HHS Public Access

Author manuscript

Infant Behav Dev. Author manuscript; available in PMC 2017 May 01.

Published in final edited form as:

Infant Behav Dev. 2016 May ; 43: 24–35. doi:10.1016/j.infbeh.2016.04.001.

The Infant Behavior Questionnaire-Revised: Factor Structure in a Culturally and Sociodemographically Diverse Sample in the United States

Michelle Bosquet Enlow and **Matthew T. White**

Boston Children's Hospital and Harvard Medical School

Katherine Hails and **Ivan Cabrera**

Boston Children's Hospital

Rosalind J. Wright

Kravis Children's Hospital, Department of Pediatrics and Mindich Child Health & Development Institute, Icahn School of Medicine at Mount Sinai

Abstract

The recommended factor structure for the Infant Behavior Questionnaire-Revised (IBQ-R), a widely used parent-report measure of infant temperament, has limited empirical support. Moreover, the recommended factors were developed using homogenous samples not representative of current United States (U.S.) sociodemographics. The objective of this study was to examine the factor structure of the IBQ-R in a culturally and sociodemographically diverse U.S. cohort (N = 380 mother-infant dyads). Mothers were assessed during pregnancy on a range of cultural and sociodemographic characteristics and completed the IBQ-R when their infants were 6 months of age. The sample was diverse on maternal marital status, educational attainment, household income, race/ethnicity, primary language spoken, and country of birth. Initial confirmatory factor analysis for the recommended three-factor model yielded a poor fit. Modifications employed in other studies failed to improve model fit. An exploratory factor analysis revealed an acceptable model fit for a three-factor solution that showed similarities to as well as differences from the originally proposed factor structure. Additional analyses suggested lack of invariance on several factor and scale scores by maternal country of birth, race/ethnicity, and household income. The findings suggest that the commonly used IBQ-R factor structure may need to be adjusted for diverse samples and deserves further study.

Keywords

Infant Behavior Questionnaire-Revised (IBQ-R); temperament; factor analysis; factor structure; sociodemographic diversity; cultural differences

1. Introduction

The Infant Behavior Questionnaire-Revised (IBQ-R) is a widely-used parent-report measure developed to assess dimensions of temperament along 14 scales (Table 1) in infants between 3 and 12 months of age (Gartstein & Rothbart, 2003). The IBQ-R was rationally derived, based on a definition of temperament as constitutionally based individual differences in reactivity (e.g., arousability of emotional, motor, and attentional responses) and self-regulation (processes that modulate reactivity, such as attention; Gartstein & Rothbart, 2003). Parent-report questionnaires of infant temperament have proven to be valuable research tools given a number of benefits they provide over laboratory observations. Such benefits include a low cost method for obtaining information on multiple temperament dimensions, the opportunity to gather data on behavior across various contexts that are not affected by the artificiality of structured observations, and simpler logistics of administration and scoring (Bornstein et al., 2015; Dragan, Kmita, & Fronczyk, 2011; Montirosso, Cozzi, Putnam, Gartstein, & Borgatti, 2011). Moreover, the IBQ-R has demonstrated good internal consistency, reliability, and validity, including correlations with laboratory observations (Gartstein & Marmion, 2008; Goldsmith & Campos, 1990; Parade & Leerkes, 2008).

1.1. IBQ-R and the Structure of Infant Temperament

In addition to being used to measure individual differences in infant temperament characteristics, the IBQ-R has been utilized to identify the structure of infant temperament via factor analysis (Clifford et al., 2013; Derauf et al., 2011; Gartstein et al., 2010; Gartstein & Rothbart, 2003; Putnam, Rothbart, & Gartstein, 2008). When the IBQ-R was first introduced (Gartstein & Rothbart, 2003), the developers presented the results of an exploratory factor analysis (EFA), which revealed three factors, labeled by the developers as (1) Surgency/Extraversion, consisting of the scales Approach, Vocal Reactivity, High Intensity Pleasure, Smiling and Laughter, Activity Level, and Perceptual Sensitivity; (2) Negative Affectivity, consisting of the scales Sadness, Distress to Limitations, Fear, and Falling Reactivity/Rate of Recovery from Distress; and (3) Orienting/Regulation, consisting of the scales Low Intensity Pleasure, Cuddliness, Duration of Orienting, and Soothability (Gartstein & Rothbart, 2003). Internal consistency ratings of the factors were high, with Cronbach's alpha values ranging from 0.91 to 0.92 (Gartstein & Rothbart, 2003).

The authors noted that this factor structure is generally consistent with broad dimensions of temperament documented for older children and adults and with theories of personality (Gartstein & Rothbart, 2003; Rothbart, Ahadi, Hershey, & Fisher, 2001). In subsequent years, investigators have utilized these factors in analyses relating infant temperament to a variety of child outcomes (e.g., Brunst et al., 2014; Rothbart, Sheese, Rueder, & Posner, 2011; Tester-Jones, O'Mahen, Watkins, & Karl, 2015; Tikotzky, Chambers, Gaylor, & Manber, 2010; van den Heuvel, Johannes, Henrichs, & Van den Bergh, 2015). Moreover, much of the recent literature regarding infant temperament has been based on IBQ measures (original Infant Behavior Questionnaire [IBQ], IBQ-R) of the structure of infant temperament (Bornstein et al., 2015).

1.2. Cultural and Sociodemographic Influences on Infant Temperament

Although temperament characteristics are hypothesized to be biologically based, they are also hypothesized to be impacted by experience (Gartstein et al., 2006), including cultural and sociodemographic influences (Gartstein, Knyazev, & Slobodskaya, 2005). Children's social settings, familial customs, and caregivers' psychosocial characteristics and beliefs may influence the meaning of children's temperament (Cozzi et al., 2013; Gartstein, Peleg, Young, & Slobodskaya, 2009; Super & Harkness, 1986). Consequently, parents may engage in caregiving practices that reinforce the characteristics consistent with their values and those of their cultural group (Dragan et al., 2011; Gartstein et al., 2006, 2009), leading to cross-cultural differences in the development of temperament (Gartstein et al., 2006). Support for this hypothesis comes from a number of studies among infants from different cultural groups that show differences on mean ratings across temperament scales (e.g., Cozzi et al., 2013; Dragan et al., 2011; Gartstein et al., 2006, 2009; Montirosso et al., 2011). Importantly, limited longitudinal data suggest that cross-cultural differences in parental report of infant temperament represent true differences in infant behavior, not differences in parental perception/reporting biases (Gartstein et al., 2006).

The propensity to encourage certain characteristics and discourage others may also influence the way that different temperament traits "hang together." For example, among infants living in groups that encourage or accept high rates of activity, activity level may correlate with extraversion/surgency/positive affect scales; among infants living in groups that discourage high activity, activity level may correlate with Negative Affectivity scales. Such variations in the value of specific traits may result in different temperament factor structures across cultural and sociodemographic groups.

