HHS Public Access

Author manuscript

Child Dev. Author manuscript; available in PMC 2016 May 01.

Published in final edited form as:

Child Dev. 2015 May ; 86(3): 844–863. doi:10.1111/cdev.12367.

Infant Temperament: Stability by Age, Gender, Birth Order, Term Status, and SES

Marc H. Bornstein,
Eunice Kennedy Shriver National Institute of Child Health and Human Development

Diane L. Putnick,
Eunice Kennedy Shriver National Institute of Child Health and Human Development

Maria A. Gartstein,
Washington State University

Chun-Shin Hahn,
Eunice Kennedy Shriver National Institute of Child Health and Human Development

Nancy Auestad, and
Dairy Research Institute

Deborah L. O'Connor
University of Toronto and The Hospital for Sick Children

Abstract

Two complementary studies focused on stability of infant temperament across the first year and considered infant age, gender, birth order, term status, and socioeconomic status (SES) as moderators. Study 1 consisted of 73 mothers of firstborn term girls and boys queried at 2, 5, and 13 months of age. Study 2 consisted of 335 mothers of infants of different gender, birth order, term status, and SES queried at 6 and 12 months. Consistent positive and negative affectivity factors emerged at all time-points across both studies. Infant temperament proved stable and robust across gender, birth order, term status, and SES. Stability coefficients for temperament factors and scales were medium to large for shorter (<9 months) inter-assessment intervals and small to medium for longer (>10 months) intervals.

Developmental Stability and Its Significance

Development is traditionally and commonly identified with transformation and change (McCall, 1986), but not all constructs, structures, functions, or processes alter in development, and development equally includes consistency over time. Two kinds of developmental consistency have been distinguished: individual-order consistency (stability) and group mean-level consistency (continuity). Order and level consistency are both developmentally informative and can co-exist conceptually and empirically as the two are

Address correspondence to: Marc H. Bornstein, Child and Family Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Suite 8030, 6705 Rockledge Drive, Bethesda MD 20892-7971, U.S.A. Marc_H_Bornstein@nih.gov.

independent (Bornstein & Bornstein, 2008). Many features of human development remain (more or less) consistent over time. Somewhat paradoxically, consistency is more parsimonious and orderly than is transformation. The two studies presented here are concerned with individual-order consistency—stability— in the early ontogeny of infant temperament.

The study of stability is central to developmental science for several reasons. First, certain biological systems in humans require stability – physical, chemical, physiological, psychological, and environmental – to survive. (Temperament is thought to have genetic and biological roots.) Second, stability provides basic information about the overall ontogenetic course of a characteristic, insofar as individuals do or do not maintain their standing relative to others in their cohort over time. Stability therefore provides links between our practical understanding of individuals when younger and older. (Is the difficult infant the troubled child?) Third, it is generally assumed that to be psychometrically as well as psychologically meaningful, a characteristic should be consistent across time, allowing antecedent assessments to serve as predictors of subsequent indicators of the same or different characteristics. (It is the presumably stable temperament that develops into personality.) Fourth, stability affects the social and physical environment and, so in turn, development; from the perspective of so-called evocative interactions, stable characteristics elicit responses that contribute to shaping later outcomes. Individuals with different stable characteristics also likely experience environments differently. (Consistently easy versus difficult infants likely have very different childhoods.) Relative to change, biological forces tend to reinforce homeostasis in the individual, and the consistent social network in which much of development normally transpires also contributes to constancy. Thus, individual ←→ environmental relational processes tilt toward stability. For these many reasons, stability has been a topic of theory and research across the history of developmental science (Lerner, Hershberg, Hilliard, & Johnson, 2015).

However, stability is not static, but dynamic, and many factors are acknowledged to moderate stability, including the characteristic studied (some characteristics are thought to be more stable than are others), developmental stage (a given characteristic may not be stable at one point in the life span, but stabilize at another), the nature of assessment (the same measure applied at different times yields higher stability estimates, whereas different measures yield lower stability estimates), the temporal interval between assessments (the shorter the inter-assessment interval, the greater the stability estimate), and the assessment context (consistent settings promote higher stability, and inconsistent ones attenuate stability). Temporal stability of infant temperament therefore might be expected to vary depending on the particular dimension of infant temperament studied, the age of the infants and the period of time over which stability is assessed, the measures of infant temperament that are employed, the consistency of settings in which infants are assessed, and a variety of other influences. In short, the stability of temperament is contingent, not absolute, underscoring the need to study its sensitivity to moderators, such as age, gender, birth order, term status, and SES, which we do here.

Infant Temperament: Structure and Stability

Although temperament is construed in multiple ways (see Bates & Pettit, 2015; Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006; Rothbart, 2011; Shiner et al., 2012), temperament in infancy is consensually defined as constitutionally based biological individual differences in attentional, motor, and emotional reactivity and self-regulation that are exhibited in different contexts in response to stimulation (Chen & Schmidt, 2015; Rothbart, 2011). Besides being biologically linked and apparent early in life, infant temperament is further characterized by two central features: first, temperament is manifest in observable individual behaviors that are structured into (at least) positive and negative higher-order factors, and, second, temperament is often hypothesized to be relatively stable.

Structure

With respect to the first property of temperament, structure, virtually all operationalizations of infant temperament rely on discrete, observable behavioral items, which are organized into scales (facets), which in turn aggregate to form overarching factors (Carey & McDevitt, 1978a; Chen & Schmidt, 2015; Gartstein & Rothbart, 2003; Riese, 1987a, 1987b; Rothbart, 1981, 1986; Thomas & Chess, 1977). As Shiner et al. (2012, p. 438) observed, “particular facets of temperament tend to covary, and the covariation among those traits is accounted for by higher order factors with greater breadth.” Within the classic Thomas, Chess, et al. (1963) model, for example, infant temperament items combine to form “difficult”, “easygoing”, “slow-to-warm-up”, and “average” temperaments. Analyses of a variety of infant temperament instruments since have resulted in more formal factor structures, as for the Revised Infant Temperament Questionnaire (RITQ; Carey & McDevitt, 1978a; McDevitt & Carey, 1981), the Infant Characteristic Questionnaire (ICQ; Bates, Freeland, & Lounsbury, 1979; Garcia Coll, Halpern, Vohr, Seifer, & Oh, 1992), and the Infant Temperament Measure (ITM; Bornstein, Gaughran, & Seguí, 1991).

Rothbart (1981, 2011; Rothbart & Derryberry, 1981) developed the Infant Behavior Questionnaire (IBQ), perhaps the most widely adopted parent-report infant temperament measure (see Gartstein, Bridgett, & Low, 2012), and the one we use here. The IBQ consists of 94 items that composite into 6 scales: *Activity Level*, gross motor activity; *Smiling and Laughter*, positive arousal under relatively quiet conditions; *Fear*, distress and/or extended latency to approach a novel stimulus; *Distress to Limitations*, reactions to frustrating conditions; *Duration of Orienting*, sustained attention in periods when there is no change in stimulation; and *Soothability*, reduction of fussing, crying, or distress to calming. Factor analyses of the IBQ have yielded overarching factors related to Positive Affectivity and Negative Affectivity (Kochanska, Coy, Tjebkes, & Husarek, 1998; Rothbart, 1986) extracted for mothers’ (and fathers’) ratings of infants from 3 to 12 months. Smiling and Laughter, Duration of Orienting, and Soothability generally comprise the Positive Affectivity factor, Distress to Limitations and Fear comprise the Negative Affectivity factor, and Activity level loads on either or both factors depending on the sample being analyzed (Carranza Carnicero, Pérez-López, Salinas-Gonzalez, & Martínez-Fuentes, 2000; Kochanska et al., 1998). Much of contemporary research concerning infant temperament is based on Rothbart’s theory and IBQ measures of the structure of infant temperament

(Carranza Carnicero et al., 2000; Gartstein, Knyazev, & Slobodskaya, 2005; Gartstein & Rothbart, 2003; Kochanska et al., 1998; Putnam, Gartstein, & Rothbart, 2006; Putnam, Rothbart, & Gartstein, 2008; Rothbart, 2011).

Stability

With respect to the second property, temporal stability, emerging scientific opinion about the development of temperament is in flux. On the one hand, as indicated, stability has traditionally been a constituent of the definition of temperament especially in the early years of life (Buss, 1989; Rothbart, 2011; Strelau, 1989; Thomas et al., 1963; Wachs & Kohnstamm, 2001). Traits that are present in infancy but not stable (or do not have an apparent enduring impact on later development) are not usually considered to be a part of temperament (Buss & Plomin, 1984). Empirically, stabilities in varying degrees have been reported for infant temperament (Bornstein et al., 1991; Carey & McDevitt, 1978b; Carranza, González-Salinas, & Ato, 2013; McDevitt & Carey, 1981; Plomin et al., 1993). As the present studies use the IBQ, most germane is Rothbart (1986) who examined infant temperament with the IBQ at 3, 6, and 9 months. She reported stability of the positivity factor across both 3- and 6-month intervals and stability of the negativity factor and overall reactivity across 3-month intervals. Later IBQ stability reports included Worobey and Blajda (1989) from 2 weeks to 2 months and from 2 months to 12 months; Carranza Carnicero et al. (2000) across the first year of life; and Putnam et al. (2006) from 6 to 12 months and from 18 to 36 months (Putnam et al., 2008). Komsi et al. (2006, 2008), in turn, examined the stability of temperament based on both mother- and father-rated infant and child temperament (assessed with the IBQ and CBQ, respectively); like Putnam et al. (2008), they found evidence for stability of temperament from infancy to middle childhood.

On the other hand, change is rapid and thoroughgoing in the first years of life (Bornstein, Arterberry, & Lamb, 2014), and fresh theoretical, methodological, and statistical perspectives contend that temperament is not fixed across infancy. The fact that features of temperament are biologically or even genetically based does not automatically mean that they are immutable or that experience has little impact on temperament. First, biological systems that control temperament (neurohormonal and brain structure) emerge across infancy, and as they come online they may alter the expression and consequently the stability of temperament (Shiner et al., 2012). For example, stability of motor activity may be attributed to genetic factors, but new genetic effects may be observed as development proceeds (Saudino, 2012). Developing biological mechanisms also signal change with respect to attention. Posner, Rothbart, Sheese, and Voelker (2012) described a change from cholinergic to dopaminergic modulation as a marker of the shift from the primary role of the orienting network early in infancy to the reliance on the executive attention network of the prefrontal cortex and anterior cingulate regions of the brain (Rothbart, Derryberry, & Posner, 1994; Rueda, 2012) that comes online at the end of the first year of life, affording more flexible attention and regulation strategies. Thus, genetic and neurobiological factors might contribute to both stability and instability in how temperament is manifest (Saudino & Wang, 2012).

