Latent profile and cluster analysis of infant temperament: Comparisons across person-centered approaches

By: Maria A. Gartstein, Amanda Prokasky, Martha Ann Bell, Susan Calkins, David J. Bridgett, Julia J. Braungart-Rieker, Esther M. Leerkes, Carol L. Cheatham, Rina D. Eiden, Krystal D. Mize, Nancy Aaron Jones, Gina Mireault, and Erich Seamon

Gartstein, M., Prokasky, A., Bell, M.A., Calkins, S., Bridgett, D., Braunguart-Rieker, J., Leerkes, E.M., Cheatham, C., Eiden, R., Mize, K., Jones, N., Mireault, G., & Seamon, E. (2017). Latent profile and cluster analysis of infant temperament: Comparisons across person-centered approaches. *Developmental Psychology, 53,* 1811-1825.

Abstract:

There is renewed interest in person-centered approaches to understanding the structure of temperament. However, questions concerning temperament types are not frequently framed in a developmental context, especially during infancy. In addition, the most common person-centered techniques, cluster analysis (CA) and latent profile analysis (LPA), have not been compared with respect to derived temperament types. To address these gaps, we set out to identify temperament types for younger and older infants, comparing LPA and CA techniques. Multiple data sets (N = 1,356; 672 girls, 677 boys) with maternal ratings of infant temperament obtained using the Infant Behavior Questionnaire–Revised (Gartstein & Rothbart, 2003) were combined. All infants were between 3 and 12 months of age (M = 7.85; SD = 3.00). Due to rapid development in the first year of life, LPA and CA were performed separately for younger (n = 731; 3 to 8 months of age) and older (n = 625; 9 to 12 months of age) infants. Results supported 3-profile/cluster solutions as optimal for younger infants, and 5-profile/cluster solutions for the older subsample, indicating considerable differences between early/mid and late infancy. LPA and CA solutions produced relatively comparable types for younger and older infants. Results are discussed in the context of developmental changes unique to the end of the first year of life, which likely account for the present findings.

Keywords: infant temperament | latent class analysis | cluster analysis

Article:

According to Rothbart’s psychobiological model, temperament represents constitutionally based individual differences in emotional, motor, and attentional reactivity, and in self-regulation, demonstrating consistency across situations and relative stability over time (Rothbart & Bates, 2006; Rothbart & Derryberry, 1981). The term “constitutional” emphasizes the connection between temperament and biology, including the link to underlying neurobehavioral systems, as well as genetic and epigenetic influences. Reactivity encompasses multiple domains of affectivity, with self-regulation, largely dependent on attentional functioning, serving to modulate reactive tendencies (Gartstein, Putnam, Aaron, & Rothbart, 2016). In the first year of life (especially early to mid-infancy), orienting attention plays a critical role, as executive functions supported by the frontal lobe maturation have not yet "come online" (Posner, Rothbart, Sheese, & Voelker, 2012). Along with more advanced attentional skills and capacity for regulation, significant increases in fear/behavioral inhibition were noted at the end of the first year of life, as for example, infants became slower, rather than faster, in reaching toward high-intensity toys (Rothbart, 1988). These increases in fearfulness have been demonstrated with respect to mean levels and individual trajectories, indicating considerable changes later in infancy (Gartstein & Rothbart, 2003; Gartstein, Hancock, & Iverson, 2017).

Temperament domains outlined on the basis of the psychobiological model have been examined primarily through a variable-centered/dimensional approach, wherein scales are combined into overarching factors. At the same time, fine-grained temperament dimensions are important in their own right, demonstrating unique predictive relationships with outcomes such as developmental psychopathology, sleep and eating/feeding problems (e.g., Gartstein, Potapova, & Hsu, 2014). For example, low levels of falling reactivity and soothability in infancy were associated with an increased risk for oppositional defiant disorder and callous-unemotional traits (Willoughby, Waschbusch, Moore, & Propper, 2011). Other investigators reported that fear and sadness made more substantial contributions to internalizing difficulties, whereas anger/frustration were related to both internalizing and externalizing problems at different ages (Lengua, 2006; Oldehinkel, Hartman, de Winter, Veenstra, & Ormel, 2004; Nigg, 2006). Although regulation and negative affect have received the most attention, positive affectivity distinctions also are important. For example, higher levels of low intensity pleasure (enjoyment of calm activities) may protect against internalizing and externalizing problems, whereas more high intensity pleasure (enjoyment of more stimulating activities) appears to convey risk for externalizing difficulties only (Gartstein, Putnam, & Rothbart, 2012).