Although the IBQ-R has been adapted into many languages and studies support its reliability and validity in a variety of cultures (Gartstein & Rothbart, 2003; Gartstein, Slobodskaya, & Kinsht, 2003; Montirosso et al., 2011), there has been relatively little examination of potential cultural and other sociodemographic influences on its factor structure (Gartstein et al., 2009). A small number of studies have attempted to replicate the original factor structure of the IBQ-R in samples representing different cultures. Findings suggest some similarities but also notable differences across studies. For example, Gartstein et al. (2005) found several dissimilarities in the IBQ-R factor structure when comparing U.S. and Russian samples; a relatively consistent pattern of factor loadings across the samples was found only following model modifications. Additionally, confirmatory factor analysis (CFA) of the original three-factor structure reported by Gartstein and Rothbart (2003) found it to be a poor fit in the U.S. sample (Gartstein et al., 2005). However, another study found that the three-factor structure was an acceptable fit for samples of U.S. and Italian infants (Montirosso et al., 2011). In a sample of Polish infants, Dragan et al. (2011) used CFA to evaluate the factor structure of the IBQ-R testing three models: the original model described by Gartstein and Rothbart (2003) and the two models obtained by Gartstein et al. (2005). They found that none of the models matched their data. Using EFA followed by CFA, they arrived at a three-factor solution, with some overlap with the original Gartstein and Rothbart (2003) factor structure but also some discrepancies (e.g., Activity Level loading with Negative Affectivity scales). Vonderlin, Ropeter, and Pauen (2012) and Mink, Henning, and Aschersleben (2013)

confirmed the dimensions of Surgency/Extraversion and Negative Affectivity but did not replicate the Orienting/Regulation dimension in separate samples of German infants. Moreover, they found that several scales did not load onto the same factors as they did in the original U.S. sample; for example, as with the Polish sample, Activity Level loaded with Negative Affectivity scales. Although they did not examine factor structure differences across groups, Gartstein, Slobodskaya, Żylicz, Gosztyla, and Nakagawa (2010) found different patterns of correlations among IBQ-R scales when comparing scores across samples from the U.S., Japan, Russia, and Poland. Such discrepancies across studies have led several investigators to call for more research into the factor structure of infant temperament, particularly in large, diverse samples (Gartstein & Rothbart, 2003; Vonderlin et al., 2012).

Parental immigration and acculturation status are also hypothesized to help shape infant temperament (Gartstein et al., 2009). Gartstein et al. (2009) posited that immigration is a significant stressor that has major implications for caregiver functioning and parenting practices, which may, in turn, influence the development of offspring temperament. Moreover, acculturating to a new social, political, and economic environment may change the child traits that parents value and want to encourage or discourage (Gartstein et al., 2009). The very limited available data support the hypothesis that acculturation status influences parent report of infant temperament. For example, Gartstein et al. (2009) compared Russian families who were living in Russia to Russian families who had immigrated to Israel or the U.S and found differences among the groups on the mean score of five IBQ-R scales and on associations between measures of acculturation and infant temperament.

The potential impact of other sociodemographic factors (e.g., race/ethnicity, parental marital status, family socioeconomic status [SES]) on infant temperament structure has received little attention. To date, the studies that have examined the IBQ-R factor structure, including those comparing different cultural groups, have focused primarily on White samples (e.g., in Gartstein et al. 2005, 87% of the U.S. sample and 99% of the Russian sample were White) with relatively homogenous sociodemographic characteristics (primarily married, educated, middle to high SES). Importantly, limited available data suggest that sociodemographic characteristics may influence parental ratings of infant temperament. For example, in a U.S. sample, Parade and Leerkes (2008) found that older, more affluent, and more highly educated mothers rated their infants lower on positive IBQ-R scales (e.g., Smiling and Laughter, High Intensity Pleasure) than did other mothers. A study using the IBQ found that infants from higher SES backgrounds were rated as showing lower levels of Soothability and higher levels of Distress to Limitations (Ventura & Stevenson, 1986). Indirect evidence comes from studies linking sociodemographic characteristics (e.g., SES, race/ethnicity, parental marital status) to the development of child self-regulation, given that the IBQ-R was created based on the definition of temperament as individual differences in reactivity and self-regulation (Bornstein et al., 2015; Gartstein et al., 2009; Lengua, Honorado, & Bush, 2007). Moreover, if temperament is influenced by the environment and different environments are differentially stable (e.g., children in lower SES environments experience more disruptions), then sociodemographic characteristics that decrease environmental stability may result in greater changes in temperament across time (Bornstein et al., 2015).

Taken together, these findings suggest that sociodemographic characteristics may influence the structure of infant temperament.

1.3. The Current Study

Research is needed to determine whether the currently recommended factors for the IBQ-R are appropriate across cultural and sociodemographic groups to maximize the utility of the measure. The goal of the current study was to conduct initial steps toward confirming the validity of the IBQ-R in a culturally and sociodemographically diverse sample. Specifically, the objectives were to (a) examine the factor structure of the IBQ-R in a sample of 6-month-old infants living in the northeastern U.S., diverse in cultural and sociodemographic characteristics (e.g., maternal country of birth, race/ethnicity, marital status, educational attainment, household income) and (b) determine if the means of any factors or if the intercepts of any scale scores of the IBQ-R showed evidence of non-invariance across cultural and sociodemographic groups.

2. Method

2.1. Participants

Participants were mothers and their 6-month-old infants (N = 380 dyads; 53.4% of infants male) enrolled in the PRogramming of Intergenerational Stress Mechanisms (PRISM) study, a prospective pregnancy cohort originally designed to examine the role of perinatal stress exposures on child development and health. Between March 2011 and July 2014, pregnant women were recruited from prenatal clinics in Boston and New York City hospitals and from a Boston community health center. Recruitment sites were chosen based on desired heterogeneity in sociodemographic and racial/ethnic characteristics. Eligibility criteria included: 1) English- or Spanish-speaking; and 2) age ≥ 18 years at enrollment. Exclusion criteria included 1) maternal endorsement of drinking ≥ 7 alcoholic drinks/week prior to pregnancy recognition or any alcohol following pregnancy recognition, as usage at or above these thresholds has been associated with increased risk for a range of health problem (Patra et al., 2011; Testa, Quigley, & Eiden, 2003); 2) high-risk infant (e.g., gestational age < 32 weeks; birth weight < 5.5 lbs; congenital abnormalities; neurological injury), due to increased risk for neurodevelopmental problems; and 3) maternal or child chronic health conditions that would impede study participation. Among 548 eligible women, 69.4% of those approached agreed to participate and continued active follow-up after delivery. Based on screening data, there were no significant differences on race/ethnicity, education, or income between women who enrolled and those who declined. Dyads with IBQ-R data, assessed when the infants were approximately 6 months of age (M = 26.7 weeks, SD = 2.5), were included in the current analyses, resulting in a sample size of N = 380.

2.2. Procedures

Procedures were approved by the relevant institutions' human studies committees; written consent was obtained in the participant's preferred language. Participant cultural and sociodemographic data were obtained shortly following recruitment. Mothers were administered the IBQ-R in a face-to-face interview during a home visit when the infants were approximately 6 months of age.