Second, temperament is now understood to be open to exogenous influences (Rothbart, 2011). For example, contextualist models focus on the roles of environment and experience in the formation and expression of temperament (Goldsmith, Buss, & Lemery, 1997; Rothbart, 2011). Higher levels of sensitive, warm parenting predict decreased child negative reactivity, even controlling for initial levels (Bates et al., 2012), and higher levels of harsh control predict increased child negative reactivity, even controlling for initial levels (Braungart-Rieker, Hill-Soderlund, & Karrass, 2010). Low parental sensitive responsiveness predicts increased fear, controlling for initial levels (Pauli-Pott et al., 2004), and the inverse (Park, Belsky, Putnam, & Crnic, 1997). Thus, one might expect changing temperament due to both maturation of self-regulation and experience.

Third, methodological issues impel a revisionist understanding of the stability of temperament. It has been common in temperament research to use different instruments in different studies, different instruments across different ages in the same study, and different levels of aggregation of temperament, thereby compromising the assessment of stability of infant temperament per se. Moreover, stability is conventionally indexed by (Pearson) correlation, and even large time-1 to time-2 correlations of, say, .50, leave unaccounted 75% of the shared developmental variance in the temperament measure. Individuals' rank-order positions on temperament indicators change detectably, and relatively low to moderate correlations between them over time may indicate stability but also accord with instability. It is also possible that temperament is less than stable early in life but stabilizes more as infancy concludes (Caspi et al., 2003; Pfeifer, Goldsmith, Davidson, & Rickman, 2002; Roberts & DelVecchio, 2000).

In consequence, some contemporary theories and models of temperament have begun to emphasize variation in expectations about the stability of infant temperament. Not surprisingly, recent studies using growth modeling techniques have confirmed that IBQ measurement is sensitive to developmental changes in infancy (Bridgett et al., 2009, 2011; Gartstein et al., 2010). Temperament appears to represent individual attentional, emotional, and behavioral characteristics of reactivity and regulation that are structured and relatively stable over time. However, temperament is influenced by biology and experience and interacts with the environment, and our understanding of its development is constrained by methodological choices. Together, these developments compel further unified study of the early structure and stability of temperament and its susceptibility or robustness to multiple moderation.

Moderation of the Structure and Stability of Infant Temperament

A related motive for a broader re-assessment of structure and stability of temperament in infancy is their possible sensitivity to moderation. Most reports of moderating factors in infant temperament research, such as those addressed here (i.e., age, gender, birth order, term status, and SES), have focused on mean levels of temperament (main effects) between groups at particular ages, a *nondevelopmental* approach. For example, the question of gender differences in infant temperament is among those most fundamental in gender research, but long-standing focus has almost exclusively fallen, not on relative stability or change of temperament in girls and boys, but on girl-boy mean-level differences in facets of

temperament (Casalin, Luyten, Vliegen, & Meurs, 2012; Else-Quest et al., 2006; Gartstein et al., 2006, 2010; Montirosso et al., 2011). Similarly, mean-level facet comparisons by birth order (Bates, 1987; Honjo et al., 1998; Keresteš, 2006) and term status (Goldberg & DiVitto, 2002; Oberklaid, Prior, & Sanson, 1986; Riese, 1987a, 1987b) have heretofore figured prominently in the infant temperament literature.

In this study, we focus on how age, gender, birth order, term status, and SES moderate *developmental stability* of temperament across infancy. Therefore, we only review mean-level group-comparison literatures as indicative of main effects that may have implications for group moderation of stability. On the one hand, strictly statistically speaking, stability of two groups (their slopes) is not tied to mean differences between the groups (their intercepts). On the other hand, it could be psychologically that a higher or lower level in one group at the start (the intercept) instigates evocative effects that maintain or discourage stability (the slope) over time. This argument applies to each moderator, and so some consideration of main effects is warranted. To the degree that patterns of stability are the same or differ according to infant age, gender, birth order, term status, and SES, failing to scrutinize stability by these moderators has resulted in a failure to fully understand the ontogeny of infant temperament.

Stability of Infant Temperament X Gender

Is temperament stable across infancy, and is it equivalently stable in infant girls and boys? Infant girls and boys are biologically dissimilar, and they are often (but not always) thought about and treated dissimilarly from the start of life (Bornstein, 2013). Mondschein, Adolph, and Tamis-LeMonda (2000) found that mothers of 11-month-old boys overestimated their babies' motor ability, whereas mothers of 11-month-old girls underestimated their babies, even when subsequent tests of ability revealed no differences between the two groups. Also, mothers may be closer to and more positively involved with their girls than boys (Bornstein et al., 2008; Clarke-Stewart, 1973), and parents show higher levels of sensitive, warm, and nonintrusive behavior toward their girls than toward their boys (Atzaba-Poria & Pike, 2008; Barnett, Deng, Mills-Koonce, Willoughby, & Cox, 2008; Bornstein et al., 2008; Lovas, 2005). Furthermore, parents respond differentially to their daughters and sons (Leaper, 2002), and mothers even respond differently to their "difficult" daughters than to their "difficult" sons, although the girls and boys may not differ on an independent assessment of difficulty (Else-Quest et al., 2006). On these accounts, then, the ontogenetic trajectories of temperament could follow different courses in infant girls and boys. However, there is evidence that girls and boys, and parenting girls and boys, are also surprisingly similar in many respects (Bornstein et al., 2015; Hyde, 2014; Leaper, 2002), and temperaments in girls and boys normally show wide and overlapping distributions, and so the ontogenies of temperament may not differ in the two genders. In the only stability study (to our knowledge) that specifically considered gender, Garcia Coll et al. (1992) assessed temperament via maternal report at 3 months and behaviorally at 3 and 7 months; they reported stability correlations for irritability for girls and for sociability and soothability for boys. We therefore hypothesized that stability of some structures of temperament in infancy might be moderated by infant gender.

Stability of Infant Temperament X Birth Order

If temperament is stable across infancy, is it equivalently stable in infant first- and laterborns? It has been argued on evolutionary grounds that the same parents might maximize fitness by producing different types of children (Ellis, Boyce, Belsky, Bakermans-Kranenburg, & van IJzendoorn, 2011). Parents also often treat children in the same family differently because those children differ in age, gender, cognition, or other idiosyncratic characteristics or life events (Furman & Lanthier, 2002; Hallers-Haalboom et al., 2014), and parenting does not affect different children in the same way (Suitor et al., 2009). One sibling might elicit more positivity and less negativity than another (Jenkins, Rasbash, & O'Connor, 2003), and siblings' different experiences (their nonshared real or perceived environments) contribute to making them temperamentally distinctive (Stoolmiller, 1999). Moreover, the presence of a sibling may alter the trajectory of either or both the older or younger sibling's temperament, as siblings engender more diverse family experiences than singletons which might affect stability. We were not able to locate any studies that compared and contrasted stability of temperament in first- versus laterborn infants per se. As a default, however, we hypothesized that first- and laterborn infants would be only moderately and equivalently stable in their temperaments.

Stability of Infant Temperament X Term Status

If temperament is stable across infancy, is it equivalently stable in preterm and term infants? Normal gestations of 38 to 42 weeks from conception are considered term (Howson, Kinney, & Lawn, 2012). However, among the roughly 4 million new births each year in the United States, approximately 12% of infants are born preterm (Martin, Hamilton, Ventura, Osterman, & Matthews, 2013). The earlier an infant is born, the more underdeveloped or medically compromised the preterm is likely to be (Behrman & Butler, 2006). Many preterm infants experience regulation difficulties, and prematurity is associated with increased negative emotionality and behavior problems (Goldberg & DiVitto, 2002; Hwang, Soong, & Liao, 2009). Thus, prematurity might disrupt individual consistency in infants, but some debate surrounds the term status moderation of stability correlations of temperament because extant research provides evidence of stability with respect to certain temperament structures but not others. Tronick, Scanlon, and Scanlon (1985) assessed preterm newborns' emotional and behavioral repertoire, utilizing the Brazelton (1973) Neonatal Behavioral Assessment Scales, twice 2 weeks apart and reported moderate to high stability correlations only for orienting, state regulation, and motor activation. Riese (1987a, 1987b) reported differential temperament stability from birth to 2 years in preterm and term infants, and later (Riese, 1988) reported that degree of prematurity affects the stability of temperament from 12 to 24 months. Garcia Coll et al. (1992) reported that patterns of temporal correlations interacted with temperament structure and intraventricular haemorrhage. Korner (1996) evaluated weekly sensory responsivity and excitability reactions in neonates born preterm and concluded that preterm might be more stable than term infants for some temperament dimensions. In a subsequent comparison of small-for-gestational-age (SGA) and appropriate-for-gestational-age (AGA) infants, few group differences emerged in temperament (Halpern, Garcia Coll, Meyer, & Bendersky, 2001), but, suggestively, when temperament was studied across the first year of life, AGA and SGA infants exhibited distinct patterns of temperament stability and instability (Halpern & Garcia Coll, 2000). We

therefore hypothesized that some but not all dimensions of temperament might be stable in preterm and term infants.

Stability of Infant Temperament X SES

If temperament is stable across infancy, is it equivalently stable in infants from different socioeconomic classes? Four lines of argumentation suggest that family SES might moderate the stability of infant temperament. First, genetically influenced traits, such as temperament is thought to be, can be expressed differently in different environments (Rowe, Jacobson, & Van den Oord, 1999). Socioeconomic class is related to aspects of child temperament, such as self-control; children from economically poor families are sometimes rated lower on self-control than children from economically well-to-do families (Chen & Schmidt, 2015). Second, under conditions of poverty, stress and neglectful and negative parenting elevate levels of glucocorticoids, fearfulness, and vigilance in children (perhaps preparing children for adverse experiences; Evans, 2004; Huizink, 2012). In this connection, Paulussen-Hoogeboom, Stams, Hermanns, and Peetsma (2007) learned that family SES moderates the effects of child negative emotionality on parental support: that is, higher-SES families may be more knowledgeable about child development and are likely to have more resources to buffer stress, thus buffering their children as well. Third, if the stability of infant temperament is carried by stability in the infant's environment, and different SES environments are differentially stable (Bradley & Corwyn, 2002), then SES could moderate the stability of temperament. Finally, parents from low-income households or with low-education levels might be less accurate in providing information on parental checklists (see, e.g., Roberts, Burchinal, & Durham, 1999) or high-SES parents may perceive and respond to children impatiently because of the relatively high opportunity cost of child care (Nelson, Kushlev, & Lyubomirsky, 2014). To our knowledge, SES moderation of the stability of infant temperament has not previously been examined. Nonetheless, on the basis of this scant literature we hypothesized that stability of temperament across infancy might be partially moderated by family SES.

Overview of Two Studies of the Stability of Infant Temperament

Infant temperament is structured, and early temperament theorists stressed the centrality and importance of stability of temperament as definitional to the construct. However, that thesis has since come into question. Studying the stability of temperament will lead to a deeper understanding of temperament and individual differences in infancy, as temperament in infancy has implications for other areas of development. For example, infant temperament has been integrated into models of emotions and developing personality, is linked to children's perceptions and interpretations of their experiences, and shapes child-caregiver relationships, processes associated with socialization, and parenting (Bates, 1987; Bates & Pettit, 2015; Chen & Schmidt, 2015; Paulussen-Hoogeboom et al., 2007; Porter & Hsu, 2003 Shiner & Caspi, 2012).

