goldman, 1997). A maternal negative emotionality composite, constructed across multiple self-report indices (e.g., neuroticism, physiological sensitivity to stress, etc.) was negatively related with mother-reported child outcomes, such as attachment security and internalization of family rules, and positively correlated with mother-reported behavioral problems (Kochanska et al., 1997). The contribution of this negative affectivity personality constellation to discrepancies between parent-report and observation-based indicators of child problem behaviors was not examined in the context of this study, but is of interest, given the reported associations with increased problem behaviors and decreased attachment security and internalization.

Maternal self-efficacy represents another factor that may influence the discrepancy between parental report and laboratory measures of infant temperament. Previous research has found that maternal self-efficacy, assessed through a self-report measure, was negatively related to infant difficulty, also evaluated through maternal report (Teti & Gelfand, 1991). Unfortunately, due to this study's reliance on maternal-report it is unclear whether low care-giver self-efficacy was indeed associated with more difficult infant temperament or whether mothers low in self-efficacy tended to perceive their infants as more difficult, even though the children may not be observed to behave in this manner by others. It is conceivable that mothers of children with more challenging temperament presentations (e.g., irritable, easily frustrated, etc.) do not perceive their self-efficacy to be at the same levels as the caregivers of less demanding infants. On the other hand, maternal perceived self-competence may act in a manner similar to depression, leading to inaccurate perceptions of child behavior that accentuate the negative aspects of his/her presentation.

The purpose of this study was to explore potential explanations for the discrepant findings based on parent-report and structured laboratory observations of infant temperament. Previous research has demonstrated low to moderate agreement between parent-report and other sources of information addressing infant temperament (e.g. Carter, Little, Briggs-Gowan, & Kogan, 1999; Field & Greenberg, 1982), making it important to address factors contributing to discrepancies, along with formulating composite scores based on convergent information provided by the different sources. Based on prior research, it was hypothesized that infant temperament data collected in the context of structured laboratory observations and information based on parent-report questionnaires would converge to a significant extent. That is, significant correlations were expected among parent-report derived measures of infant temperament and laboratory observation indicators. Second, it was anticipated that maternal depression would be related to increased reports of infant fearfulness, compared to fearfulness measured in the laboratory.

Additional analyses were aimed at exploring factors potentially contributing to the discrepancies between laboratory observation and parent-report based temperament score. Specifically, a question of whether maternal depression would be associated with decreased reporting of positive affectivity (i.e., Smiling/Laughter), thus leading to another area of discrepancy between the two methodological approaches in addition to the already identified negative emotionality, was addressed. The relationship between perceived maternal efficacy and potential discrepancies between parent-report and laboratory observations of infant temperament was also examined, in order to determine if lower levels of parenting efficacy would be related to increased reports of Fear and decreased reporting of Smiling/Laughter, as compared to laboratory measurements. The relationship between maternal temperament and discrepancies observed between the two sources of information regarding infant temperament was also explored. Specifically, the contributions of maternal positive and negative affectivity to discrepancies between parental reports of infant Fear and Smiling/Laughter respectively, relative to observation-based indicators, were examined.

Methods

Participants

The sample consists of 68 families with infants (19 infants were 6 months, 25 were 9 months, and 24 were 12 months of age), who completed the laboratory assessment, as well as the questionnaires. This group of families was recruited to ensure an approximately equivalent age and gender distribution (36 of the infants were male and 32 were female). Eight additional families completed the laboratory assessment only due to time constraints (i.e., they did not complete the questionnaire portion of the assessment prior to the laboratory observations, citing a lack of time), and thus are not included in this study. The laboratory assessment scores for infants whose caregivers completed the full assessment did not differ from the scores of those whose families did not (i.e., families who did not complete the questionnaire portion of the study). Birth announcements from hospitals published in the local newspapers were used to recruit families from the San Francisco Bay Area, who were called about two weeks before their infants were eligible to participate in the study (i.e., were within two weeks of being 6, 9, or 12 months of age). The majority of participating parents (94.1%) were married, 1.5% were single, 1.5% were remarried, and 1.5% of the participants were cohabitating. This sample was primarily Caucasian/European American (82.4%), with 8.8% of Asian/American participants, 4.4% of Hispanic/Latino, 2.9% of Filipino, and 1.5% of African American families. All participants finished at least high school with 5.9% of participants finishing high school without attaining any postsecondary education, 17.7% attaining some postsecondary education, 41.2% attained four years of postsecondary education, and 33.9% attaining graduate or professional training. One participant did not disclose educational attainment. The mean family socioeconomic status according to the Revised Duncan Sociometric Index (TSEI2; Stevens & Featherman, 1981) was 37.29 (SD= 25.41), with the most prevalent occupations being related to management in medicine and health. Mothers completed 97.1% of the questionnaires, whereas 2.9% were completed by fathers.