2.3. Measures

2.3.1. Cultural and sociodemographic variables—Participants provided information on maternal age, relationship status, educational attainment, annual household income, race/ethnicity, primary language spoken in the home, country of birth, maternal age when moved to the U.S., length of time living in the U.S. among those born outside the U.S, and parity. Maternal age was scored continuously. Maternal relationship status was categorized as follows: married, living with partner, never married/single, divorced, separated, or other. Maternal educational attainment was scored categorically: not completed high school, completed high school/general equivalency diploma (GED), completed associate's degree or some college, completed a college degree, or completed a graduate degree. Annual household income was categorized into seven mutually exclusive categories ranging from less than $10,000 to greater than $100,000. Maternal race/ethnicity was categorized as White, Black/Haitian, Hispanic, or other; the other group consisted primarily of participants who self-identified as Asian or multi-racial. Primary language spoken in the home was categorized as English, Spanish, or other. Maternal birth country was categorized by geographic region. Maternal age when she moved to the U.S. was categorized as younger than 5 years of age, 5 to 10 years old, 11 to 17 years old, 18 to 25 years old, or older than 25 years. Length of time living in the U.S. was categorized as less than one year, one to two years, three to five years, or greater than five years. Parity was scored as primiparous or multiparous.

2.3.2. Infant Behavior Questionnaire-Revised (IBQ-R; Gartstein & Rothbart, 2003)—The IBQ-R is a caregiver-report instrument developed to assess temperament in infants from 3 to 12 months of age. The IBQ-R has demonstrated good internal consistency (Cronbach's alphas typically range from .70 to .90; Gartstein et al., 2009), and reliability and validity have been supported in samples from different cultures (Gartstein et al., 2005; Parade & Leerkes, 2008). The measure is available in nearly twenty languages, including English and Spanish (Gartstein et al., 2006). The 191 items were designed specifically to reduce the influence of reporter biases by inquiring about concrete infant behaviors rather than asking for abstract judgments.

The IBQ-R was administered as an interview in the mother's preferred language (English or Spanish). Mothers rated the frequency that their infant engaged in specific day-to-day behaviors in the prior one to two weeks using a 7-point scale, with responses ranging from 1 (*never*) to 7 (*always*). Mothers indicated their scores using a card with response choices. Item scores were summed according to IBQ-R scoring rules to create scores on the 14 scales, with higher scores indicating greater levels of that temperament dimension. See Table 1 for a description of each scale.

2.4. Data Analytic Plan

Tests for normality, descriptive statistics, and correlation coefficients were evaluated using SAS version 9.3 or SPSS version 22 prior to conducting factor analysis modeling. The Shapiro-Wilk test was used to evaluate the normality of each IBQ-R scale; only three of the scales (Activity Level, Sadness, Distress to Limitations) were normally distributed. Due to the non-normality of several scales, maximum likelihood estimation with robust standard

errors was utilized when conducting the factor analyses; these models were run using Mplus version 7.3.

The factor analysis plan was as follows: First, CFA was used to test previously published models of the IBQ-R factor structure and assess their goodness of fit. Measures of fit included the model Chi-square (with the caveat that this measure may be overly sensitive with a large sample size), the Tucker-Lewis Index (TLI), the Comparative Fit Index (CFI), and the root mean square error of approximation (RMSEA). As recommended by Hu and Bentler (1999), a good model fit was defined as TLI > 0.95, CFI > 0.95, and RMSEA < 0.06. If these models failed to provide a good fit, an EFA would be used to find an appropriate factor structure. After an appropriate factor structure was determined, in a Multiple Indicator Multiple Causes (MIMIC) model, sociodemographic covariates were included in the factor analysis, specifying direct effects of the covariates on the latent factors as well as allowing potential direct effects of the covariates on the manifest variables. The MIMIC model can therefore be used to assess group differences (i.e., non-invariance) in factor means and scale score intercepts. Maternal country of birth (U.S. vs. non-U.S. [reference group]), maternal race/ethnicity (Hispanic, Black/Haitian, or White [reference group]), and annual household income (≥ $40,000 vs. < $40,000 [reference group]) were considered as covariates. Mothers identifying their race as “other” were excluded from the MIMIC analysis, as the subsample size (n = 25) was too small to conduct meaningful analyses. To assess the direct effects of the covariates on the manifest variables, a forward-selection approach (Brown, 2006) was used in which the loading between each covariate and each manifest variable is constrained to zero and then modification indices are used to successively free loadings that significantly improve model fit. All tests were 2-sided at the .05 level of significance.

3. Results

3.1. Sample Cultural and Sociodemographic Characteristics

As displayed in Table 2, the sample was culturally and sociodemographically diverse. Approximately half of the mothers were married. Maternal educational attainment was varied, with less than half obtaining a college degree. Annual household income was also varied, with approximately half earning < $40,000 per year. Maternal race/ethnicity was approximately equally distributed among Hispanic (34.7%), Black/Haitian (30.5%), and White (28.2%), with a small percentage identifying as “other” (6.6%). Participants were primarily English-speaking (72.9%), with a substantial minority primarily Spanish-speaking (21.1%) and a small minority primarily speaking a language other than English or Spanish (6.1%). Forty-one percent of maternal participants were born outside of the U.S.; non-U.S.-born participants originated from various geographic regions, with the largest subgroup from Central America and the remaining from North and South America, Africa, Europe, and Asia. Among participants born outside the U.S., 35.5% moved to the U.S. during childhood or adolescence, 40.1% during early adulthood, and 21.9% during later adulthood. The vast majority of non-U.S.-born participants (80.0%) had lived in the U.S. for at least 5 years. Approximately one quarter of the mothers were primiparous.

3.2. IBQ-R Scale Analyses

Table 3 presents the means, standard deviations, and Cronbach's alpha values for each of the IBQ-R scales as well as all Spearman bivariate correlation coefficients among the scales. The Cronbach's alpha values indicate acceptable to excellent internal consistency across scales in this sample, similar to those reported in other samples (Gartstein et al., 2003, 2009). The magnitude of the correlation coefficients among scales ranged from negligible to large, with stronger correlations among the scales that were assigned to the same factor in the original IBQ-R study (Gartstein & Rothbart, 2003).

3.3. Confirmatory and Exploratory Factor Analyses

Initial CFA results for the three-factor model presented in Gartstein and Rothbart (2003) yielded a poor fit to the sample (Table 4). Notably, Gartstein et al. (2005) also found that this model did not fit well in a sample of U.S. infants. Following the approach by Gartstein et al. (2005), the Cuddliness scale was then excluded from the three-factor model; this step did not improve the model fit (Table 4). The models proposed by Gartstein et al. (2005) using two U.S. samples (see Figures 1 and 2 in Gartstein et al., 2005) were then fit. These models also failed to provide an acceptable fit to the current data (Table 4).