Because stability putatively represents a central feature of temperament, we set out to revisit this critical element, examining the moderating roles of infant age, gender, birth order, term status, and SES. To address whether and how these factors inflect the stability of temperament across infancy and to advance the existing literature, we used a common

established measure of infant temperament and studied the factor structure of temperament in different groups tracing the stability of identical Positive and Negative Affectivity factors (and the scales that compose those factors) across the first year of life. The majority of available empirical evidence points to overall moderate stability of temperament factors and scales, so our general default expectations favored some degree of stability, but reviews of the meager available literature with respect to important moderators of stability in infancy suggested some residual variation in stability. The first goal of each of the two complementary studies reported here was to explore the structure and stability of temperament at and between different points in infancy; the second goal was to evaluate potential moderation of stability to explore more deeply the developmental nature of infant temperament.

General Methods

We undertook two longitudinal studies of infant temperament to reach our two goals. Their common methods are described here.

Participants

Families that participated in Study 1 consisted of a European American community sample that varied in terms of maternal education and family SES. This sample was specifically chosen to be ethnically homogenous because infant temperament has been shown to vary with ethnicity and culture (Chen & Schmidt, 2015). An ethnically homogenous sample would therefore constitute a first step in understanding the ontogeny of infant temperament and its moderation that logically precedes embarking on studies and analyses with ethnically diverse samples, such as Study 2 (Bornstein, Jager, & Putnick, 2013).

Procedures

In both studies, families were recruited from patient populations of private obstetric and pediatric groups in urban and semi-urban catchment areas. At each wave, mothers completed a demographic questionnaire about the family and the Infant Behavior Questionnaire (IBQ; Rothbart, 1981, 1986). IBQ Scales are computed as means of all scale items. The IBQ was normed on 463 infants who were 3 to 12 months old; however, the IBQ has been used with infants from 2 weeks to 19 months of age (Goldsmith & Rothbart, 1991). In most IBQ studies, interscale correlations are low enough to avoid collinearity problems (Goldsmith & Rothbart, 1991), which justified considering each temperament scale independently. We explored individual scales and derived Positive and Negative Affectivity factors in each study. Rothbart (1981) reported alpha coefficients ranging from .72 to .85 at 3 months, .67 to .81 at 6 months, and .72 to .84 at 12 months. Satisfactory psychometric properties of the IBQ have been documented in multiple reports (Bridges, Palmer, Morales, Hurtado, & Tsai, 1993; Carranza et al., 2013; Goldsmith & Rothbart, 1991; Reznick, Gibbons, Johnson, & McDonough, 1989). The IBQ has also been used successfully to examine individual differences in reactivity and regulation via structured laboratory tasks, physiological recordings, and other biomarkers (Kochanska et al. 1998; Leppänen et al., 2011; Rothbart, Derryberry, & Hershey, 2000) and to explore relations among temperament,

socialization, parenting, and family functioning (Puura et al., 2013; Seifer, Schiller, Sameroff, Resnick, & Riordan, 1996).

Power Analyses

Post-hoc power analyses were computed for both studies to ascertain whether there was significant power to detect medium (r = .30) and large (r = .50) effects in two-tailed correlation analyses with α = .05 (Faul, Erdfelder, Lang, & Buchner, 2007). The power to detect medium and large stabilities in the full samples was sufficient (power = .74–1.00) in both studies. The power to detect medium and large stabilities in subgroups was also sufficient (power = .95–.99) in Study 2 but only sufficient for large effects in Study 1 (power = .89).

Preliminary Analyses and Data Analytic Plan

We first report factor analyses for Positive and Negative Affectivity and then descriptive statistics and stability estimates (zero-order Pearson correlation coefficients) across ages for IBQ Positive and Negative Affectivity factors and scale scores. Stability estimates are reported for the full sample as well as by age, gender, birth order, term status, and SES as appropriate. Average stability across scales (excluding the factors) was computed using Fisher's r-to-z transformation. Differences in stability between younger and older infants, girls and boys, firstborns and laterborns, term and preterm infants, and infants with mothers of lower and higher SES were also tested using Fisher's transformed z. In describing effect sizes, we follow Cohen's (1988) terminology, small effect size: estimate of population correlation, $r \approx .10$, medium effect size: $r \approx .30$, large effect size: $r \approx .50$.

Study 1: Infant Temperament from 2 to 5 to 13 Months in Girls and Boys

Participants

Mothers of 73 infants (36 girls, 37 boys) provided temperament ratings when their infants were 2, 5, and 13 months of age. All children were firstborn, term (M birth weight = 3569.79 g, SD = 446.28; M birth length = 52.22 cm, SD = 2.13), and healthy. All mothers and infants were European American, and mothers averaged 33.10 years of age (SD = 4.60) at the birth of the child. At the first wave, 97% of families were intact, and all were middle to upper SES (Hollingshead, 1975, Four-Factor Index of Social Status M = 56.32, SD = 8.28; see Bornstein, Hahn, Suwalsky, & Haynes, 2003). Ten percent of mothers completed only partial college or less, 53% completed a 4-year college degree, and 37% completed a graduate or professional degree.

Measures

Besides the IBQ, mothers completed the Social Desirability Scale (SDS; Crowne & Marlowe, 1960), which uses 33 items to assess adults' tendencies to respond to questions in a socially desirable fashion. Crowne and Marlowe reported that test-retest reliability for the SDS is .89. We included the SDS as a potential covariate to determine whether socially desirable response patterns affect maternal reports of stability of infant temperament (Bornstein, Putnick, et al., 2014). Maternal social desirability bias was associated with infant Positive Affectivity at 2 and 5 months, rs(68) = .25 and .35; duration of orienting at 2 and 5

months, *r*s(68) = .35 and .33; and smiling and laughter at 13 months, *r*(68) = .25, *p*s ≤ .05. Consequently, we computed residual scores for these 5 scales, controlling for maternal social desirability bias. When we explored these residual scores in analyses, all statistical decisions and effect sizes in Study 1 were unchanged from analyses of uncontrolled scores; for ease of interpretation we report descriptive statistics and analyses using untransformed scale scores.

Results

Structure: Positive and Negative Affectivity factors—Principal components analyses (PCA) with Varimax rotation on the 6 IBQ scales at each of the 3 waves supported the two-component solution: Positive Affectivity, consisting of smiling and laughter (rotated factor loadings = .70–.79), soothability (rotated factor loadings = .58–.73), and duration of orienting (rotated factor loadings = .76–.82), and Negative Affectivity, consisting of activity level (rotated factor loadings = .49–.74), fear (rotated factor loadings = .80–.89), and distress to limitations (rotated factor loadings = .62–.80). Positive Affectivity accounted for between 28.33% and 28.69% of the variance in the scales, and Negative Affectivity accounted for between 26.12% and 28.19% of the variance in the scales. Based on the PCA, we computed Positive Affectivity as the mean of smiling and laughter, soothability, and duration of orienting, and Negative Affectivity as the mean of activity level, fear, and distress to limitations.

Descriptive statistics and stability—The magnitudes of stability coefficients for the IBQ temperament factors and scales were medium to large from 2 to 5 months and from 5 to 13 months, but small to medium from 2 to 13 months (Table 1). Average stability did not differ across time. Girls' and boys' stability coefficients were similar, except that girls' Positive Affectivity, smiling and laughter, and soothability were more stable than boys' from 2 to 13 months, *z*s = 2.60 – 3.47, *p*s ≤ .05 – .001 (Table 2). Only 1 of 6 (17%) gender comparisons on the scales, and 2 of 18 (11%) comparisons on the items, were significant. The stability of girls' Positive Affectivity was also greater than the stability of their Negative Affectivity from 2 to 13 months, *z* = 2.47, *p* ≤ .05 (Raghunathan, Rosenthal, & Rubin, 1996).

Study 2: Infant Temperament from 6 to 12 Months in Girls and Boys, Firstborns and Laterborns, Terms and Preterms, and Lower- and Higher-SES Families

Participants

Mothers of 335 infants (165 girls, 170 boys; 135 firstborns, 200 laterborns; 194 term, 141 preterm infants; 111 lower-SES, 221 higher-SES) provided temperament ratings when their children were 6 and 12 months (term-corrected age was used for preterm infants; Aylward, 2002). Children varied in ethnicity, with 80% European American, 6% African American, 4% Latin American, and 10% biracial or other.

Term infants (37–42 weeks gestation) of normal birth weights (≥ 2500 g) and preterm infants (< 33 weeks gestation) with birth weights of 750 to 1805 g were eligible. Of the 141

preterm infants in the study sample, 43 were of low birth weight (<2500 g and ≥ 1500 g), 72 were of very low birth weight (<1500 g and ≥ 1000 g), and 26 were of extremely low birth weight (< 1000 g). Both term and preterm infants were required to have 5-min Apgar scores ≥ 7 to be included. Term pregnancy was defined as 40 weeks of gestation for the calculation of term corrected age for preterm infants. Infants with serious medical conditions (e.g., congenital abnormalities, periventricular/intraventricular haemorrhage greater than Grade II, significant cardiac, respiratory, ophthalmologic, gastrointestinal, hematologic disease, etc.) that could affect growth and development were excluded. By design, preterm infants had shorter gestations (*M* = 29.80 weeks, *SD* = 2.12) than term infants (*M* = 39.32 weeks, *SD* =1.20), $t(204.26) = -47.99$, $p \leq .001$, $\eta^2_p = .89$; preterm infants weighed less at birth (*M* = 1296.72 g, *SD* = 288.39) than term infants (*M* = 3499.73 g, *SD* = 444.50), $t(329.11) = -54.93$, $p \leq .001$, $\eta^2_p = .89$; and preterm infants scored lower on the 5-min Apgar (*M* = 8.32, *SD* = .72) than term infants (*M* = 9.10, *SD* = .43), $t(206.63) = -11.27$, $p \leq .001$, $\eta^2_p = .31$. There were more firstborns in the preterm group (52.5%) than in the term group (31.4%), $\chi^2(1, N = 335) = 15.02$, $p \leq .001$. The two groups did not differ in the distribution of child gender, $\chi^2(1, N = 335) = 0.01$, *ns*. Mothers of preterm infants had fewer years of education (*M* = 13.20, *SD* = 2.30) than mothers of the term infants (*M* = 14.83, *SD* = 2.13), $t(330) = -6.67$, $p \leq .001$, $\eta^2_p = .12$.

Results

Structure: Positive and Negative Affectivity factors—PCA with Varimax rotation at both waves supported the two-component solution reported in Study 1: Positive Affectivity, consisting of smiling and laughter (rotated factor loadings = .83 and .85 at 6 and 12 months, respectively), soothability (rotated factor loadings = .70 and .69), and duration of orienting (rotated factor loadings = .77 and .70), and Negative Affectivity, consisting of activity level (rotated factor loadings = .58 and .67), fear (rotated factor loadings = .76 and .68), and distress to limitations (rotated factor loadings = .85 and .83). Positive Affectivity accounted for 33.56% and 29.68% of the total variance in the scales at 6 and 12 months, respectively. Negative Affectivity accounted for 26.24% and 26.46% of the total variance in the scales at 6 and 12 months, respectively. Based on the PCA, Positive Affectivity was computed as the mean of smiling and laughter, soothability, and duration of orienting, and Negative Affectivity was computed as the mean of activity level, fear, and distress to limitations.