Measures

The Infant Behavior Questionnaire-Revised (IBQ-R; Gartstein & Rothbart, 2003)—The IBQ-R, which is based on the IBQ (Rothbart, 1981), represents a rationally derived, fine-grained parent-report assessment tool, grounded in the definition of temperament proposed by Rothbart & Derryberry (1981), work with the Child Behavior Questionnaire (CBQ; Rothbart, Ahadi, & Hershey, 1994), comparative studies, as well as other developmental research that had identified significant temperament dimensions and associated behavioral tendencies. The IBQ-R consists of 14 scales, which in turn compose three overarching factors: (1) Positive Affectivity/Surgency - Smiling and Laughter, High Intensity Pleasure, Activity Level, Approach, Perceptual Sensitivity, and Vocal Reactivity; (2) Negative Emotionality - Fear, Distress to Limitations, Sadness, and Falling Reactivity (negatively loading); (3) Orienting/Regulatory Capacity - Duration of Orienting, Soothability, Cuddliness/Affiliation, and Low Intensity Pleasure (Gartstein & Rothbart, 2003). Reliability and validity of this parent-report instrument have been documented, with Cronbach's alphas ranging from .77 to .96 (Gartstein & Rothbart, 2003; Gartstein, Slobodskaya & Kinsht, 2003). Fear as well as Smiling and Laughter scales were utilized for the purposes of this study, with the following Chronach's alphas computed for the present sample: IBQ Fear = .81; IBQ Smiling and Laughter = .80.

The Temperament Laboratory Assessment (TLA; Gonzalez, Gartstein, Carranza, & Rothbart, 2003)—This laboratory protocol represents an adaptation of the widely used Lab-TAB episodes (Goldsmith, & Rothbart, 1996), designed to elicit reactive and regulatory aspects of temperament consistent with the IBQ-R, in order to maximize the potential for convergence across these two sources of temperament-related information. The TLA consists of 10 different episodes, which take about 1 hour to administer: warm-up, play with toys, toy

retraction, embrace with examiner, visual perceptual sensitivity, unstructured play with the parent, separation, peek-a-boo, auditory perceptual sensitivity, presentation of masks. This protocol was developed for children between 6 and 12 months of age. The Lab-TAB (Goldsmith, & Rothbart, 1996) that provides the bases for the TLA has been consistently described as reliable and valid, with inter-rater agreement ranging from 88% to 99%. Analyses of the TLA have also provided satisfactory inter-rater reliability estimates, with r 's ranging from .64 to 1.00. These correlation coefficients represent reliability estimates computed on the basis of the ratings provided by the authors, responsible for providing codes in the episodes included in this study. These coders took part in comprehensive training, aimed at achieving an adequate agreement, operationalized as Pearson Product Moment correlation coefficients of at least .60 associated with each pair-wise association reflecting agreement between raters. Scores derived from the TLA provided indices of threshold, latency, intensity, time-to-peak intensity, and recovery time for reactions that involved arousability of affect, motor activity, and related responses. Fear and smiling/laughter eliciting episodes of the TLA were utilized in the context of this study. Both the warm-up, when a stranger (i.e., an unfamiliar experimenter) approached the child for the first time, and the presentation of masks, were expected to elicit fearful reactions from the infants. During the warm-up the stranger first talked to the child, who was sitting next to his/her mother, for one minute and then introduced a novel toy during the second minute. The mother was instructed not to talk to the child during the task. For the presentation of masks, the child was presented with four different masks, with each presentation lasting 10 seconds. The mother was seated to the side and slightly behind the child, and was instructed not to comment on the masks.

Both fear-eliciting episodes were coded in a manner that produced a number of behavioral indices, including Intensity of Bodily Fear, Intensity of Facial Fear, Intensity of Distress Vocalizations, and Intensity of Escape. These were rated according to a predetermined set of criteria. For example, Intensity of Bodily Fear was judged based on the following scale: 0= No sign of bodily fear; 1= Decreased activity: an apparent and/or sudden decrease in the activity; sense of bodily apprehension and ambiguous body fear; 2= Tensing: visible tensing of the muscles, associated with decreased activity; 3= Freezing or trembling: tensing of the entire body with no motion, or trembling due to extreme muscular tension; with all ratings assigned every 5 seconds, and averaged over the epochs. On the other hand, Intensity of Facial Fear was coded in accordance to the following scheme: 0= No facial region shows identifiable fear movement; 1= Only one facial region shows identifiable movement, indicating a low intensity fear, or fear expression appears ambiguous (e.g., eyes widen slightly; mouth opens slightly with corners retracting back; child appears to be mildly afraid facially); 2= Only 2 facial regions show identifiable movement, or expression in one region is definite (e.g., eyes widen, brows may be raised; mouth open or closed, with corners retracted straight back; nasal root narrowed, jaw drops); 3= An appearance change occurs in all 3 facial regions, or the coder otherwise has an impression of strong facial fear (eyes definitely widen, mouth corners retracted straight back; jaw may drop; eyebrows straight or raised up; may be horizontal wrinkle above the child's nose and near the inside of the eyebrows). Similarly scaled coding schemes were implemented in rating the remaining fear-related observation codes: Intensity of Distress Vocalizations and Escape.