Next, an EFA was conducted to determine the best fitting factor structure for the current sample. Results revealed that a three-factor solution was appropriate; an examination of the factor loadings led to the exclusion of the Cuddliness and Soothability scales from further consideration due to their low factor loadings. An examination of modification indices then led to the exclusion of the Approach scale from further consideration due to the substantial residual covariation of the error term with a number of other variables. Finally, the examination of modification indices led to the addition of several cross-loadings (for the Vocal Reactivity, Activity Level, and Fear scales) and an error covariance term between the High Intensity Pleasure scale and the Low Intensity Pleasure scale (Figure 1). These final steps led to a good model fit (Table 4). The final factors showed acceptable internal consistency: Surgency Factor Cronbach's $\alpha = .71$, Negative Affectivity Factor Cronbach's $\alpha = .67$, and Orienting/Regulation Factor Cronbach's $\alpha = .63$, although these values are lower than those reported by Gartstein and Rothbart (2003).

3.4. MIMIC Model

A MIMIC model was built to assess the effect of covariates on the factor structure obtained from the EFA (Figure 2). In addition to the 25 participants who reported their race as "other," 19 participants were excluded due to having missing data on one or more covariates. The final model ($N = 336$) demonstrated a good fit to the data (Table 4).

Infants of U.S.-born mothers had a significantly lower mean score for the Orienting/Regulation factor than infants of non-U.S.-born mothers ($p = 0.012$). Infants of U.S.-born mothers also had significantly lower than expected scores for the Smiling and Laughter and Activity Level scales than infants of non-U.S.-born mothers ($p < 0.001$ and $p = 0.003$, respectively). Therefore, the Orienting/Regulation factor mean and the intercepts for the Smiling and Laughter and Activity Level scales were not invariant between the maternal U.S.- and non-U.S.-born groups.

Hispanic and Black/Haitian participants had significantly higher mean scores for the Surgency and Orienting/Regulation factors than White participants ($p < 0.005$ for all comparisons). Hispanic participants also had significantly higher mean scores for the Negative Affectivity factor than White participants ($p = 0.038$). Hispanic and Black/Haitian participants had significantly lower than expected scores on the Sadness scale than White participants ($p = 0.003$ and $p < 0.001$, respectively). Black/Haitian participants also had significantly lower than expected scores on the Falling Reactivity scale than White participants ($p = 0.015$). Hispanic and Black/Haitian participants did not significantly differ on their factor mean scores or in their association with Sadness and Falling Reactivity. These findings indicate that the three factor means and the intercept for the Sadness scale were not invariant between White and Hispanic participants and that the Surgency and Orienting/Regulation factor means and the intercepts for the Sadness and Falling Reactivity scales were not invariant between the Black/Haitian and White participants.

Annual household income ≥ $40,000 was associated with significantly lower mean scores for the Surgency and Orienting/Regulation factors than annual household income < $40,000 ($p = 0.012$ and $p = 0.011$, respectively). Higher income was also associated with significantly higher than expected scores on the High Intensity Pleasure and Low Intensity Pleasure scales ($p = 0.005$ and $p < 0.001$, respectively) and with a significantly lower than expected score on the Fear scale ($p = 0.025$), indicating that the Surgency and Orienting/Regulation factor means and the intercepts for the High Intensity Pleasure and Low Intensity Pleasure scales were not invariant between the income groups.

4. Discussion

The objectives of this study were to examine the factor structure of the IBQ-R in a culturally, sociodemographically diverse sample of infants living in the northeastern U.S and to determine if any of the IBQ-R factors or scales showed evidence of non-invariance across cultural and sociodemographic groups. Validating the factor structure of the IBQ-R in diverse samples is critical for several reasons. The IBQ-R is a widely used instrument, with recommended factors in use by researchers studying infant temperament. However, evidence suggests that the factor structure may be influenced by cultural and sociodemographic factors, and the currently recommended factors were based on a sample of primarily White, married, middle to high SES families. Thus, investigators utilizing the recommended IBQ-R factors without validating them within their own sample may be using a factor structure inappropriate for their sample. Given the current demographics of the U.S. population (e.g., 37% racial/ethnic minority, 13% foreign-born, 15% below poverty line; www.census.gov), confirming the validity of the recommended factor structure in diverse samples is necessary for the measure to have maximum utility in research studies going forward. This issue will become increasingly important as U.S. demographics continue to evolve.

Moreover, assumptions have been made about the fundamental nature of temperament and the developmental course of temperament/personality on the basis of factor analysis results of the IBQ-R and other temperament/personality measures (Clifford et al., 2013; Gartstein et al., 2003). Such assumptions may be flawed if the structure is not universal across cultural and sociodemographic groups. Different structures in infant temperament (and by extension,

later personality) across cultural and/or sociodemographic groups would suggest limits to the generalizability of theories regarding the structure of temperament and personality (Montirosso et al., 2011). Such differences would need to be considered in future studies of temperament and personality and theories of socioemotional development. Conversely, establishing that there is a consistent temperament structure across cultural and sociodemographic groups would contribute to unified definitions of temperament constructs that could be widely applied (Montirosso et al., 2011). Finally, differences in temperament structure across cultural and other sociodemographic factors may have clinical implications by laying the foundation for later psychopathology, including the manner in which clinical distress is manifested as emotional and behavioral symptoms and disorders (Gartstein et al., 2006, 2009). Thus, elucidating any differences in the factor structure of infant temperament across cultural and other sociodemographic groups may provide insight into group differences in the expression of mental illness and relative rates of various forms of psychopathology.

The current study's findings suggest some confirmation of the general factor structure originally reported by Gartstein and Rothbart (2003), but with some notable differences. In the current study, the best model fit identified three factors that showed similarities to those described by Gartstein and Rothbart, including a Surgency factor, a Negative Affectivity factor, and an Orienting/Regulation factor. However, unlike Gartstein and Rothbart's study, the Approach scale did not load on the Surgency factor. In fact, the Approach scale had to be excluded from analyses due to substantial residual covariation of the error term with a number of other variables. The Negative Affectivity factor consisted of the same scales as those identified by Gartstein and Rothbart as well as cross-loadings from the Vocal Reactivity and Activity Level scales. The Orienting/Regulation factor consisted of the Duration of Orientation, Low Intensity Pleasure, and Fear scales; the Fear scale, however, had a significant cross-loading with the Negative Affectivity factor. These patterns were not observed in the Gartstein and Rothbart factor structure. The Gartstein and Rothbart Orienting/Regulation factor also included the Cuddliness and Soothability scales. These scales were ultimately excluded in the current study due to their low factor loadings. Notably, these differences are similar to those found in other studies that have attempted to replicate the IBQ-R factors. For example, Gartstein et al. (2005) excluded Cuddliness from their factor analysis after determining it had a relatively low factor loading and an error term with substantial residual covariation with several other variables. Although Vonderlin et al. (2012) and Mink et al. (2013) were able to generally confirm the Surgency/Extraversion and Negative Affectivity factors with some changes (e.g., Activity Level loading on Negative Affectivity instead of Surgency/Extraversion), they could not replicate the Orienting/Regulation factor. The findings from this and previous studies suggest that the Cuddliness and Soothability scales may be less reliable contributors to underlying temperament factors. Interestingly, a number of studies that have examined cross-cultural differences in infant temperament have documented group differences on the mean scores of these scales (e.g., Cozzi et al., 2013; Dragan et al., 2011; Gartstein et al., 2006; Mink et al., 2013; Montirosso et al., 2011), suggesting that cultural differences in how these domains of behavior are encouraged and/or expressed may contribute to differences in the IBQ-R factor structure across studies.