Descriptive statistics and stability—The magnitudes of stability coefficients for the IBQ temperament factors and scales were generally large from 6 to 12 months in all sub-samples (Tables 3 and 4). No differences emerged in stability between girls and boys (except that stability for distress to limitations was stronger in girls than in boys, $z = 2.46$, $p < .05$) or between lower (maternal education was high school graduate or lower) and higher (maternal education was some college or higher education) SES groups (except that stability of activity level was stronger in higher-SES than in lower-SES groups, $z = 2.36$, $p < .05$). Only 1 of 6 (17%) gender comparisons on the items was significant, and 1 of 6 (17%) SES comparisons was significant. No differences emerged between firstborns and laterborns or between term and preterm infants. Given the significant differences in maternal education between preterm and term children and the significant difference in stability on activity level between higher- and lower-SES groups, we examined stability across lower- and higher-SES

groups within preterm and term children. Children showed medium to strong stability in all 4 sub-samples (ranging from .33 to .69 in preterm infants and from .30 to .68 in term infants). Stability of activity level was stronger in higher SES (r = .58) than in lower SES (r = .30), z = 1.98, $p < .05$, for term infants; stability of activity level was only marginally stronger in higher SES (r = .69) than in lower SES (r = .47), z = 1.95, p = .051, for preterm infants. No other significant group differences were found.

To assess whether stability differed in preterms with varying medical characteristics, we controlled birth weight, gestational age, number of days in the hospital, chronic lung disease, 5-min Apgar scores, being small for gestational age, and being on a ventilator. No differences in stability were found between controlled and uncontrolled analyses, except preterm infants who had never been on a ventilator had stronger stability of Positive Affectivity (after controlling for birth weight and gestational age, r = .81, $p < .001$) and smiling and laughter (r = .70, $p < .001$) than infants who had been on a ventilator (r = .60 for Positive Affectivity, and r = .55 for smiling and laughter, both ps < .001), Fs(1, 137) = 5.72 and 3.92, ps = .02 and .05, respectively. However, stabilities were still substantial and significant in both groups.

General Discussion

Two coordinated studies examined the stability of infant temperament by maternal report in girls and boys, firstborns and laterborns, terms and preterms, and lower- and higher-SES families over a range of 2 to 13 months using a standard and well-established measure. One principal advance of this study was the systematic exploration of the moderation of developmental stability of infant temperament. A second advance was the use of the same instrument (the IBQ) at different ages and with different groups, thereby reducing method variance that has undermined conclusions about infant temperament to date. Factor analyses consistently supported a two-factor structure, with Positive Affectivity consisting of smiling and laughter, soothability, and duration of orienting, and Negative Affectivity consisting of activity level, fear, and distress to limitations across two samples, and the two factors and their scales were generally stable by gender, birth order, term status, and SES. These results were based on maternal report, but ancillary analysis (in Study 1) showed essentially identical results controlling for maternal social desirability bias (Bornstein, Putnick, et al., 2014). Whereas infants' emotions are fleeting or reactive, their temperament is commonly thought of as more enduring. Moreover, short-term fluctuations in an infant's emotional state are seldom thought to have long-term implications, whereas temperament is believed to be meaningful precisely because it is considered stable and prognostic, and often viewed as an early foundation for personality (Caspi et al., 2003; Chen & Schmidt, 2015; Rothbart, 2011). For example, Positive and Negative Affectivity show strong similarities with two of the "Big Five" factors that have emerged from analyses of personality in adults, *viz*. Extraversion and Neuroticism, respectively.

Structure and Stability of Infant Temperament

Stability of temperament in infancy is generally medium to large in magnitude. However, not unexpectedly (on the Guttman simplex) lower stability estimates were observed over longer inter-assessment periods (i.e., small to medium in magnitude from 2 to 13 months in

Study 1). Our stability findings across infancy accord with other research examining longer-term prediction from early appearing temperament (Kagan, 2013; Rothbart, 2011; Schwartz, Wright, Shin, Kagan, & Rauch, 2003). For example, Putnam et al. (2008) investigated longitudinal stability for factors and scales of the revised IBQ with 3- to 12-month-olds, the Early Childhood Behavior Questionnaire (ECBQ; Putnam et al., 2006) with 18- to 32-month-olds, and the Children's Behavior Questionnaire (CBQ; Rothbart et al., 2001) when children were on average 50 months old. All scales exhibited stability across adjacent measurement periods, and equivalent scales in the ECBQ and CBQ were correlated. At the factor level, Positive Affect (Surgency) and Negative Affect were stable across all time points. Similarly, Komsi and colleagues (2006) described stability of temperament from 6 months (IBQ) to 5.5 years (CBQ) in Finnish children.

With respect to moderation of the stability of temperament, we detected few meaningful effects. Similarities in stability coefficients for girls and boys were prevalent; however, girls showed somewhat more pervasive stability for positive reactivity. This gender effect could reflect girls' greater relatedness to mothers where consistency of mother-child interactions maintains expressions of girls' positivity over time. For instance, Keating and Manning (1974) observed that maternal acceptance was more closely associated with girls' emotional equilibrium (consisting of ratings assigned for crying intensity/duration, affective stability, vulnerability to stress, and ease of recovery from being upset). Greater stability in girls might be encouraged as well by higher levels of maternal acceptance to maintain a more positive disposition compared to boys. Cameron (1978) reported a number of correlations between temperament attributes (adaptability, intensity, threshold, distractibility, persistence, approach-withdrawal, mood and rhythmicity) in early childhood and parenting factors (disapproval-rejection, inconsistency, conflict regarding childrearing, strictness, and concern-protectiveness), with the majority of associations emerging for girls. This pattern of results was interpreted as indicating that girls might be more influenced by caregiving with respect to temperament development. However, average stability of all IBQ scales did not differ for girls and boys in either study reported here. Nor were any differences between firstborns and laterborns in the stability of temperament detected. Perhaps the 50% of genes that siblings share on average, reinforced by constituents of the family environment they share (parenting, home), promote the similar stabilities in siblings we observed. Term and healthy preterm infants did not differ with respect to stability of temperament either. In addition, few medical factors (8 of 72 or 11% of possible correlations) pertinent to preterm status (birth weight, gestational age, number of days in the hospital, chronic lung disease, 5-min Apgar scores, size for gestational age, and being on a ventilator) correlated with temperament scale scores, and they were generally low in magnitude. Only the use of a ventilator emerged as a significant factor in the stability of infant temperament, with infants who did not require a ventilator demonstrating higher stability of Positive Affectivity and smiling and laughter relative to infants who did. This lower stability for infants requiring a ventilator seems to indicate that ventilator use is somehow disruptive to developmental processes involved in the production of smiles, laughter, other vocalizations, and motor activity that signal joy/pleasure. This disruption may be the lingering result of physical limitations imposed by ventilators on infants' ability to vocalize, control facial muscles, and interact normally with caregivers, as the latter is likely an ingredient in the development of

positive affectivity and the social smile. However, healthy prematurity per se appears not necessarily to moderate the stability of infant temperament. Finally, with one exception, temperament was similarly stable in groups of infants whose mothers varied in educational achievement. The present study therefore provides evidence of considerable consistency in temperament structure and stability across multiple contrasting samples.

Stability Redux: A Glass Half-Full and Half-Empty

Stability correlations index the degree to which the rank order of individuals in a group is preserved from one assessment occasion to the next. Our developmentally sensitive analyses comparing children 2 to 13 months of age demonstrated consistent stability even across distinct developmental periods of infancy associated with rapidly growing and changing capabilities in motor, cognitive, language, and socioemotional functioning (Bornstein, Arterberry, et al., 2014). What sorts of developmental processes might account for preservation of infant standing on temperament despite these thoroughgoing transformations?

Development is governed by genetic and biological factors in inextricable combination with environmental influences and experiences (Lerner et al., 2015). Thus, stability of temperament might best be seen as a joint reflection of constitutional characteristics of the child and of the family system. Some evidence points to heritability of (presumably stable) individual differences in temperament (Gagne & Goldsmith, 2011; Plomin, DeFries, Craig, & McGuffin, 2003; Rothbart, 2011), but genetic underpinnings do not imply immutability (Saudino, 2012). Stability might also emerge through the individual's experiential transactions with a consistent environment. Unsurprisingly in this light, Goldsmith, Lemery, Buss, and Campos (1999) identified genetic/biological *and* environmental/experiential sources of individual differences in temperament in a sample of 3- to 16-month-old twins: Using a quantitative behavior-genetic design, they learned that both additive genetic/biological and shared environmental/experiential effects were needed to best represent subscales of mother-completed IBQs. As Goldsmith et al. observed, the ontogeny of temperament reflects transactions of the biological organism with the social environment through time.

Although temperament may be stable in some significant degree, we should not expect overt expressions of temperament to be invariant through time; expressions of temperament may still change due to maturation, experience, the contexts in which temperament is manifest, as well as the integration of temperament into emerging personality. Relative stability should not be interpreted as indicative of lifelong consistency because developmental change and individual differences in developmental timetables must always be taken into account. Moreover, even large stability estimates leave substantial variance unaccounted for. The stability correlations we observed were statistically significant, but the range of shared variance in average stability between adjacent time-points across both studies was only 20–29%, suggesting that 71–80% of the variance in temperament at a later time point was not explained by temperament at an earlier time point. To be stable does not mean to be impervious. A great deal of change in individual differences in temperament across time occurs, and change in temperament might reflect children's changing and maturing biology

(Riese, 1987a). Insofar as temperament is influenced by genetics, genetic expression changes during ontogeny (Naumova, Lee, Rychkov, Vlasova, & Grigorenko, 2013; Szyf & Bick, 2012) and likely contributes to developmental variability in individuals. The rate of expression of the human genome in the prefrontal cortex, for example, decreases substantially after birth, remaining higher in infancy relative to subsequent developmental periods (Colantuoni et al., 2011), and expression for genes associated with dendritic and synaptic development (e.g., genes controlling ionic channels) across 16 brain regions (including 11 neocortex areas) increases across infancy (Kang et al., 2011). Temperament also reacts to environmental change. The life-span perspective in psychology further specifies that human beings are open systems. For example, a temperamentally fearful infant may learn self-soothing strategies that may make the same child appear less fearful later; an infant who is rewarded for smiling may smile even more later on. Another explanation for change in temperament over time is that mothers (who rated temperament) may change in their perceptions of their child. For example, an infant who scores high on distress to limitations might be rated lower on this dimension later in childhood because the mother has learned how to structure the child's environment so as to curb limitations and/or overt displays of distress and so she now perceives her child as lower on distress to limitations.

In short, organic systems work to maintain homeostasis, but the plastic nature of psychological functioning ensures adaptability. Thus, people (infants included) exhibit both stability and instability in many characteristics (temperament included) throughout the life course. This finding leaves latitude for early intervention and may inform prevention and remediation strategies aimed at infant temperament and early parent–infant interactions.