On the other hand, the peek-a-boo game and unstructured play with the parent episodes were expected to produce positive reactivity (i.e., smiling/laughter). The unstructured play period with the mother lasted four minutes and the mother was simply instructed to play with child without getting him/her too excited, utilizing a play telephone provided for this episode. The peek-a-boo game involves the mother looking through a series of windows created in a wooden board, according to the instructions provided by the experimenter, while saying "peek-a-boo." Similar to the fear-eliciting episodes, activities designed to elicit positive emotionality were coded in a manner that produced a number of behavioral indices, including Presence of Smile,

Intensity of Smile, Positive Vocalizations, Positive Motor Activity, and Laughter. In the case of Intensity of Smile, the ratings were made on the basis of the following scale: 0= No smiling at all; 1= Small smile, with lips only slightly upturned, little or no involvement of cheeks, no crinkling about eyes; 2= Medium smile, with lips definitely upturned, mouth perhaps open, some bulging of the cheeks, perhaps slight crinkling about eyes; 3= Large smile, with lips stretched and quite upturned or perhaps mouth open, cheeks bulging, definite crinkling about eyes. The Positive Vocalizations indicator was determined on the basis of identifiable positively toned babbling, squealing, etc. (rated as 1= present, 0= not present), and represents another example of the coding scheme utilized in conjunction with the TLA in this study.

The TLA battery represents an alternative to the established and widely utilized Lab-TAB procedures (Goldsmith, & Rothbart, 1996), in so far as it corresponds more closely to the dimensions of temperament addressed by the IBQ-R. That is, the IBQ-R includes a number of additional temperament scales, relative to its predecessor (i.e., the IBQ), and the TLA provides structured laboratory episodes enabling researchers to observe parallels of these domains of temperament, evaluated via parent-report with the help of the IBQ-R.

In this study, this structured laboratory observations were conducted after participating parents had completed the questionnaire portion of the assessment (i.e., IBQ-R, etc.).

The Parenting Stress Index (PSI; Abidin, 1983)—The PSI provides indices of child and parent characteristics, family context, and stressful life events (Abidin, 1995), including the child and parent domains. Two of the parental domain subscales, namely Depression and Competence, were utilized in this study. Higher scores on the Depression scale are indicative of significant parental depression (Webster-Stratton & Hammond, 1988), and this scale has been utilized as an indicator of depressive symptomatology in previous investigations addressing difficult child behaviors, parental and family adjustment (Gartstein & Sheeber, 2004; Sheeber & Johnson, 1992). The Competence subscale assesses factors related to perceived parenting efficacy, with high scores interpreted as indicating low self-efficacy in the parenting domain (Abidin, 1995). Cronbach's alpha coefficients as high as .84 have been reported for the PSI subscales, and test-retest reliability estimates have ranged from .88 to .96 (Abidin, 1995). In the present sample Cronbach's alphas were somewhat lower, but still generally adequate: Competence = .55; Depression= .65.

The Adult Temperament Questionnaire (ATQ; Evans & Rothbart, 2003)—The most current version of the ATQ, the 177-item ATQ-2, contains 13 scales, which load onto four broad factors: Negative Affect (Fear, Discomfort, Frustration, Sadness), Extraversion/Surgency (Sociability, Positive Affect, High Intensity Pleasure), Effortful Control (Inhibitory Control, Attentional Control, Activation Control), and Orienting Sensitivity (Neutral Perceptual Sensitivity, Affective Perceptual Sensitivity, Associative Sensitivity). Satisfactory internal consistency has been demonstrated for ATQ-2, with Cronbach's alpha coefficients exceeding .80 on 13 of the 18 scales (Evans & Rothbart, 2003). For this study the over-arching factors of Extraversion/Surgency and Negative Affectivity were utilized, with Cronbach's alphas of .73 and .80, respectively.