Analyses also examined whether three cultural/sociodemographic characteristics— maternal country of birth (U.S. or not U.S.), race/ethnicity (White, Black/Haitian, Hispanic), and annual household income (greater or less than $40,000)—were associated with differences on factor and individual scale scores. The findings indicated that each of these characteristics was associated with differences on factor as well as individual scale scores, suggesting lack of invariance in these measures across groups. Specifically, differences were found for maternal country of birth on the Orienting/Regulation factor and the Smiling and Laughter and Activity Level scales; for maternal race/ethnicity for all three factors and the Sadness and Falling Reactivity scales; and for annual household income on the Surgency and Orienting/Regulation factors and the High Intensity Pleasure, Low Intensity Pleasure, and Fear scales.

This study is unique in a number of ways. The sample was diverse on several cultural and sociodemographic factors, including race and ethnicity, immigration and acculturation status, maternal marital status and education level, and household income. To date, IBQ-R factor analysis studies have been performed on largely White, middle to high SES families, which does not represent current U.S. demographics. Understanding the functioning of the IBQ-R in samples more similar to current U.S. demographics is critical for confirming its validity for use in diverse samples, which are becoming more common in developmental research. The study was based on a sample size comparable to or larger than most similar studies.

Notably, many prior studies included a mix of ages, most frequently 3- to 12-month-olds (e.g., Gartstein & Rothbart, 2003; Gartstein et al., 2005, 2009; Montirosso et al., 2011), despite evidence that temperament characteristics demonstrate change over development (Bornstein et al., 2015). These changes are hypothesized to occur due to rapid development in biological systems underlying temperament (e.g., neuroendocrine, brain structures) as well as the heightened malleability of these systems to environmental influences in early life (Bornstein et al., 2015; Gartstein & Rothbart, 2003). The majority of data suggest only moderate stability of temperament factors and scales across infancy (Bornstein et al., 2015). Different patterns of change have been documented for the various temperament domains, including increases in expressions of positive emotionality (e.g., smiling), activity level, approach, vocal reactivity, perceptual sensitivity, and distress to limitations between 3 and 12 months; increases in fear, particularly between 6 and 12 months; and decreases in duration of orienting between 6 and 9 months followed by either increases or further decreases between 9 and 12 months (Gartstein & Rothbart, 2003). Domains of temperament that are more heavily influenced by normative developmental growth and/or are more susceptible to environmental influences may show the least stability over time. For example, attentional and regulation strategies improve as the executive attention network for the prefrontal cortex and anterior cingulate regions of the brain become active (Bornstein et al., 2015; Rothbart et al., 2011). Quality of caregiving appears to influence the expression of negative temperamental domains (Bornstein et al., 2015). In addition, temperamental domains may influence each other; for example, higher orienting is associated with greater positive and less negative affect (Rothbart et al., 2011). Moreover, some studies have found interactive effects of age and culture on temperament characteristics (e.g., Montirosso et al., 2011). Thus, one strength of the current study is its focus on one age (6 months), as limiting

age likely reduced variability in findings due to developmental changes. However, because temperament in general changes over the first year of life and different domains experience different rates/patterns of change, the current findings may not generalize to younger or older infants. Additional study is needed to determine how the IBQ-R factor structure may be influenced by cultural and sociodemographic characteristics across infancy.

This study also has limitations. Due to the relatively small size of subgroups across cultural and sociodemographic variables, multiple-group analysis could not be conducted to test whether the IBQ-R factor structure found in the sample as a whole (i.e., factor means, loadings, variances and covariances, intercepts of manifest variables, and residual variances and covariances) was consistent across the various subgroups (e.g., White versus Black versus Hispanic, high versus low SES, U.S.-born versus non-U.S.-born). To evaluate group differences in light of these sample size constraints, a MIMIC model was utilized. The MIMIC model can assess group differences in factor means and in intercepts of manifest variables. Ideally, a multiple-group analysis (which can additionally assess group differences in factor variances/covariances, factor loadings, and observed residual variances/covariances) to evaluate non-invariance would have been employed, but this approach requires larger group subsamples than were available. Future studies with larger subsamples should directly test the specific impact of various cultural and sociodemographic factors on the IBQ-R factor structure. Also, the influence of infant sex was not examined. Notably, Gartstein et al. (2003) did not find differences between male and female infants on factor scores. Further, there has been some agreement in the literature that sex differences in temperament are relatively minor during the first year of life, with greater discrepancies emerging in later childhood (Montirosso et al., 2011). However, recent data suggest that there may be some sex differences observable in infancy (e.g., greater Surgency-based traits in males and greater Fear and Orienting/Regulatory capacity traits in females; Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006; Montirosso et al., 2011). Future studies should examine whether infant sex influences IBQ-R factor structure, particularly when other cultural and sociodemographic factors are taken into account, as there is some evidence for interactions between sex and culture on temperament characteristics (e.g., Montirosso et al., 2011). Finally, future research should consider how age, sex, and cultural and sociodemographic characteristics jointly interact to influence the structure of infant temperament.

4.1. Conclusions

This study is an initial step in confirming the validity of the commonly used IBQ-R factor structure in culturally and sociodemographically diverse samples. The current sample is more representative of current U.S. demographics than the samples used to date to identify the IBQ-R factor structure. Next steps in this line of research should include examining whether the factor structure is upheld within specific cultural and sociodemographic groups (e.g., specific racial/ethnic minority groups; low versus high SES), as any group differences may be obscured in analyses that combine participants from different cultural and sociodemographic backgrounds. Results from the current study suggest that cultural and sociodemographic characteristics such as maternal country of birth, race/ethnicity, and SES may influence IBQ-R ratings. Further study is also needed to determine if any differences in

factor structure are developmentally or clinically meaningful and predictive of later outcomes, such as personality or behavioral problems/psychopathology. Confirming the structure of infant temperament across cultural and sociodemographic groups is critical to the conduct of valid infant temperament research in diverse populations and for refining current theories of temperament and personality.

Acknowledgements

The authors thank the participating families, whose generation donation of time made the study possible.

Role of the Funding Source

The research was supported by a grant from the National Heart, Lung, and Blood Institute (R01HL095606; Wright, Bosquet Enlow); neurobehavioral phenotyping in PRISM is supported by P30 ES023515. During preparation of this manuscript, the authors were supported by R01HL095606 (Bosquet Enlow and Wright), the Program for Behavioral Science in the Department of Psychiatry at Boston Children's Hospital (Bosquet Enlow), and the Mindich Child Health & Development Institute at the Icahn School of Medicine at Mount Sinai (Wright). None of the funding sources had any role in the study design (with the exception of the grant review process), in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to submit the manuscript for publication.