The fine-grained focus of the present study is thus a function of important distinctions among more narrowly defined attributes, often combined for convenience (e.g., reducing the number of analyses) or due to sample size limitations, rather than theoretical reasons. In the context of person-centered techniques, fine-grained temperament attributes can be expected to result in more differentiated typologies, likely increasing effectiveness of classification. Refining classification is of interest in part because it could enhance targeting for temperament-based prevention efforts, identifying children who face high versus low levels of risk as a result of their temperament profiles. However, person-centered approaches have not been widely used to distinguish types based on fine-grained temperament attributes, particularly during infancy. Thus, the primary goal of this study was to identify typologies of infant temperament, at the same time comparing the two most widely used person-centered techniques, cluster analysis (CA) and latent profile analysis (LPA).

Person-Centered Approaches in Temperament: Cluster and Latent Profile Analyses

The study of children's temperament has a longstanding tradition of relying on typologies. Notably, Thomas and Chess (1977) identified three infant temperament types: difficult, easy, and slow-to-warm up, relying on parental perceptions of nine underlying fine-grained temperament dimensions. These temperament types have the inherent appeal of answering the question: "What kind of kid is she?" Yet the efforts to understand children's temperament within the psychobiological framework have relied primarily upon the variable-centered perspective; person-centered approaches, by comparison, have received relatively little attention (Zentner & Bates, 2008). A holistic interactionist perspective, wherein an individual is viewed as the unit of analysis, represents the conceptual foundation for person-centered approaches, with all variables considered simultaneously (von Eye & Bergman, 2003). Applying a person-centered perspective to child temperament in a quantitative manner requires that combinations of multiple temperament dimensions be considered. Typologies based on these combinations can be compared as to their ability to explain the observed pattern of results, differentiating between individuals.

Cluster analysis (CA) represents the most frequently used person-centered technique for the identification of temperament types. CA is a data-driven approach, which begins by randomly assigning cases to a specified number of clusters, and subsequently reassigning cases to minimize the distance to the cluster center (Mooi & Sarstedt, 2011). CA has produced mixed results with respect to the number of temperament types. Caspi and Silva (1995) identified five temperament types: undercontrolled, inhibited, confident, reserved, and well-adjusted, using CA with investigator behavior ratings provided for a sample of 1,037 3-year-old children. Sanson et al. (2009), on the other hand, derived four temperament types based on maternal ratings on the Child Temperament Questionnaire (CTQ; Thomas & Chess, 1977) in a sample of 1,662 3- to 4-year-olds: nonreactive/outgoing, high attention regulation, poor attention regulation, and reactive/inhibited. Also using CTQ maternal reports, Martin, Bridger, and Huttunen (2000) identified seven clusters inhibited, impulsive, highly emotional, typical, reticent, uninhibited, and passive. Recently, Prokasky et al. (2017) concluded that six temperament types were optimal using maternal report on the Children's Behavior Questionnaire (CBQ; Rothbart, Ahadi, Hershey, & Fisher, 2001): unregulated, bold, high reactive, average, well adjusted, and regulated, replicating these groups with independent samples. In the only investigation employing CA to derive temperament types from infancy to middle childhood, Komsi et al. (2006) identified three clusters: overcontrolled, undercontrolled, and resilient. These typologies were based on two broadly defined temperament dimensions: positive and negative affectivity, and the fine-grained attribute of activity level.

Variations of latent class analysis (LCA) were utilized to identify children's temperament typologies during toddlerhood and middle childhood. Relative to CA, latent class analysis (LCA) is a newer, model-based, person-centered approach that has started to gain use in identification of temperament types. LCA determines the optimal number of latent subsets of children who share similar patterns of temperament attributes based on scale scores. Using a variation of LCA for continuous variables, latent profile analysis (LPA), van den Akker, Dekovic, Prinzie, and Asscher (2010) identified three profiles (typical, fearful, and expressive) based on maternal ratings on the Toddler Behavior Assessment Questionnaire (TBAQ; Goldsmith, 1996). In a sample of 787 twin pairs (mean age = 7.4 years), Scott et al. (2016) employed twin factor mixture modeling (LCA which allows simultaneous modeling of profile and factor structure) and identified four temperament profiles (regulated/typical reactive, well-regulated/positive reactive, regulated/surgent, and dysregulated/negative reactive), with mother/father rating composites obtained via a modified CBQ.