Procedures

Potential participants, identified on the basis of the infant's age (6, 9, or 12 months), were contacted by telephone. The parents of the infants participating in this study were asked to complete the IBQ-R (Gartstein & Rothbart, 2003) no more than two weeks prior to the laboratory visit, to ensure a similar time-frame for the two assessment approaches. Additionally, parents completed a demographic questionnaire, the PSI (Abidin, 1983), and the ATQ (Evans & Rothbart, 2003). Participating families were also asked to come into the

laboratory for the Temperament Laboratory Assessment (TLA; Gonzales et al., 2003) administration, at a time convenient to them. Participants were told that taking part in this research would require approximately one and a half hours for questionnaires and one hour for the TLA procedure. Participating families were required to complete all of these activities within 2 weeks of the infant turning 6, 9, or 12 months of age, and were reimbursed for their participation (\$25.00).

Results

Development of Construct

First, Pearson Product Moment correlation coefficients were examined for fear and smiling/laughter laboratory-based indicators. Indices associated with significant correlations were subsequently subjected to a factor analytic examination and an evaluation of internal consistency (i.e., Cronbach's alphas were computed). Exploratory factor analysis (Principal Axis extraction) was conducted to develop the laboratory-based fear composite, using the previously coded fear-related behavioral indices with significant correlations (e.g., intensity of fearful facial expression, bodily fear, presence of escape behaviors, etc.). A single observation-based fear factor emerged ($\alpha = .74$), consisting of five fear-related indicators: intensity of bodily fear for the Warm-Up episode; intensity of facial fear, intensity of distress vocalization, intensity of escape, and intensity of bodily fear, all from the Masks episode. An identical factor analytic procedure was conducted to create the positive affectivity/smiling and laughter laboratory composite factor ($\alpha = .67$). The indicators reflecting the presence and intensity of smiling, positive motor activity, and positive vocalizations from the Interaction episode, as well as the presence and intensity of smiling, positive vocalizations, and laughter from the Peek-a-Boo episode, demonstrated significant correlations, and subsequently loaded onto this factor (Table 1). The laboratory based composite scores were subsequently computed by summing the standardized indicators, which have been demonstrated as contributing the fear and positive emotionality composite scores, respectively (Table 2).

Convergence between Temperament Measures

Next, correlation coefficients were computed among the laboratory composite scores and the parallel IBQ-R scales. Whereas the correlation between the laboratory-based positive affectivity/smiling and laughter composite and the Smiling and Laughter subscale of the IBQ-R ($r = .20$; $p = .11$) did not reach statistical significance, the laboratory fear composite and the IBQ-R Fear subscale were significantly correlated ($r = .28$; $p < .05$). Hence, this finding lends partial support to the hypothesis that the two sets of measures (i.e., laboratory-based and parent-report temperament indicators) would converge to a significant degree.

Analyses of Discrepancies

Pearson Product Moment correlation coefficients between parent characteristics and infant temperament indicators were computed first (Table 3). Next, Hierarchical Multiple Regression Analyses (HMRA's) were conducted, predicting differences between the laboratory-based composites and parent-report indicators of smiling/laughter and fearfulness. Difference scores were computed in order to provide an index of discrepancy, and were utilized as the dependent variables in these analyses. The fear difference score was created by subtracting the standardized IBQ-R Fear subscale from the standardized laboratory fear composite factor score. Thus, a standardized difference score was formed after a z-score transformation was applied to each of the indicators (i.e., laboratory composite and parent-report rating), due to the difference in scale, and served as the dependent variable. Hence, a negative difference score meant that parents reported their children to be more fearful than was deemed on the bases of the structured laboratory observations. A positive difference score, on the other hand, indicated that observers coded a greater level of fear than was reported by the parents. Although, different

analytical procedures could be utilized to derive the discrepancies of interest, the difference score approach has been advocated by those examining the contribution of maternal depression to discrepancies between mothers and other sources of information regarding child psychopathology (Briggs-Gowan, Carter, Schwab-Stone, 1996; Duhig, Renk, Epstein, & Phares, 2000; Richters, 1992; Treutler & Epkins, 2003). The difference score approach addressed the contribution of maternal characteristics to discrepancies between different sources of information, thus incorporating the inherent dissimilarities in the measurement approaches (i.e., parent-report and structured laboratory observations). In addition, this metric directly provided information regarding which methodological approach demonstrated higher vs. lower levels infant temperament attributes, central to our ability to evaluate the contribution of maternal characteristics to discrepancies in the ratings of infant fear and positive emotionality. For example, if higher levels of maternal depression were associated with to lower (i.e., negative) difference scores for infant fear, it would suggest that parents who reported higher levels of depression, also reported greater levels of fearfulness for their infants, relative to observation-based indices. The latter pattern of results would be significant in so far as it would support the plausibility of the distortion explanation, suggesting that parents experiencing symptoms of depression tend to overemphasize the negative aspects of their children's behavior.