Abbreviations

IBQ-R	Infant Behavior Questionnaire-Revised
EFA	exploratory factor analysis
CFA	confirmatory factor analysis
SES	socioeconomic status
GED	general equivalency diploma
App	Approach
VR	Vocal Reactivity
HP	High Intensity Pleasure
SL	Smiling and Laughter
Act	Activity Level
PS	Perceptual Sensitivity
Sad	Sadness
DL	Distress to Limitations
Fall	Falling Reactivity/Rate of Recovery from Distress
LP	Low Intensity Pleasure
Cud	Cuddliness
DO	Duration of Orienting

Sooth	Soothability
SUR	Surgency
NEG	Negative Affectivity
REG	Orienting/Regulation

References

Bornstein MH, Putnick DL, Gartstein MA, Hahn C, Auestad N, O'Connor DL. Infant temperament: stability by age, gender, birth order, term status, and socioeconomic status. Child Development. 2015; 86:844–863. doi:10.1111/cdev.12367. [PubMed: 25865034]

Brown, T. Confirmatory factor analysis for applied research. Guildford; New York: 2006.

Brunst KJ, Enlow MB, Kannan S, Carroll KN, Coull BA, Wright RJ. Effects of prenatal social stress and maternal dietary fatty acid ratio on infant temperament: Does race matter? Epidemiology (Sunnyvale). 2014; 4:167. doi:10.4172/2161-1165.1000167.

Clifford SM, Hudry K, Elsabbagh M, Charman T, Johnson MH, The BASIS Team. Temperament in the first 2 years of life in infants at high-risk for autism spectrum disorders. Journal of Autism and Developmental Disorders. 2013; 43:673–686. doi:10.1007/s10803-012-1612-y. [PubMed: 22918859]

Cozzi P, Putnam SP, Menesini E, Gartstein MA, Aureli T, Calussi P, et al. Studying cross-cultural differences in temperament in toddlerhood: United States of America (US) and Italy. Infant Behavior & Development. 2013; 36:480–483. http://dx.doi.org/10.1016/j.infbeh.2013.03.014. [PubMed: 23735346]

Derauf C, LaGasse L, Smith L, Newman E, Shah R, Arria A, et al. Infant temperament and high risk environment relate to behavior problems and language in toddlers. Journal of Developmental and Behavioral Pediatrics. 2011; 32:125–135. doi:10.1097/DBP.0b013e31820839d7. [PubMed: 21200329]

Dragan WL, Kmita G, Fronczyk K. Psychometric properties of the Polish adaptation of the Infant Behavior Questionnaire-Revised (IBQ-R). International Journal of Behavioral Development. 2011; 35:542–549. doi:10.1177/0165025411422181.

Else-Quest NM, Hyde JS, Goldsmith HH, Van Hulle CA. Gender differences in temperament: A meta-analysis. Psychological Bulletin. 2006; 132:33–72. doi:10.1097/0033-2909.132.1.33. [PubMed: 16435957]

Gartstein MA, Bridgett DJ, Rothbart MK, Robertson C, Iddins E, Ramsay K, et al. A latent growth examination of fear development in infancy: contributions of maternal depression and the risk for toddler anxiety. Developmental Psychology. 2010; 46:651–658. doi:10.1037/a0018898. [PubMed: 20438177]

Gartstein MA, Gonzalez C, Carranza JA, Ahadi SA, Ye R, Rothbart MK, et al. Studying cross-cultural differences in the development of infant temperament: People's Republic of China, the United States of America, and Spain. Child Psychiatry and Human Development. 2006; 37:145–161. doi: 10.1007/s10578-006-0025-6. [PubMed: 16874564]

Gartstein MA, Knyazev GG, Slobodskaya HR. Cross-cultural differences in the structure of infant temperament: United States of America (U.S.) and Russia. Infant Behavior & Development. 2005; 28:54–61. doi:10.1016/j.infbeh.2004.09.003.

Gartstein MA, Marmion J. Fear and positive affectivity in infancy: convergence/discrepancy between parent-report and laboratory-based indicators. Infant Behavior & Development. 2008; 31:227–238. doi:10.1016/j.infbeh.2007.10.012. [PubMed: 18082892]

Gartstein MA, Peleg Y, Young BN, Slobodskaya HR. Infant temperament in Russia, United States of America, and Israel: differences and similarities between Russian-speaking families. Child Psychiatry and Human Development. 2009; 40:241–256. doi:10.1007/s10578-009-0123-3. [PubMed: 19165594]

Gartstein MA, Rothbart MK. Studying infant temperament via the Revised Infant Behavior Questionnaire. Infant Behavior & Development. 2003; 26:64–86. doi:10.1016/S0163-6383(02)00169-8.

Gartstein MA, Slobodskaya HR, Kinsht LA. Cross-cultural differences in temperament in the first year of life: United States of America (U.S.) and Russia. International Journal of Behavioral Development. 2003; 27:316–328.

Gartstein MA, Slobodskaya HR, Żylicz P, Gosztyla D, Nakagawa A. A cross-cultural evaluation of temperament: Japan, USA, Poland, and Russia. International Journal of Psychology & Psychological Therapy. 2010; 10:55–75.

Goldsmith HH, Campos JJ. The structure of temperamental fear and pleasure in infants: a psychometric perspective. Child Development. 1990; 61:1944–1964. [PubMed: 2083507]

Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal. 1999; 6:1–55.

Lengua LJ, Honorado E, Bush NR. Contextual risk and parenting as predictors of effortful control and social competence in preschool children. Journal of Applied Developmental Psychology. 2007; 28:40–55. doi:10.1016/j.appdev.2006.10.001. [PubMed: 21687825]

Mink D, Henning A, Aschersleben G. Visual habituation tasks: the role of infant temperament. Infant Behavior & Development. 2013; 36:377–390. http://dx.doi.org/10.1016/j.infbeh.2013.03.005. [PubMed: 23598166]

Montirosso R, Cozzi P, Putnam SP, Gartstein MA, Borgatti R. Studying cross-cultural differences in temperament in the first year of life: United States and Italy. International Journal of Behavioral Development. 2011; 35:27–27. doi:10.1177/0165025410368944.

Parade SH, Leerkes EM. The reliability and valdity of the Infant Behavior Questionnaire-Revised. Infant Behavior & Development. 2008; 31:637–646. doi:10.1016/jinfbeh.2008.07.009. [PubMed: 18804873]

Patra J, Bakker R, Irving H, Jaddoe VW, Malini S, Rehm J. Dose-response relationship between alcohol consumption before and during pregnancy and the risks of low birthweight, preterm birth and small for gestational age (SGA)-a systematic review and meta-analyses. British Journal of Obstetrics and Gynaecology. 2011; 118:1411–1421. doi:10.1111/j.1471-0528.2011.03050.x. [PubMed: 21729235]

Putnam SP, Rothbart MK, Gartstein MA. Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. Infant and Child Development. 2008; 17:387–405.

Rothbart MK, Ahadi SA, Hershey KL, Fisher P. Investigations of temperament at three to seven years: the Children's Behavior Questionnaire. Child Development. 2001; 72:1394–1408. [PubMed: 11699677]

Rothbart MK, Sheese BE, Rueda R, Posner MI. Developing mechanisms of self-regulation in early life. Emotion Review. 2011; 3:207–213. doi:10.1177/1754073910387943. [PubMed: 21892360]

Super CM, Harkness S. The developmental niche: a conceptualization at the interface of child and culture. International Journal of Behavioral Development. 1986; 9:545–569.