Maternal Report of Stability of Infant Temperament

Infant temperament has been measured through caregiver reports, naturalistic home and structured laboratory observations, as well as physiological indicators (Bornstein, 2014a; Gartstein et al., 2012; Rothbart, 2011). Each of these approaches offers relative advantages, but each is also constrained by limitations. Parent-report questionnaires (used here) represent the most widely adopted and appreciated approach (Rothbart, 2011), primarily because they enable researchers to access an extensive and rich knowledge base of infants from caregivers who occupy a unique position in having experienced how their child behaves in a wide variety of situations and across multiple contexts over the child's entire lifetime. Indeed, Thomas and Chess instigated The New York Longitudinal Study and infant temperament research by interviewing mothers who (they argued) are likely to provide the most insightful reports of infant temperament based on their long-term, intimate, and broad experiences with their children. Reports are also convenient, relatively easy and inexpensive to administer and analyze, and allow the study of multiple dimensions of temperament (Bates, 1989; Bornstein, 2014a).

However, caution regarding parental report has been voiced, with concerns raised about potential sources of error (Kagan, 2013; Rothbart, 2011). For example, caregivers' limited knowledge of the meaning of infant behavior, unfamiliarity with the behavior of other infants with whom the infant might be compared, and inaccurate memories of events involving the infant all may compromise caregiver report. Subjective reports are also

vulnerable to biases of various sorts; for example, behavioral characteristics of a child may be amplified or attenuated in the parent's mind when one child is compared with another child in the same family. Other factors affecting parent-report accuracy include inadequate attention to instructions and characteristics of the reporter (e.g., mothers' personality is associated with their ratings of their infants; Mangelsdorf, McHale, Diener, Goldstein, & Lehn, 2000). Furthermore, relying on parental report means that observed stabilities perforce include shared variance associated with using the same informant and measure at different times (e.g., Winstanley et al., 2014), which of course may raise reported stability.

The assessment of infant temperament ineluctably filters its subject, and so a central endeavor in measurement is to minimize error. Many concerns with parent reports are at least partially mitigated by careful construction and presentation of items, only asking about recently occurring events, and inquiring about concrete behaviors rather than requiring parents to make abstract, global, or comparative judgments (Rothbart & Goldsmith, 1985). For example, vaguely impressionistic items (such as "My child is always on the go.") are less informative and predictive than more objective items that more precisely describe specific behaviors and the spatial and temporal settings in which they occur (such as "My child sits still (moves little) during procedures like hair brushing or nail cutting."). The IBQ was specifically designed to address these concerns, asking caregivers to report, on a 7-point scale, the relative frequency of occurrence of specific infant reactions in concretely defined situations during recent fixed reference periods. The response format of the IBQ presents sets of items based on the context eliciting the infant's reactions (e.g., bathing and dressing), which is thought to reduce social desirability; to account for that potential bias of maternal report, social desirability was statistically controlled in Study 1 with no significant changes in maternal IBQ reports of infant temperament stability. Thus, the IBQ format relies on recognition and augments report accuracy by minimizing problems associated with recall and limiting biases associated with more global questions. Nor does the IBQ require caregivers to make comparative judgments with respect to other children, diminishing the need for familiarity with other infants.

Overall, despite potential limitations, parent-report instruments offer a developmentally appropriate assessment tool with strong reliability and superior predictive validity relative to other temperament measures (Pauli-Pott, Mertesacker, Bade, Haverkock, & Beckmann 2003). No one measurement approach can boast of being limitation-free, and parent-report instruments are no exception. However, parents are uniquely positioned to provide information regarding their infant's temperament, given that others do not have the necessary access to babies to describe them as fully. Finally, maternal report is valuable *sui generis* insofar as caregiver perceptions (accessed via reports) constitute a prime social milieu of infant rearing (Bornstein, 2014a).

Limitations and Future Directions

The two studies presented here make a unique contribution to the field of infant temperament by focusing on structure and stability, by closely examining multiple potential significant moderators of each, and by doing so in a systematic, unified, and comprehensive way. Measuring stability calls for longitudinal within-subject designs that provide powerful

analyses as they control for variance due to additional individual factors. Some limitations should be noted, however. The samples were not representative of the U.S. population. Study 1 used an ethnically homogeneous sample, but Study 2 included a broader range of ethnic and SES groups. The power of the statistical tests performed would have been enhanced by still larger sample sizes; nonetheless, our analyses were adequately powered and provided consistently significant findings. A number of potential moderators of stability of infant temperament was examined, but additional ones, such as maternal depression or culture (Bornstein, 1989; Wang & Dix, 2013), could be important to understanding stability of infant temperament. Birth order was assessed, but ours was a between- not within-family design, which should be considered for future methodological refinements. Stability of temperament in preterm infants was studied, but our sample was healthy, and stability of temperament also warrants investigation in fragile preterms and in other medically compromised infant populations (Mayes, Bornstein, Chawarska, Haynes, & Granger, 1996).

Further research on possible biological and environmental factors influencing stability and change of infant temperament is clearly needed. If features of temperament in infancy can be altered by maturation and experience, stability over long periods of time will be affected and may be subject to intervention and (if negative) remediation. Furthermore, multi-method approaches in which contributions associated with parent ratings and other avenues of measurement can be studied and compared are called for (Gagne & Goldsmith, 2011). In this situation aggregating responses from multiple sources might achieve still greater convergence. Some investigators have compared and combined home observation and matching parent-report procedures to assess infant temperament (Bornstein et al., 1991), whereas others have compared laboratory-based assessments with home observations (Gagne, Van Hulle, Aksan, Essex, & Goldsmith, 2011). For future investigations, it would be developmentally revealing as well to continue researching stability and prediction of temperament into later stages of development and examine their association with the personality, educational, and social development of the child.

Mothers and fathers are likely influenced by somewhat different factors in perceiving their infant's temperaments (Parade & Leerkes, 2008; Pesonen et al., 2006). Here, we used mothers as informants. But fathers are believed to have more leeway in choosing how to enact the parental role (Cabrera, Tamis-LeMonda, Bradley, Hofferth, & Lamb, 2000). Last, the majority of studies involving stability of temperament have concerned homotypic stability, as here, *viz*. the extent to which specific temperamental traits endure across different stages of development. A complementary approach refers to heterotypic stability and involves the expectation that some early temperamental characteristics would share common underlying attributes with later phenotypically different aspects of temperament or personality (Carranza et al., 2013).

Conclusions

Despite the seeming randomness and reactivity of infancy, infant temperament is commonly thought of as organized and enduring. In two studies, we focused on structure and stability of infant temperament development from multiple perspectives of infant age, gender, birth order, term status, and SES. Analyses consistently supported a two-factor structure of

Positive and Negative Affectivity across samples, and infant temperament factors and scales were generally stable across age (if somewhat more stable over shorter intervals), gender, birth order, term status, and SES. Given the importance of replicating findings in science (Duncan, Engel, Claessens, & Dowsett, 2014), the reproduction of both structure and stability of factors and scales of infant temperament and extension across multiple moderators lend credence to the robustness of the structure and stability of infant temperament.

Temperament in infancy is considered meaningful because it is considered stable and prognostic, viewed as an early foundation of personality (Chen & Schmidt, 2015). Stability is important for any individual-difference characteristic and has multiple implications for later development. For example, infants' enduring individuality influences their own development, and even small early differences in stable temperaments may have larger-than-expected later effects (Bornstein, 2014b). Temperament in infancy can affect children's interactions with the world and their learning, color the way children interpret their experiences, shape how children compare themselves to others and the manner in which others (i.e., adults and peers) perceive and respond to them, guide the choices children make, and modify the ways children interpret and manipulate their environment. Parenting differs for infants with different temperaments. Successful adjustment likely depends on interactions between temperament and setting. By studying stable child temperament vis-à-vis social interaction, cognition, behavioral problems, and other domains, researchers will be on firmer ground in reaching an even deeper understanding of the significant and complex roles played by infant temperament in later child development.

Acknowledgments

We thank A. Dovidio for assistance and Abbott Nutrition for Study 2 data.

References

Atzaba-Poria N, Pike A. Correlates of parenting for mothers and fathers from English and Indian backgrounds. Parenting: Science and Practice. 2008; 8:17–40.10.1080/15295190701665698

Aylward GP. Methodological issues in outcome studies of at-risk infants. Journal of Pediatric Psychology. 2002; 27:37–45.10.1093/jpepsy/27.1.37 [PubMed: 11726678]

Barnett MA, Deng M, Mills-Koonce WR, Willoughby M, Cox M. Interdependence of parenting of mothers and fathers of infants. Journal of Family Psychology. 2008; 22:561–573.10.1037/0893-3200.22.3.561 [PubMed: 18729670]

Bates, JE. Temperament in infancy. In: Osofsky, JD., editor. Handbook of infant development. 2. Oxford, England: Wiley; 1987. p. 1101-1149.

Bates, JE. Concepts and measures of temperament. In: Kohnstamm, GA.; Bates, JE.; Rothbart, MK., editors. Temperament in childhood. New York: Wiley; 1989. p. 3-26.

Bates JE, Freeland CB, Lounsbury ML. Measurement of infant difficultness. Child Development. 1979; 50:794–803.10.2307/1128946 [PubMed: 498854]

Bates, JE.; Pettit, GS. Temperament, parenting, and social development. In: Grusec, J.; Hastings, P., editors. Handbook of socialization. 2. New York: Guilford; 2014. p. 372-397.

Bates, JE.; Schermerhorn, AC.; Petersen, IT. Temperament and parenting in developmental perspective. In: Zentner, M.; Shiner, RL., editors. Handbook of temperament. New York: Guilford Press; 2012. p. 425-441.

Behrman, RE.; Butler, AS., editors. Preterm birth: Causes, consequences, and prevention. Washington, DC: National Academies Press; 2006.

Bornstein MH. Cross-cultural developmental comparisons: The case of Japanese--American infant and mother activities and interactions. What we know, what we need to know, and why we need to know. Developmental Review. 1989; 9:171–204.10.1016/0273-2297(89)90028-2

Bornstein, MH. Parenting x gender x culture x time. In: Wilcox, WB.; Kline, KK., editors. Gender and parenthood. New York: Columbia University Press; 2013. p. 91-119.

Bornstein, MH. Parents' reports about their children's lives. In: Ben-Arieh, A.; Cashmore, J.; Goodman, GS.; Melton, GB., editors. The SAGE handbook of child research. London, England: SAGE; 2014a. p. 486-533.

Bornstein MH. Human infancy … and the rest of the lifespan. Annual Review of Psychology. 2014b; 65:121–158.10.1146/annurev-psych-120710-100359

Bornstein, MH.; Arterberry, ME.; Lamb, ME. Development in infancy: A contemporary introduction. 5. New York: Psychology Press; 2014.

Bornstein, MH.; Bornstein, L. Psychological stability. In: Darity, WA., Jr, editor. International Encyclopedia of Social Sciences. 2. Vol. 8. Detroit, MI: Macmillan Reference USA; 2008. p. 74-75.