During the first step of the HRA, SES, parental age, infant's age, and infant's gender were entered into the equation, because these were all considered to be demographic/background characteristics, and the decision was made to remove, or control for, their influence prior to examining the remaining independent variables. It should be noted that preliminary analyses (ANOVA's) indicated that the difference scores, parental depression, efficacy, negative emotionality, and positive emotionality did not differ significantly between the three age groups, or by gender, enabling us to simply statistically control for the potential contribution of these factors. Subsequently, parental characteristic's scores (i.e., depression, parenting competence, negative affectivity and positive emotionality) were entered into the regression equation in a single step (Table 4). Parental negative affectivity was the only attribute associated with a statistically significant effect ($\text{Beta} = -.34, p < .05$), having controlled for SES, parental age, infant's age, infant's gender, depression, and parental competency. Thus, parents higher in negative affect report a higher level of fear for their infants, relative to the results of the laboratory observation.

A similar procedure is followed to analyze which variables contribute to the discrepancy between the positive affectivity laboratory composite and the parent-report indicator of smiling/laughter. The standardized IBQ-R Smiling and Laughter subscale score was subtracted from the standardized laboratory positive affectivity/smiling and laughter composite. Significant predictors did not emerge in the analyses addressing the Smiling and Laughter difference score (Table 5).

Discussion

This study was designed to address two inter-related goals: (1) establish convergence between the laboratory-based temperament composites and parent-report ratings of temperament; (2) determine which factors influenced discrepancies between these two approaches to the measurement of infant temperament. As anticipated, fear assessed in the laboratory was significantly correlated with the fear indicator derived from parental report on the IBQ-R. However, contrary to our hypothesis, the laboratory composite of smiling and laughter was not significantly correlated with smiling and laughter assessed via parent-report. This lack of a significant relationship may be at least in part due to the novel laboratory environment not being conducive to expressing positive emotionality (i.e., smiling and laughing). Rothbart and Goldsmith (1985) warned that infants become wary in the laboratory environment due to

factors such as novelty of the situation/surroundings and the lack of familiarity with the adult experimenters. Thus, infant behavior related to positive emotionality in the laboratory may not be representative of the behavioral tendency (e.g., smiling, laughing) in the home environment. These findings provided evidence consistent with the cautionary statements of Rothbart and Goldsmith (1985), demonstrating the challenges of assessing positive affectivity in a reliable/valid manner in the laboratory setting.

The hypothesis that parental depression would predict increased reporting of fear, as compared to the levels of fear measured through laboratory observations, was not supported by the results of this investigation, likely due to the nature of the included sample. Specifically, a community sample was recruited for the present study, demonstrating both low rates of depression and little variance in symptom severity (mean = 1.95; SD = .587; range = 1 – 3.33 out of 5 possible). A similar explanation can be posed for the lack of explanatory power for the perceived parenting efficacy score (mean = 2.00; SD = .417; range = 1 – 3). However, results of this study indicate that parental negative affectivity was predictive of increased parental report of infant fear, relative to the level of fearfulness observed during the structured laboratory observation. This finding could be interpreted in light of the previously reported association between maternal depression and increased reporting of negative child temperament characteristics (e.g., Whiffen, 1990), as well as links between negative affectivity and depression. The Negative Affect factor of the ATQ consisted of four subscales: Fear, Sadness, Discomfort, and Frustration. The Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (DSM-IV-TR; American Psychiatric Association, 2000) stated that some frequently occurring associated features of depression included feeling tearful, irritable, and anxiety. Thus, considerable overlap exists between the signs and symptoms associated with depression and the experience of negative emotionality conceptualized as a broadband dimension of temperament, given that a person who is frequently sad and tearful, irritable and frustrated, and/or anxious/fearful would receive high negative affectivity scores and also be deemed as showing signs of depression. The nature of the present sample (i.e., community-based recruitment) is likely the reason for a significant effect associated with the parent temperament domain of Negative Affectivity, but not the indicator of depression utilized in this study.

Although parental characteristics as a whole (i.e., entered into the regression equation simultaneously) did not account for a statistically significant amount of variance in infant fear difference scores, parental negative emotionality was associated with a significant effect. This significant effect reflects the fact that parents higher in Negative Affectivity reported more fear for their children, relative to the fear observed during the laboratory assessment, and suggests that parental temperament influences the way caregivers interpret their children's behaviors. The presence of higher Negative Affectivity may influence parents' perception of their children in such a way that they are more likely to pay attention to cues of negative emotionality in their children. Thus, the finding suggests that we need to take parental temperament into consideration when interpreting parental report of child temperament attributes.