Testa M, Quigley BM, Eiden RD. The effects of prenatal alcohol exposure on infant mental development: A meta-analytical review. Alcohol and Alcoholism. 2003; 38:295–304. doi:http//dx.doi.org.ezp-prod1.hul.harvard.edu/10.1093/alcalc/agg087. [PubMed: 12814894]

Tester-Jones M, O'Mahen H, Watkins E, Karl A. The impact of maternal characteristics, infant temperament and contextual factors on maternal responsiveness to infant. Infant Behavior & Development. 2015; 40:1–11. http://dx.doi.org/10.1016/j.infbeh.2015.02.014. [PubMed: 25913568]

Tikotzky L, Chambers AS, Gaylor E, Manber R. Maternal sleep and depressive symptoms: links with infant Negative Affectivity. Infant Behavior & Development. 2010; 33:605–612. doi:10.1016/j.infbeh.2010.07.012. [PubMed: 20723998]

van den Heuvel MI, Johannes MA, Henrichs J, Van den Bergh BRH. Maternal mindfulness during pregnancy and infant socio-emotional development and temperament: the mediating role of

maternal anxiety. Early Human Development. 2015; 91:103–108. http://dx.doi.org/10.1016/j.earlhumdev.2014.12.003. [PubMed: 25577496]

Ventura JN, Stevenson MB. Relations of mothers' and fathers' reports of infant temperament, parents' psychological functioning, and family characteristics. Merrill-Palmer Quarterly. 1986; 32:275–289.

Vonderlin E, Ropeter A, Pauen S. [Assessment of temperament with the Infant Behavior Questionnaire Revised (IBQ-R) – the psychometric properties of a German version. Zeitschrift für Kinder- und Jugendpsychiatrie. 2012; 40:307–314. doi:10.1024/1422-4917/a000187.

Highlights

- The Infant Behavior Questionnaire (IBQ-R) is a widely used temperament measure.
- Recommended IBQ-R factors were derived in a sociodemographically homogenous sample.
- The factor structure needs to be confirmed in sociodemographically diverse samples.
- This study identified recommended factors, with adjustments, in a diverse sample.
- More research is needed to ensure maximal validity of the IBQ-R in diverse samples.

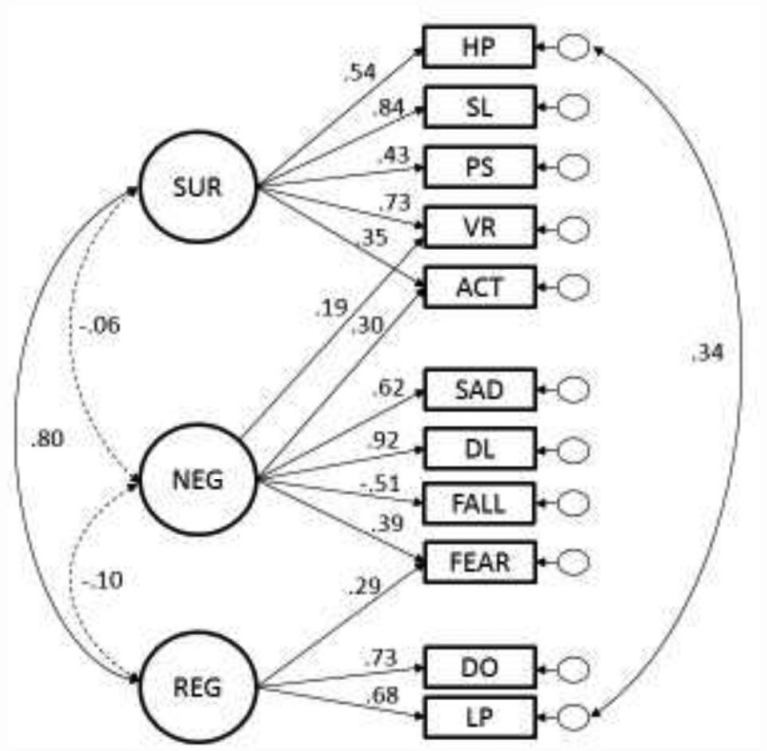

Figure 1.

Exploratory factor analysis model of the IBQ-R structure in the current sample ($N = 380$). Standardized parameter estimates are presented. Solid lines indicate significant paths ($p < .05$) and dashed lines non-significant paths ($p \geq .05$). SUR = Surgency; NEG = Negative Affectivity; REG = Orienting/Regulation; VR = Vocal Reactivity; HP = High Intensity Pleasure; SL = Smiling and Laughter; ACT = Activity Level; PS = Perceptual Sensitivity; SAD = Sadness; DL = Distress to Limitations; FALL = Falling Reactivity/Rate of Recovery from Distress; LP = Low Intensity Pleasure; DO = Duration of Orienting.

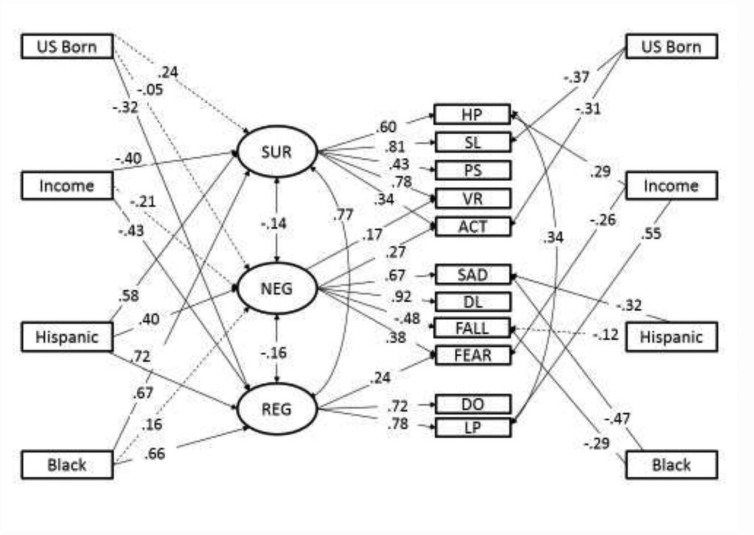

Figure 2.

Multiple Indicator Multiple Causes (MIMIC) model examining the impact of cultural and sociodemographic covariates on the three latent factors and on the manifest variables. Standardized parameter estimates are presented. Solid lines indicate significant paths ($p < .05$) and dashed lines non-significant paths ($p \geq .05$). All covariates are binary. Maternal country of birth was dichotomized into U.S.-born and non-U.S.-born (reference group). Annual household income was dichotomized using a median split as ≥ $40,000 versus < $40,000 (reference group). For race/ethnicity, White was treated as the reference group; participants reporting a race/ethnicity other than White, Black/Haitian, or Hispanic were excluded. SUR = Surgency; NEG = Negative Affectivity; REG = Orienting/Regulation; VR = Vocal Reactivity; HP = High Intensity Pleasure; SL = Smiling and Laughter; ACT = Activity Level; PS = Perceptual Sensitivity; SAD = Sadness; DL = Distress to Limitations; FALL = Falling Reactivity/Rate of Recovery from Distress; LP = Low Intensity Pleasure; DO = Duration of Orienting.