Bornstein MH, Gaughran JM, Seguí I. Multimethod assessment of infant temperament: Mother questionnaire and mother and observer reports evaluated and compared at five months using the Infant Temperament Measure. International Journal of Behavioral Development. 1991; 14:131–151.10.1177/016502549101400202

Bornstein, MH.; Hahn, CS.; Suwalsky, JTD.; Haynes, OM. Socioeconomic status, parenting, and child development: The *Hollingshead Four-Factor Index of Social Status*and the *Socioeconomic Index of Occupations*. In: Bornstein, MH.; Bradley, RH., editors. Socioeconomic status, parenting, and child development. Mahwah, NJ: Erlbaum; 2003. p. 29-82.

Bornstein MH, Jager J, Putnick DL. Sampling in developmental science: Situations, shortcomings, solutions, and standards. Developmental Review. 2013; 33:357–370.10.1016/j.dr.2013.08.003 [PubMed: 25580049]

Bornstein MH, Putnick DL, Heslington M, Gini M, Suwalsky JT, Venuti P, Zingman de Galperin C. Mother-child emotional availability in ecological perspective: Three countries, two regions, two genders. Developmental Psychology. 2008; 44:666–680.10.1037/0012-1649.44.3.666 [PubMed: 18473635]

Bornstein MH, Putnick DL, Lansford JE, Deater-Deckard K, Bradley R. Gender in low- and middle-income countries. Monograph of the Society for Research in Child Development. In-Press.

Bornstein MH, Putnick DL, Lansford JE, Pastorelli C, Skinner AT, Sorbring E, Oburu P. Mother and father socially desirable responding in nine countries: Two kinds of agreement and relations to parenting self-reports. International Journal of Psychology. In-Press. 10.1002/ijop.12084

Bradley RH, Corwyn RF. Socioeconomic status and child development. Annual Review of Psychology. 2002; 53:371–399.10.1146/annurev.psych.53.100901.135233

Braungart-Rieker JM, Hill-Soderlund AL, Karrass J. Fear and anger reactivity trajectories from 4 to 16 months. Developmental Psychology. 2010; 46:791–804.10.1037/a0019673 [PubMed: 20604602]

Brazelton, T. Neonatal behavioral assessment scale. London: Spastics International Medical Publications; 1973.

Bridges LJ, Palmer SA, Morales M, Hurtado M, Tsai D. Agreement between affectively based observational and parent-report measures of temperament at infant age 6 months. Infant Behavior and Development. 1993; 16:501–506.10.1016/0163-6383(93)80008-V

Bridgett DJ, Gartstein MA, Putnam SP, Lance KO, Iddins E, Waits R, Vanvleet J, Lee L. Emerging effortful control in toddlerhood. Infant Behavior and Development. 2011; 34:189–199.10.1016/j.infbeh.2010.12.008 [PubMed: 21186061]

Bridgett DJ, Gartstein MA, Putnam SP, McKay T, Iddins E, Robertson C, Rittmueller A. Maternal and contextual influences and the effect of temperament development during infancy on parenting in toddlerhood. Infant Behavior and Development. 2009; 32:103–116.10.1016/j.infbeh.2008.10.007 [PubMed: 19111913]

Buss, A. Temperaments as personality traits. In: Kohnstamm, GA.; Bates, JE.; Rothbart, MK., editors. Temperament in childhood. Oxford, England: Wiley; 1989. p. 49-58.

Buss, AH.; Plomin, R. Temperament. Hillsdale, NJ: Erlbaum; 1984.

Cabrera N, Tamis-LeMonda CS, Bradley RH, Hofferth S, Lamb ME. Fatherhood in the twenty-first century. Child Development. 2000; 71:127–136.10.1111/1467-8624.00126 [PubMed: 10836566]

Cameron JR. Parental treatment, children's temperament and the risk of childhood behavioral problems. Annual Progress in Child Psychiatry and Child Development. 1978; 47:233–244.10.1111/j.1939-0025.1977.tb01268.x

Carey WB, McDevitt SC. Revision of the infant temperament questionnaire. Pediatrics. 1978a; 61:735–739. [PubMed: 662513]

Carey WB, McDevitt SC. The measurement of temperament in 3–7 year old children. Child Psychology & Psychiatry & Allied Disciplines. 1978b; 19:245–253.10.1111/j.1469-7610.1978.tb00467.x

Carranza JA, Gonzalez-Salinas C, Ato E. A longitudinal study of temperament continuity through IBQ, TBAQ and CBQ. Infant Behavior and Development. 2013; 36:749–761.10.1016/j.infbeh.2013.08.002 [PubMed: 24036224]

Carranza Carnicero JA, Pérez López J, Salinas-González MC, Martínez-Fuentes MT. A longitudinal study of temperament in infancy: Stability and convergence of measures. European Journal of Personality. 2000; 14:21–37.10.1002/(SICI)1099-0984(200001/02)14:1<21::AID-PER367>3.0.CO;2-A

Casalin S, Luyten P, Vliegen N, Meurs P. The structure and stability of temperament from infancy to toddlerhood: A one-year prospective study. Infant Behavior and Development. 2012; 35:94–108.10.1016/j.infbeh.2011.08.004 [PubMed: 21962396]

Caspi A, Harrington H, Milne B, Arnell JW, Theodore RF, Moffitt TE. Children's behavioral styles at age 3 are linked to their adult personality traits at age 26. Journal of Personality. 2003; 71:495–514.10.1111/1467-6494.7104001 [PubMed: 12901429]

Chen, X.; Schmidt, LA. Temperament and personality. In: Lamb, ME.; Lerner, RM., editors. Social, emotional, and personality development. Volume 3 of the Handbook of child psychology and developmental science. 7. Hoboken, NJ: Wiley; 2015. p. xx-xx.

Clarke-Stewart KA. Interactions between mothers and their young children: Characteristics and consequences. Monographs of the Society for Research in Child Development. 1973; 38:1–109. [PubMed: 4783875]

Cohen, J. Statistical power analysis for the behavior sciences. Hillsdale, NJ: Erlbaum; 1988.

Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT, Kleinman JE. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature. 2011; 478:519–524.10.1038/nature10524 [PubMed: 22031444]

Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology. 1960; 24:963–968.10.1037/h0047358

Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society Series B (Methodological). 1977; 39:1–38.

Duncan GJ, Engel M, Claessens A, Dowsett CJ. The value of replication for developmental science. Developmental Psychology. 2014; 50:2417–2425. http://dx.doi.org/10.1037/a0037996. [PubMed: 25243330]

Ellis BJ, Boyce WT, Belsky J, Bakermans-Kranenberg MJ, van IJzendoorn MH. Differential susceptibility to the environment: An evolutionary-neurodevelopmental theory. Development and Psychopathology. 2011; 23:7–28.10.1017/S0954579410000611 [PubMed: 21262036]

Else-Quest NM, Hyde JS, Goldsmith HH, Van Hulle CA. Gender differences in temperament: A meta-analysis. Psychological Bulletin. 2006; 132:33–72.10.1037/0033-2909.132.1.33 [PubMed: 16435957]

Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. Behavior Research Methods. 2007; 39:175–191.10.3758/BF03193146 [PubMed: 17695343]

Furman, W.; Lanthier, R. Parenting siblings. In: Bornstein, MH., editor. Handbook of Parenting. 2. Mahwah, NJ: Erlbaum; 2002. p. 165-188.

Gagne JR, Goldsmith HH. A longitudinal analysis of anger and inhibitory control in twins from 12 to 36 months of age. Developmental Science. 2011; 14:112–124.10.1111/j.1467-7687.2010.00969.x [PubMed: 21159093]

Gagne JR, Van Hulle CA, Aksan N, Essex MJ, Goldsmith HH. Deriving childhood temperament measures from emotion-eliciting behavioral episodes: scale construction and initial validation. Psychological Assessment. 2011; 23:337–353.10.1037/a0021746 [PubMed: 21480723]

Garcia Coll CT, Halpern LF, Vohr BR, Seifer R, Oh W. Stability and correlates of change of early temperament in preterm and full-term infants. Infant Behavior & Development. 1992; 15:137–153.10.1111/j.1467-7687.2010.00969.x

Gartstein, MA.; Bridgett, DJ.; Low, CM. Self- and other-report measures of temperament. In: Zentner, M.; Shiner, RL., editors. Handbook of temperament. New York: Guilford Press; 2012. p. 183-208.

Gartstein MA, Bridgett DJ, Rothbart MK, Robertson C, Iddins E, Ramsay K, Schlect S. A latent growth examination of fear development in infancy: contributions of maternal depression and the risk for toddler anxiety. Developmental Psychology. 2010; 46:651–668.10.1037/a0018898 [PubMed: 20438177]

Gartstein MA, Gonzalez C, Carranza JA, Ahadi SA, Ye R, Rothbart MK, Yang SW. Studying cross-cultural differences in the development of infant temperament. Child Psychiatry and Human Development. 2006; 37:145–161.10.1007/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 (US) and Russia. Infant Behavior and Development. 2005; 28:54–61.10.1016/j.infbeh.2004.09.003

Gartstein MA, Rothbart MK. Studying infant temperament via the revised infant behavior questionnaire. Infant Behavior & Development. 2003; 26:64–86.10.1016/S0163-6383(02)00169-8

Goldberg, S.; DiVitto, B. Parenting children born preterm. In: Bornstein, MH., editor. Handbook of parenting. 2. Vol. 1. Mahwah, NJ: Erlbaum; 2002. p. 329-354.

Goldsmith HH, Buss KA, Lemery KS. Toddler and childhood temperament: expanded content, stronger genetic evidence, new evidence for the importance of environment. Developmental Psychology. 1997; 33:891–905.10.1037/0012-1649.33.6.891 [PubMed: 9383612]

Goldsmith HH, Lemery KS, Buss KA, Campos JJ. Genetic analyses of focal aspects of infant temperament. Developmental Psychology. 1999; 35:972–985.10.1037/0012-1649.35.4.972 [PubMed: 10442866]

Goldsmith, HH.; Rothbart, MK. Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In: Strelau, J.; Angleitner, A., editors. Explorations in temperament. New York: Plenum Press; 1991. p. 249-272.

Hallers-Haalboom ET, Mesman J, Groeneveld MG, Endendijk JJ, van Berkel SR, van der Pol LD, Bakermans-Kranenburg MJ. Mothers, fathers, sons and daughters: Parental sensitivity in families with two children. Journal of Family Psychology. 2014; 28:138–147.10.1037/a0036004 [PubMed: 24635666]

Halpern LF, Garcia Coll CT. Temperament of small-for-gestational-age and appropriate-for-gestational-age infants across the first year of life. Merrill-Palmer Quarterly. 2000; 46:738–765.

Halpern LF, Garcia Coll CT, Meyer EC, Bendersky K. The contributions of temperament and maternal responsiveness to the mental development of small-for-gestational-age and appropriate-for-gestational-age infants. Journal of Applied Developmental Psychology. 2001; 22:199–224.10.1016/s0193-3973(01)00077-6

Hollingshead, AB. Unpublished manuscript. Yale University; 1975. The four-factor index of social status.