In terms of the limitations of this work, first, the reliance on difference scores should be noted, given that the use such scores has been criticized on the bases of concerns with reduced reliability, ambiguities in interpretation, and potential for confounded effects (Edwards, 2002). The difference is said to be inevitably lower in its reliability, relative to its component measures; however, Tisak and Smith (1994) remarked that the reliability of difference scores would be acceptable if the component measures were reliable and not highly positively correlated, which is the case with the laboratory-based temperament composites and parent-report IBQ-R indicators utilized in this study. Furthermore, it has been noted that reduced reliability was not a weakness unique to analyses utilizing difference scores (Sheeber, Sorensen, & Howe, 1996). In this study, the parent-report and laboratory-based indicators utilized to compute difference scores were standardized, due to the differences in scale across

the two sets of scores. The resulting standardized difference scores seem readily interpretable given that the questionnaire and the laboratory assessment tools were designed to measure manifestations of the same underlying aspects of temperament, and the laboratory measure was designed to parallel the questionnaire in terms of content. The use of difference scores became necessary in the context of this study because of an attempt to refrain from assigning one of the sources of information (i.e., parent-report vs. laboratory-based indicators) as the dependent variable, thereby implying that it represented a “gold standard” against which all other measures should be evaluated.

It is of utmost importance to continue the study of the development of temperament, as well as the strengths and weaknesses of the different measurement approaches available for the assessment of temperament in childhood. The present study represents a step in this direction, although, additional research is needed to replicate and extend our findings. This study was necessarily limited in its scope, in that only the fear domain of negative emotionality and the smiling and laughter aspect of positive affectivity were studied. It would be important for future research to examine the convergence and discrepancies between the remaining subscales of the IBQ-R and the TLA components, in order to determine whether or not factors considered in this study influence discrepancies between the two measurement approaches in assessing additional temperament attributes (e.g., Perceptual Sensitivity). It should also be acknowledged that the non-significant relationship between the laboratory composite and the parent-report indicator addressing smiling/laughter could have resulted from the relatively small sample size, and subsequently limited power, of the present investigation. Thus, future research would benefit from including larger sample, leading to greater power to detect potentially meaningful effects. Furthermore, this study addressed convergence/discrepancies among only two possible methodological approaches for assessing infant temperament. However, home observations have also been used to assess infant temperament, and it would be beneficial to study the convergence between home observations and the laboratory assessment, as well as the parent-report measures. The latter addition would allow us to gain further insight into “untangling” effects of setting (home vs. laboratory) and method (parent report vs. observation) on childhood temperament scores, their convergence and discrepancies. Furthermore, we chose to focus on parent factors that may influence discrepancies between parent-report and laboratory observation. Future research should consider child characteristics that could also influence the discrepancies between the two sources of information, possibly contributing to the deviation of the laboratory observation score from the parent-report indicators. For example, children who exhibit lower levels of regulation may be particularly vulnerable to the effects of encountering a novel laboratory setting and unfamiliar adults, limiting the range of positive emotionality expressed during the observation period. Finally, it could also be argued that the fact that some of the questionnaires were completed by fathers (N=2) represents a limitation of the present investigation. However, analyses aimed at evaluating hypotheses of this study were also examined with a sub-sample including only mothers, and did not differ from the results of those analyses conducted with an entire sample.

References

- Abidin, RR. Parenting Stress Index. Charlottesville, VA: Pediatric Psychology Press; 1983.
- Abidin, RR. Parenting Stress Index—Manual. 3rd. Charlottesville, Virginia: Pediatric Psychology Press; 1995.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th, text revision. Washington, DC: Author; 2000.
- Beck CT. A meta-analysis of the relationship between postpartum depression and infant temperament. *Nursing Research* 1996;45(4):225–230. [PubMed: 8700656]