Table 1

IBQ-R Scales

IBQ-R Scale	Scale Description
Approach	Rapid approach, excitement, and positive anticipation of pleasurable activities
Vocal Reactivity	Amount of vocalization exhibited in daily activities
High Intensity Pleasure	Amount of pleasure or enjoyment related to high stimulus intensity, rate, complexity, novelty, and incongruity
Smiling and Laughter	Smiling or laughter in general caregiving and play situations
Activity Level	Movement of arms and legs, squirming, and locomotor activity
Perceptual Sensitivity	Amount of detection of slight, low intensity stimuli from external environment
Sadness	General low mood; lowered mood and activity related to personal suffering, physical state, object loss, or inability to perform a desired action
Distress to Limitations	Fussing, crying, or showing distress when in a confining place or position, involved in caregiving activities, or unable to perform a desired action
Fear	Startle or distress to sudden changes in stimulation or novel physical objects or social stimuli; inhibited approach to novelty
Falling Reactivity/Rate of Recovery from Distress	Rate of recovery from peak distress, excitement, or general arousal; ease of falling asleep
Low Intensity Pleasure	Amount of pleasure or enjoyment related to low stimulus intensity, rate, complexity, novelty, and incongruity
Cuddliness	Expression of enjoyment and molding of body to being held by caregiver
Duration of Orienting	Attention to and/or interaction with a single object for extended periods of time
Soothability	Reduction of fussing, crying, or distress when caregiver uses soothing techniques

Table 2

Sample Cultural and Sociodemographic Characteristics (N = 380 Dyads)

	N	%	*M*	*SD*
Maternal age (years)			30.2	5.8
Maternal relationship status				
Married	194	51.0		
Living with partner	94	24.7		
Single	50	13.2		
Divorced	6	1.6		
Separated	4	< 1		
Other/not reported	32	8.4		
Maternal educational attainment				
Not completed high school	81	21.3		
High school diploma/GED	46	12.1		
Associate's degree/some college	89	23.4		
College degree	74	19.5		
Graduate degree	82	21.6		
Not reported	8	2.1		
Annual household income				
< $10,000	51	13.4		
$10,000-$20,000	59	15.5		
$20,000-$35,000	55	14.5		
$35,000-$50,000	55	14.5		
$50,000-$70,000	20	5.3		
$70,000-$100,000	31	8.2		
> $100,000	88	23.2		
Not reported	21	5.5		
Maternal race/ethnicity				
Hispanic	132	34.7		
Black/Haitian	116	30.5		
White	107	28.2		
Other	25	6.6		
Primary language in home				
English	277	72.9		
Spanish	80	21.1		
Other	23	6.1		
Mother born outside U.S.	155	40.8		
Maternal region of birth				
Central America	63	16.6		
Caribbean	29	7.6		

	N	%	*M*	*SD*
South America	27	7.1		
Africa	9	2.4		
Eastern Europe	7	1.8		
Western Europe	5	1.3		
South/Southeast Asia	5	1.3		
East Asia	4	1.1		
Canada	4	1.1		
Other	2	0.5		
Maternal age when moved to U.S.				
< 5 years old	11	2.9		
5 to 10 years old	22	5.8		
11 to 17 years old	22	5.8		
18 to 25 years old	63	16.6		
> 25 years old	34	8.9		
Not reported	3	0.8		
Length of time in U.S.				
< 1 year	6	1.6		
1 to 2 years	6	1.6		
3 to 5 years	18	4.7		
> 5 years	124	32.6		
Not reported	1	0.3		
Primiparous	93	24.5		
Infant age (weeks)			26.7	2.5
Infant % male	203	53.4		

Infant Behav Dev. Author manuscript; available in PMC 2017 May 01.

Table 3

Spearman Correlation Coefficients and Means, Standard Deviations, and Cronbach Alphas for IBQ-R Scale Scores in Current Sample

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. App	-													
2. VR	.40***	-												
3. HP	.37***	.40***	-											
4. SL	.35***	.63***	.47***	-										
5. Act	.23***	.29***	.18***	.27***	-									
6. PS	.34***	.37***	.30***	.33***	.16**	-								
7. Sad	−.02	−.03	−.08	−.14**	.10*	.01	-							
8. DL	−.02	.13*	−.06	−.02	.25***	.01	.54***	-						
9. Fear	−.11*	.17**	−.07	.19***	.19***	.13*	.22***	.31***	-					
10. Fall	.17**	−.05	.07	.06	−.05	.06	−.27***	−.44***	−.14**	-				
11. LP	.33***	.43***	.52***	.48***	.11*	.26***	−.08	−.08	.08	.11*	-			
12. Cud	.09+	.09+	.23***	.06	−.18***	.06	−.08	−.17**	−.15**	.16**	.28***	-		
13. DO	.21***	.42***	.30***	.51***	.15**	.24***	−.12*	−.05	.21***	.00	.47***	.05	-	
14. Sooth	.19***	.09+	.19***	.09+	.02	.07	−.15**	−.20***	−.11*	.34***	.18**	.22***	.05	-
Mean	5.53	4.78	6.13	5.34	4.64	4.40	3.25	3.59	2.55	5.16	5.64	5.86	4.53	5.31
SD	0.91	1.03	0.70	1.02	0.80	1.22	0.92	0.90	1.09	0.90	0.79	0.63	1.16	0.77
Cronbach's α	.79	.73	.76	.80	.71	.81	.76	.78	.91	.81	.73	.74	.81	.80

Note. App = Approach; VR = Vocal Reactivity; HP = High Intensity Pleasure; SL = Smiling and Laughter; Act = Activity Level; PS = Perceptual Sensitivity; Sad = Sadness; DL = Distress to Limitations; Fall = Falling Reactivity/Rate of Recovery from Distress; LP = Low Intensity Pleasure; Cud = Cuddliness; DO = Duration of Orienting; Sooth = Soothability.

$^{+}p < .10$.

$^{*}p < .05$.

$^{**}p < .01$.

$^{***}p < .001$.

Table 4

Fit Statistics for All Models Tested

Model	d.f.	χ^2	TLI	CFI	RMSEA	90% CI
Gartstein et al. (2003; *N* = 360)	74	302.62	0.76	0.81	0.090	0.080-0.101
Gartstein et al. (2003), excluding Cuddliness (*N* = 360)	62	245.73	0.79	0.84	0.088	0.077-0.100
Gartstein et al. (2005), California subsample (*N* = 229)	58	177.93	0.86	0.89	0.074	0.062-0.086
Gartstein et al. (2005), California + Oregon subsamples (*N* = 608)	42	110.49	0.89	0.93	0.066	0.051-0.081
EFA Model (*N* = 380, Figure 1)	37	62.37	0.96	0.97	0.042	0.023-0.060
MIMIC Model (*N* = 336, Figure 2)	60	66.28	0.99	0.99	0.018	0-0.039

Note. TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = root mean square error of approximation; CI = confidence interval; EFA = exploratory factor analysis; MIMIC = Multiple Indicator Multiple Causes.