Honjo S, Mizuno R, Ajiki M, Suzuki A, Nagata M, Goto Y, Nishide T. Infant temperament and child-rearing stress: Birth order influences. Early Human Development. 1998; 51:123–135.10.1016/S0378-3782(97)00102-3 [PubMed: 9605465]

Howson, CP.; Kinney, MV.; Lawn, JE. Born too soon. The global action report on preterm birth. Geneva, Switzerland: March of Dimes, Partnership for Maternal, Newborn, & Child Health, Save the Children, World Health Organization; 2012.

Huizink, AC. Prenatal factors in temperament: The role of prenatal stress and substance use exposure. In: Zentner, M.; Shiner, RL., editors. Handbook of temperament. New York: Guilford Press; 2012. p. 297-314.

Hwang AW, Soong WT, Liao HF. Influences of biological risk at birth and temperament on development at toddler and preschool ages. Child: Care, Health, and Development. 2009; 35:817–825.10.1111/j.1365-2214.2009.01001.x

Hyde JS. Gender similarities and differences. Annual Review of Psychology. 2014; 65:373–398.10.1146/annurev-psych-010213-115057

Jenkins JM, Rasbash J, O'Connor TG. The role of the shared family context in differential parenting. Developmental Psychology. 2003; 39:99–113.10.1037/0012-1649.39.1.99 [PubMed: 12518812]

Kagan, J. The human spark. New York: Basic Books; 2013.

Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sestan N. Spatio-temporal transcriptome of the human brain. Nature. 2011; 478:483–489.10.1038/nature10523 [PubMed: 22031440]

Keating, GW.; Manning, TM. Infant temperament and sex of infant: Effects on maternal behavior. In: Shereshefsky, PM.; Yarrow, LJ., editors. Psychological aspects of a first pregnancy and early postnatal adaptation. New York: Raven Press; 1974.

Keresteš G. Birth order and maternal ratings of infant temperament. Studia Psychologica. 2006; 48:95–106.

Kochanska G, Coy KC, Tjebkes TL, Husarek SJ. Individual differences in emotionality in infancy. Child Development. 1998; 64:375–390.10.1111/j.1467-8624.1998.tb06196.x [PubMed: 9586213]

Komsi N, Räikkönen K, Heinonen K, Pesonen AK, Keskivaara P, Järvenpää AL, Strandberg TE. Transactional development of parent personality and child temperament. European Journal of Personality. 2008; 22:553–573.10.1002/per.690

Komsi N, Raikkönen K, Pesonen AK, Heinonen K, Keskivaara P, Jarvenpaa AL, Strandberg TE. Continuity of temperament form infancy to middle childhood. Infant Behavior & Development. 2006; 29:494–508.10.1016/j.infbeh.2006.05.002 [PubMed: 17138302]

Korner AF. Reliable individual differences in preterm infants' excitation management. Child Development. 1996; 67:1793–1805.10.1111/j.1467-8624.1996.tb01828.x [PubMed: 8890508]

Leaper, C. Parenting girls and boys. In: Bornstein, MH., editor. Handbook of parenting. 2. Vol. 1. Mahwah, NJ: Erlbaum; 2002. p. 189-225.

Leppänen JM, Peltola MJ, Puura K, Mäntymaa M, Mononen N, Lehtimäki T. Serotonin and early cognitive development: Variation in the tryptophan hydroxylase 2 gene is associated with visual attention in 7-month-old infants. Journal of Child Psychology and Psychiatry. 2011; 52:1144–1152.10.1111/j.1469-7610.2011.02391.x [PubMed: 21418063]

Lerner, RM.; Hershberg, RM.; Hilliard, LJ.; Johnson, SK. Concepts and theories of human development. In: Bornstein, MH.; Lamb, ME., editors. Developmental science: An advanced textbook. 7. New York: Psychology Press; 2015. p. x-xx.

Lovas GS. Gender and patterns of emotional availability in mother-toddler and father-toddler dyads. Infant Mental Health Journal. 2005; 26:327–353.10.1002/imhj.20056

Mangelsdorf SC, McHale JL, Diener M, Goldstein LH, Lehn L. Infant attachment: Contributions of infant temperament and maternal characteristics. Infant Behavior and Development. 2000; 23:175–196.10.1016/S0163-6383(01)00035-2

Martin, JA.; Hamilton, BE.; Ventura, SJ.; Osterman, MJK.; Matthews, TJ. Births: Final data for 2011. National Vital Statistics Reports. Vol. 62. Hyattsville, MD: National Center for Health Statistics; 2013.

Mayes LC, Bornstein MH, Chawarska K, Haynes OM, Granger RH. Impaired regulation of arousal in 3-month-old infants exposed prenatally to cocaine and other drugs. Development and Psychopathology. 1996; 8:29–42.10.1017/S0954579400006957

McCall, RB. Issues of stability and continuity in temperament research. In: Plomin, R.; Dunn, J., editors. The study of temperament. Hillsdale, NJ: Erlbaum; 1986. p. 13-26.

McDevitt S, Carey W. Stability of ratings versus perceptions of temperament from early infancy to 1–3 years. American Journal of Orthopsychiatry. 1981; 11:342–345.10.1111/j.1939-0025.1981.tb01379.x [PubMed: 7223877]

Mondschein ER, Adolph KE, Tamis-LeMonda CS. Gender bias in mothers' expectations about infant crawling. Journal of Experimental Child Psychology. 2000; 77:304–316.10.1006/jecp.2000.2597 [PubMed: 11063631]

Montirosso R, Fedeli C, Murray L, Morandi F, Brusati R, Perego GG, Borgatti R. The role of negative maternal affective states and infant temperament in early interactions between infants with cleft lip and their mothers. Journal of Pediatric Psychology. 2011; 37:241–250.10.1093/jpepsy/jsr089 [PubMed: 22004886]

Naumova OY, Lee M, Rychkov SY, Vlasova NV, Grigorenko EL. Gene expression in the human brain: The current state of the study of specificity and spatiotemporal dynamics. Child Development. 2013; 84:76–88.10.1111/cdev.12014 [PubMed: 23145569]

Nelson SK, Kushlev K, Lyubomirsky S. The pains and pleasures of parenting: When, why, and how is parenthood associated with more or less well-being? Psychological Bulletin. 2014; 140:846–895.10.1037/a0035444 [PubMed: 24491021]

Oberklaid F, Prior M, Sanson A. Temperament of preterm versus full-term infants. Journal of Developmental and Behavioral Pediatrics. 1986; 7:159–162.10.1097/00004703-198606000-00005 [PubMed: 3722391]

Parade SH, Leerkes EM. The reliability and validity of the Infant Behavior Questionnaire-Revised. Infant Behavior & Development. 2008; 31:637–646.10.1016/j.infbeh.2008.07.009 [PubMed: 18804873]

Park SY, Belsky J, Putnam S, Crnic K. Infant emotionality, parenting, and 3-year inhibition: Exploring stability and lawful discontinuity in a male sample. Developmental Psychology. 1997; 33:218–227.10.1037/0012-1649.33.2.218 [PubMed: 9147831]

Pauli-Pott U, Mertesacker B, Bade U, Haverkock A, Beckmann D. Parental perceptions and infant temperament development. Infant Behavior & Development. 2003; 26:27–48.10.1016/S0163-6383(02)00167-4

Pauli-Pott U, Mertesacker B, Beckmann D. Predicting the development of infant emotionality from maternal characteristics. Development and Psychopathology. 2004; 16:19–42.10.1017/S0954579404044396 [PubMed: 15115063]

Paulussen-Hoogeboom MC, Stams GJJM, Hermanns JMA, Peetsma TTD. Child negative emotionally and parenting from infancy to preschool: A meta-analytic review. Developmental Psychology. 2007; 43:438–453.10.1037/0012-1649.43.2.438 [PubMed: 17352551]

Pesonen AK, Raïkkönen K, Strandberg TE, Järvenpää AL. Do gestational age and weight for gestational age predict concordance in parental perceptions of infant temperament? Journal of Pediatric Psychology. 2006; 31:331–336.10.1093/jpepsy/jsj084 [PubMed: 16221953]

Pfeifer M, Goldsmith HH, Davidson RJ, Rickman M. Continuity and change in inhibited and disinhibited children. Child Development. 2002; 73:1474–1485.10.1111/1467-8624.00484 [PubMed: 12361313]

Plomin, R.; DeFries, JC.; Craig, IW.; McGuffin, P. Behavioral genetics. Washington, DC: American Psychological Association; 2003.

Plomin R, Emde R, Braungart J, Campos J, Corley R, Fulker D, DeFries J. Genetic change and continuity from fourteen to twenty months: The MacArthur Longitudinal Twin Study. Child Development. 1993; 64:1354–1376.10.1111/j.1467-8624.1993.tb02957.x [PubMed: 8222877]

Porter CL, Hsu HC. First-time mothers' perceptions of efficacy during the transition to motherhood: Links to infant temperament. Journal of Family Psychology. 2003; 17:54–64.10.1037/0893-3200.17.1.54 [PubMed: 12666463]

Posner MI, Rothbart MK, Sheese BE, Voelker P. Control networks and neuromodulators of early development. Developmental Psychology. 2012; 48:827–835.10.1037/a0025530 [PubMed: 21942663]

Putnam SM, Gartstein MA, Rothbart MK. Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. Infant Behavior & Development. 2006; 29:386–401.10.1016/infbeh.2006.01.004 [PubMed: 17138293]

Putnam SM, Rothbart MK, Gartstein MA. Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. Infant and Child Development. 2008; 17:387–405.10.1002/icd.582

Puura K, Mäntymaa M, Leppänen J, Peltola M, Salmelin R, Luoma I, Tamminen T. Associations between maternal interaction behavior, maternal perception of infant temperament, and infant social withdrawal. Infant Mental Health Journal. 2013; 34:586–593.10.1002/imhj.21417

Raghunathan TE, Rosenthal R, Rubin DB. Comparing correlated but nonoverlapping correlations. Psychological Methods. 1996; 1:178–183.10.1037/1082-989X.1.2.178

Reznick, JS.; Gibbons, JL.; Johnson, MO.; McDonough, PM. Behavioral inhibition in a normative sample. In: Reznick, JS., editor. Perspectives on behavioral inhibition. Chicago, IL: The University of Chicago Press; 1989. p. 25-49.

Riese ML. Temperament stability between the neonatal period and 24 months. Developmental Psychology. 1987a; 23:216–222.10.1037/0012-1649.23.2.216

Riese ML. Longitudinal assessment of temperament from birth to 2 years: A comparison of full-term and preterm infants. Infant Behavior and Development. 1987b; 10:347–363.10.1016/0163-6383(87)90022-1

Riese ML. Temperament in full-term and preterm infants: Stability over ages 6 to 24 months. Developmental and Behavioral Pediatrics. 1988; 9:6–1.10.1097/00004703-198802000-00002

Roberts BW, DelVecchio WF. The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. Psychological Bulletin. 2000; 126:25–30.10.1037/0033-2909.126.1.3

Roberts JE, Burchinal M, Durham M. Parents' report of vocabulary and grammatical development of African American preschoolers: Child and environmental associations. Child Development. 1999; 70:92–106.10.1111/1467-8624.00008 [PubMed: 10191517]

Rothbart MK. Measurement of temperament in infancy. Child Development. 1981; 52:569–578.10.2307/1129176

Rothbart MK. Longitudinal observation of infant temperament. Developmental Psychology. 1986; 22:356–365.10.1037/0012-1649.22.3.356

Rothbart, MK. Becoming who we are. New York: Guilford Press; 2011.