- Briggs-Gowan MJ, Carter AS, Schwab-Stone M. Discrepancies among mother, child, and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology* 1996;24:749–765. [PubMed: 8970908]
- Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 1959;56(2):81–105. [PubMed: 13634291]
- Carter AS, Little C, Briggs-Gowan MJ, Kogan N. The infant-toddler social and emotional assessment (ITSEA): Comparing parent ratings to laboratory observations of task mastery, emotion regulation, coping behaviors, and attachment status. *Infant Mental Health Journal* 1999;20(4):375–392.
- Duhig AM, Renk K, Epstein MK, Phares V. Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice* 2000;7:435–453.
- Edborg M, Seimyr L, Lundh W, Widström AM. Fussy child-difficult parenthood? Comparisons between families with a “depressed” mother and non-depressed mother 2 months postpartum. *Journal of Reproductive and Infant Psychology* 2000;18(3):225–238.
- Edwards, JR. Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In: Drasgow, F.; Schmitt, N., editors. *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis*. San Francisco: Jossey-Bass; 2002. p. 350-400.
- Evans, D.; Rothbart, MK. A hierarchical model of temperament and the Big Five. 2003. Manuscript in preparation
- Fergusson DM, Lynskey MT, Horwood LJ. The effect of maternal depression on maternal ratings of child behavior. *Journal of Abnormal Child Psychology* 1993;21(3):245–269. [PubMed: 8335763]
- Field T, Greenberg R. Temperament ratings by parents and teachers of infants, toddlers, and preschool children. *Child Development* 1982;53:160–163.
- Gartstein MA, Rothbart MK. Studying infant temperament via a revision of the Infant Behavior Questionnaire. *Journal of Infant Behavior and Development* 2003;26:64–86.
- Gartstein MA, Sheeber LB. Maternal dysphoria and child behavior problems: A mediational model. *Journal of Child and Adolescent Psychiatric Nursing* 2004;17:141–150. [PubMed: 15742795]
- Gartstein MA, Slobodskaya HR, Kinsht IA. Cross-cultural differences in the first year of life: United States of America (U.S.) and Russia. *International Journal of Behavioural Development* 2003;27(4):316–328.
- Goldsmith HH, Campos JJ. The structure of temperamental fear and pleasure in infants: A psychometric perspective. *Child Development* 1990;61(6):1944–1964. [PubMed: 2083507]
- Goldsmith, HH.; Rothbart, MK. Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In: Angleitner, A.; Strelau, J., editors. *Explorations in temperament: International perspectives on theory and measurement*. New York: Plenum; 1991. p. 249-272.
- Goldsmith, HH.; Rothbart, MK. The Laboratory Temperament Assessment Battery (LAB-TAB): Prelocomotor Version 3.0. Personality Development Laboratory, Department of Psychology, University of Wisconsin; Madison, WI 53706: 1996. Available from Hill H. Goldsmith, Ph.D.
- Gonzalez, C.; Gartstein, MA.; Carranza, JA.; Rothbart, MK. Temperament Laboratory Assessment: Procedures parallel to the Infant Behavior Questionnaire-Revised. Department of Psychology; Washington State University, P.O. Box 644820, Pullman, WA: 2003. p. 99164-4829. Manual available from the second author: M.A. Gartstein, Ph.D
- Kagan, J. Galen’s Prophecy: Temperament in human nature. New York: Basic Books; 1994.
- Kagan, J. Biology and the child. In: Damon, W.; Eisenberg, N., editors. *Handbook of child psychology: Vol. 3. Social, emotional and personality development*. 5th. New York: Wiley; 1998. p. 177-235.
- Kochanska G, Clark LA, Goldman MS. Implications of mothers’ personality for their parenting and their young children’s developmental outcomes. *Journal of Personality* 1997;62(2):387–420. [PubMed: 9226943]
- Kochanska G, Murray K, Coy KC. Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development* 1997;68(2):263–277. [PubMed: 9180001]
- Leerkes EM, Crockenberg SC. The impact of maternal characteristics and sensitivity on the concordance between maternal reports and laboratory observations of infant negative emotionality. *Infancy* 2003;4:517–539.

- Marmion, J.; Gartstein, M.; Dockery, N. Fear and positive affectivity in infancy: Convergence/ discrepancy between parent-report and laboratory based indicators. Poster session presented at the conference of the International Society on Infant Studies; Chicago, IL. 2004 May.
- Richters JE. Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin* 1992;112:485–499. [PubMed: 1438639]
- Richters J, Pellegrini DS. Depressed mothers' judgments about their children: An examination of the depression-distortion hypothesis. *Child Development* 1989;60(5):1068–1075. [PubMed: 2805884]
- Rothbart MK. Measurement of temperament in infancy. *Child Development* 1981;52:569–578.
- Rothbart MK, Ahadi SA, Hershey KL. Temperament and social behavior in childhood. *Merrill-Palmer Quarterly* 1994;40:21–39.
- Rothbart, MK.; Bates, JE. Temperament. In: Damon, W.; Lerner, R.; Eisenberg, N., editors. *Handbook of child psychology: Vol. 3. Social, emotional, and personality development*. 6th. New York: Wiley; 2006. p. 99-166.
- Rothbart, MK.; Derryberry, D. Development of individual differences in temperament. In: Lamb, ME.; Brown, AL., editors. *Advances in developmental psychology*. 1. Hillsdale, NJ: Erlbaum; 1981. p. 37-86.
- Rothbart MK, Goldsmith HH. Three approaches to the study of infant temperament. *Developmental Review* 1985;5:237–260.
- Sheeber LB, Johnson JH. Child temperament, maternal adjustment, and changes in family life style. *American Journal of Orthopsychiatry* 1992;62(2):178–185. [PubMed: 1580337]
- Sheeber LB, Sorensen ED, Howe SR. Data analytic techniques for treatment outcome studies with pretest/posttest measurements: An extensive primer. *Journal of Psychiatric Research* 1996;30(3):185–199. [PubMed: 8884657]
- Stevens G, Featherman DL. A revised socioeconomic index of occupational status. *Social Science Research* 1981;10:365–395.
- Teti DM, Gelfand DM. Behavioral competence among mothers of infants in the first year: The mediational role of maternal self-efficacy. *Child Development* 1991;62:918–929. [PubMed: 1756667]
- Tisak J, Smith CS. Defending and extending difference score methods. *Journal of Management* 1994;20:675–682.
- Treutler CM, Epkins CC. Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent-child relationships? *Journal of Abnormal Child Psychology* 2003;31:13–27. [PubMed: 12597696]
- Webster-Stratton C, Hammond M. Maternal depression and its relationship to life stress, perceptions of child behavior problems, parenting behaviors, and child conduct problems. *Journal of Abnormal Child Psychology* 1988;16:299–315. [PubMed: 3403812]
- Whiffen VE. Maternal depressed mood and perceptions of child temperament. *Journal of Genetic Psychology* 1990;151(3):329–339. [PubMed: 2266357]