Rothbart, MK.; Derryberry, D. Development of individual differences in temperament. In: Lamb, ME.; Brown, AL., editors. Advances in developmental psychology. Vol. 1. Hillsdale, NJ: Erlbaum; 1981. p. 37-86.

Rothbart, MK.; Derryberry, D.; Hershey, K. Stability of temperament in childhood: Laboratory infant assessment to parent report at seven years. In: Molfese, VJ.; Molfese, DL., editors. Temperament and personality development over the life span. Hillsdale, NJ: Erlbaum; 2000. p. 85-119.

Rothbart, MK.; Derryberry, D.; Posner, MI. A psychobiological approach to the development of temperament. In: Bates, JE.; Wachs, TD., editors. Temperament: Individual differences at the interface of biology and behavior. Washington, DC: American Psychological Association; 1994. p. 83-116.

Rothbart MK, Goldsmith HH. Three approaches to the study of infant temperament. Developmental Review. 1985; 5:237–260.10.1016/0273-2297(85)90012-7

Rowe DC, Jacobson KC, Van den Oord EJCG. Genetic and environmental influences on vocabulary IQ: Parental education level as moderator. Child Development. 1999; 70:1151–1162.10.1111/1467-8624.00084 [PubMed: 10546338]

Rueda, MR. Effortful control. In: Zentner, M.; Shiner, R., editors. The handbook of temperament. New York: Guilford Press; 2012. p. 145-167.

Saudino KJ. Sources of continuity and change in activity level in early childhood. Child Development. 2012; 83:266–281.10.1111/j.1467-8624.2011.01680.x [PubMed: 22103336]

Saudino, KJ.; Wang, M. Quantitative and molecular genetic studies of temperament. In: Zentner, M.; Shiner, RL., editors. Handbook of temperament. New York: Guilford Press; 2012. p. 315-346.

Schwartz CE, Wright CI, Shin LM, Kagan J, Rauch SL. Inhibited and uninhibited infants "grown up": Adult amygdalar response to novelty. Science. 2003; 300(5627):1952–1953.10.1126/science.1083703 [PubMed: 12817151]

Seifer R, Schiller M, Sameroff AJ, Resnick S, Riordan K. Attachment, maternal sensitivity, and temperament during the first year of life. Developmental Psychology. 1996; 32:3–11.10.1037/0012-1649.32.1.12

Shiner RL, Buss KA, McClowry SG, Putnam SP, Saudino KJ, Zentner M. What is temperament now? Child Development Perspectives. 2012; 6:436–444.10.1111/j.1750-8606.2012.00254.x

Shiner, RL.; Caspi, A. Temperament and the development of personality traits, adaptations, and narratives. In: Zentner, M.; Shiner, RL., editors. Handbook of temperament. New York: Guilford Press; 2012. p. 497-516.

Stoolmiller M. Implications of the restricted range of family environments for estimates of heritability and nonshared environment in behavior–genetic adoption studies. Psychological Bulletin. 1999; 125:392–409.10.1037/0033-2909.125.4.392 [PubMed: 10414224]

Strelau, J. The regulative theory of temperament as a result of East-West influences. In: Kohnstamm, GA.; Bates, JE.; Rothbart, MK., editors. Temperament in childhood. Oxford, England: Wiley; 1989. p. 35-48.

Suitor JJ, Sechrist J, Plikuhn M, Pardo ST, Gilligan M, Pillemer K. The role of perceived maternal favoritism in sibling relations in midlife. Journal of Marriage and Family. 2009; 71:1026–1038.10.1111/j.1741-3737.2009.00650.x [PubMed: 20104251]

Szyf M, Bick J. DNA methylation: A mechanism for embedding early life experiences in the genome. Child Development. 2012; 84:49–57.10.1111/j.1467-8624.2012.01793.x [PubMed: 22880724]

Thomas, A.; Chess, S. Temperament and development. Oxford, England: Brunner/Mazel; 1977.

Thomas, A.; Chess, S.; Birch, HG.; Hertzig, ME.; Korn, S. Behavioral individuality in early childhood. New York: New York University; 1963.

Tronick EZ, Scanlon KB, Scanlon JW. A comparative analysis of the validity of several approaches to the scoring of the behavior of the preterm infant. Infant Behavior & Development. 1985; 8:395–411.10.1016/0163-6383(85)90004-9

Wachs, TD.; Kohnstamm, GA., editors. Temperament in context. Mahwah, NJ: Erlbaum; 2001.

Wang Y, Dix T. Patterns of depressive parenting: Why they occur and their role in early developmental risk. Journal of Family Psychology. 2013; 27:884–895.10.1037/a0034829 [PubMed: 24294931]

Winstanley A, Sperotto RG, Putnick DL, Cherian S, Bornstein MH, Gattis M. Consistency of maternal cognitions and principles across the first five months following preterm and term deliveries. Infant Behavior and Development. 2014; 37:760–771.10.1016/j.infbeh.2014.09.005 [PubMed: 25459794]

Worobey J, Blajda VM. Temperament ratings at 2 weeks, 2 months, and 1 year. Developmental Psychology. 1989; 25:257–263.10.1037/0012-1649.25.2.257

Table 1

Study 1: Descriptive Statistics and Stabilities Across Time for Infant Behavior Questionnaire (IBQ) Factors and Scales

	Descriptive Statistics						Stability		
	2 months		5 months		13 months		2–5 months	5–13 months	2–13 months
	M	*SD*	*M*	*SD*	*M*	*SD*	*r*	*r*	*r*
Positive Affectivity	3.61	.81	4.48	.67	4.52	.68	.56***	.59***	.44***
Smiling and Laughter	3.14	1.14	4.52	.82	5.08	.66	.41***	.56***	.29*
Soothability	4.32	.97	5.11	.78	5.10	.89	.43***	.46***	.23*
Duration of Orienting	3.36	1.20	3.79	1.08	3.37	1.13	.53***	.39***	.23
Negative Affectivity	3.22	.57	3.21	.52	3.65	.54	.58***	.53***	.37***
Activity Level	3.24	.77	3.81	.68	4.28	.81	.54***	.45***	.35**
Fear	2.40	.88	2.47	.77	2.92	.69	.47***	.44***	.15
Distress to Limitations	4.01	.72	3.35	.66	3.75	.69	.47***	.39***	.28*
Average							.47***	.45***	.25***

Note. IBQ scale scores range from 1 to 7 (1 = *never*, 2 = *very rarely*, 3 = *less than half the time*, 4 = *about half the time*, 5 = *more than half the time*, 6 = *almost always*, 7 = *always*).

* $p \leq .05$.

** $p \leq .01$.

*** $p \leq .001$.

Table 2

Study 1: Stabilities Across Time for Infant Behavior Questionnaire (IBQ) Factors and Scales by Gender

	Stability					
	2–5 months		5–13 months		2–13 months	
	Girls	Boys	Girls	Boys	Girls	Boys
Positive Affectivity	.63***	.55***	.70***	.40*	**.73*** *a***	**.08**
Smiling and Laughter	.46**	.44**	.64***	.39*	**.55** *b*	**−.05**
Soothability	.52***	.30	.51**	.39*	**.52*** *c***	**−.06**
Duration of Orienting	.59***	.54***	.49**	.20	.31	.18
Negative Affectivity	.40*	.71***	.51***	.55***	.31	.41*
Activity Level	.38*	.60***	.46**	.38*	.31	.34*
Fear	.36*	.59***	.52***	.35*	.09	.22
Distress to Limitations	.48**	.46**	.33	.45**	.24	.32
Average	.46***	.49***	.49***	.36***	.34***	.16***

Note. Bolded coefficients differ.

a $z = 3.47$, $p \leq .001$.

b $z = 2.74$, $p \leq .01$.

c $z = 2.60$, $p \leq .01$.

* $p \leq .05$.

** $p \leq .01$.

*** $p \leq .001$.

Table 3

Study 2: Descriptive Statistics and Stabilities Across Time for Infant Behavior Questionnaire (IBQ) Factors and Scales

	Descriptive Statistics				*Stability*
	6 months		**12 months**		
	M	***SD***	***M***	***SD***	***r***
Positive Affectivity	5.01	0.75	4.78	0.71	.67***
Smiling and laughter	5.42	0.86	5.36	0.75	.63***
Soothability[a]	5.38	0.85	5.30	0.88	.47***
Duration of orienting	4.23	1.17	3.65	1.15	.56***
Negative Affectivity	3.40	0.61	3.71	0.59	.59***
Activity level	4.41	0.87	4.49	0.89	.56***
Fear	2.59	0.85	2.99	0.76	.45***
Distress to limitations	3.19	0.77	3.65	0.78	.57***
Average					.54***

Note. IBQ scale scores range from 1 to 7 (1 = *never*, 2 = *very rarely*, 3 = *less than half the time*, 4 = *about half the time*, 5 = *more than half the time*, 6 = *almost always*, 7 = *always*).

[a]Five term infants had missing data on Soothability at either one or two ages, Little's MCAR test: $\chi^2(32) = 40.51$, $p = .14$, and 1 preterm infant had missing data on Soothability at 6 months, Little's MCAR test: $\chi^2(11) = 6.83$, $p = .81$. As these data were all missing completely at random, the missing data points were imputed using the Expectation-Maximization algorithm (Dempster, Laird, & Rubin, 1977).

***$p \leq .001$.

Table 4

Study 2: Stabilities from 6 to 12 Months for Infant Behavior Questionnaire (IBQ) Factors and Scales by Gender, Birth Order, Term Status, and Socioeconomic Status (SES)

	Gender		Birth Order		Term Status		SES	
	Girls	**Boys**	**Firstborns**	**Laterborns**	**Terms**	**Preterms**	**Lower**	**Higher**
Positive Affectivity	.72***	.61***	.63***	.71***	.68***	.66***	.69***	.66***
Smiling and laughter	.66***	.59***	.59***	.66***	.67***	.59***	.61***	.65***
Soothability	.44***	.48***	.40***	.51***	.47***	.47***	.50***	.42***
Duration of orienting	.59***	.53***	.52***	.59***	.56***	.56***	.56***	.54***
Negative Affectivity	.60***	.58***	.57***	.61***	.59***	.57***	.58***	.61***
Activity level	.52***	.61***	.56***	.57***	.54***	.59***	**.42*** *b***	**.62***
Fear	.47***	.40***	.43***	.44***	.48***	.39***	.46***	.44***
Distress to limitation	**.66*** *a***	**.48***	.55***	.60***	.59***	.51***	.55***	.59***
Average	.56***	.52***	.51***	.56***	.55***	.52***	.52***	.55***

Note. Bolded coefficients differ.

[a] $z = 2.46$, $p \leq .05$.

[b] $z = 2.36$, $p \leq .05$

*** $p \leq .001$.