Table 1
Factor Loadings for the Laboratory Composites: Fear and Positive Affectivity/Smiling and Laughter

Individual Fear Indicators	Factor Loading
Intensity of Bodily Fear (Warm-Up)	.429
Intensity of Facial Fear (Masks)	.873
Intensity of Distress Vocalization (Masks)	.849
Intensity of Escape (Masks)	.635
Intensity of Bodily Fear (Masks)	.817
Individual Smiling and Laughter Indicators	Factor Loading
Presence of Smile (Interaction)	.661
Intensity of Smile (Interaction)	.821
Positive Vocalization (Interaction)	.512
Positive Motor Activity (Interaction)	.372
Presence of Smile (Peek-a-Boo)	.797
Intensity of Smile (Peek-a-Boo)	.805
Laughter (Peek-a-Boo)	.421
Positive Vocalization (Peek-a-Boo)	.658

Table 2
Descriptive Information for the Temperament Laboratory Assessment (TLA) Scores

Score	Mean	Standard Deviation
Laboratory-based Fear Composite	3.57	2.68
<u>Codes from Masks Episode:</u>		
Facial Fear Intensity	1.33	.86
Distress Vocalization Intensity	.84	1.30
Body Fear Intensity	1.04	.59
Escape Intensity	.35	.37
<u>Codes from Warm-Up Episode:</u>		
Body Fear Intensity	.01	.09
Laboratory-based Smiling/Laughter Composite	20.07	10.66
<u>Codes from Interaction Episode:</u>		
Presence of Smiling	.27	.26
Intensity of Smiling	.46	.52
Positive Motor Activity	.42	.33
Positive Vocalizations	.18	.23
<u>Codes from Peek-A-Boo Episode:</u>		
Presence of Smiling	.79	.35
Intensity of Smiling	1.81	1.22
Positive Vocalizations	.28	.35

Note. The composite scores represent sums of standardized individual laboratory-based indicators; these individual codes are presented in their unstandardized form.

Table 3

Pearson Product Moment correlation coefficients between parent characteristics and infant temperament variables.

	Fear Laboratory Composite	Positive Affectivity Laboratory Composite	IBQ Smiling/ Laughter	IBQ Fear
Parental Positive Affect	.01	.08	-.00	.04
Parental Negative Affect	-.12	.07	.23	.20
Parenting Competence	.11	-.06	-.10	.17
Parental Depression	-.00	-.05	-.09	.13

*
p < .05,

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

Table 4
Hierarchical Regression Equation for the Fear Difference Scores: Change Statistics

Step and Independent Variable	R	R ²	R ² _Δ	F _Δ ^a	B
1. Background variables					
SES	.13	.02	.02	.27	-.09
Parental age					-.06
Infant's gender					.02
Infant's age					-.06
2. Parent variables					
Parental Depression	.37	.14	.12	1.92	-.23
Parenting Competence					.07
Parental Negative Affect					-.34*
Parental Positive Affect					-.05

* $p < .05$,

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

^aF tests addressing significance of change in R², not the entire regression equation.

Table 5 Hierarchical Regression Equation for the Positive Emotionality Difference Scores: Change Statistics

Step and Independent Variable	R	R ²	R _A ²	F _A ^a	B
1. Background variables					
SES	.28	.08	.08	1.27	.13
Parental age					-.10
Infant's gender					-.21
Infant's age					.10
2. Parent variables	.30	.09	.01	.14	.05
Parental Depression					.03
Parenting Competence					-.06
Parental Negative Affect					.01
Parental Positive Affect					

* $p < .05$,

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

^aF tests addressing significance of change in R², not the entire regression equation.