In studies with infants, LCA has been applied to laboratory observations of reactivity at 4 months of age (*N* = 169; Loken, 2004). Results supported at least three temperament classes, corresponding to high reactive (high distress/activity, low smiling), low reactive (low distress/activity, high smiling), and a category characterized as "aroused" (low distress/high activity). More recently, Beekman et al. (2015) utilized LPA to identify temperament profiles when children were 9, 18, and 27 months of age (*N* = 561). Typical/low expressive, typical/expressive, negative reactive, and positive reactive profiles were identified at 9 months. Positive reactive, negative reactive, active reactive (marked by high levels of activity and above average levels of both pleasure and anger), and fearful profiles emerged at 18 and 27 months of age.

Although these person-centered findings may seem disparate at first, a number of themes emerge across existing studies. First, there is a consistent grouping marked by reactivity/negative affect, also reminiscent of "difficult temperament" (Beekman et al., 2015; Loken, 2004; Martin et al., 2000; Prokasky et al., 2017; Sanson et al., 2009; Scott et al., 2016). Another theme has to do with children being well regulated and/or presenting with high levels of positive affectivity, sometimes combined under labels referring to adjustment or resilience (Beekman et al., 2015; Caspi & Silva, 1995; Komsi et al., 2006; Prokasky et al., 2017; Scott et al., 2016). In addition, several typologies included fear-based groups (Beekman et al., 2015; Martin et al., 2000; Sanson et al., 2009; van den Akker, et al., 2010) and those defined by fearlessness (Martin et al., 2000) or undercontrol (Caspi & Silva, 1995; Komsi et al., 2006). Existing research has pointed to a number of themes, yet unanswered questions remain, in part due to the relatively limited scope of infant temperament attributes considered to date.

Beekman et al. (2015) and Komsi et al. (2006) assessed infant temperament using the Infant Behavior Questionnaire (IBQ; Rothbart, 1981), the predecessor of the IBQ-R, which includes only two regulation-related scales and one positive emotionality dimension. Loken (2004) just considered distress, smiling, and activity level aspects of infant temperament, measured using laboratory observations. The age range of infants in existing studies is restricted as well, as Loken (2004) collected temperament data at 4 months, Komsi et al. (2006) at 6 months, and Beekman et al. (2015) evaluated infants at 9 months of age. As a result, these studies are not able to inform about potential shifts in typologies that stem from changes in temperament at the end of the first year. Temperament typologies are likely not impervious to rapid developmental transitions, such as those evident between early/mid and late infancy. This developmental period is defined by marked locomotor advances (Shonkoff & Phillips, 2000), the emergence of different domains of reactivity, such as anger/frustration earlier in infancy (Carranza, Perez-Lopez, Gonzalez, & Martinez-Fuentes, 2000; Rothbart & Bates, 2006), and notable increases in fear at the end of the first year of life (Gartstein et al., 2017). Later infancy, relative to early/mid infancy, is also marked by the "coming online" of more advanced attentional capabilities (Bridgett, Burt, Edwards, & Deater-Deckard, 2015; Posner et al., 2012), linked with improved regulation. Thus, while existing studies employing CA and LCA are informative, additional work sensitive to developmental shifts in temperament at the end of the first year of life is needed.

Finally, no study to date has directly compared CA and LCA in the same sample of children to determine if these approaches result in comparable temperament typologies. Only two studies comparing CA and LCA solutions have been conducted. Eshghi, Haughton, Legrand, Skaletsky, and Woolford (2011) examined groupings of countries ($N = 160$) formed on the basis on 10 sociodemographic variables (per capita income, education, percent urban population, etc.). CA was found superior in terms of within-group homogeneity (i.e., producing types with the most similar members). DiStefano and Kamphaus (2006) compared CA and LCA deriving child behavioral adjustment types for 6- to 11-year-olds using teacher ratings. CA results supported seven, and LCA three, adjustment categories. Thus, questions concerning differences among these analytic techniques require consideration in deriving temperament types.

The Current Study

Given the relative dearth of research addressing temperament types in a fine-grained manner across infancy, the primary goal of this study was to identify infant temperament types based on the 14 IBQ-R Scale scores. Developmental considerations, including the overall rapid rate of growth in infancy (Shonkoff & Phillips, 2000) and noted changes in temperament (e.g., Gartstein et al., 2010), dictated that typologies be derived for younger and older infants separately. As CA has been most widely utilized to investigate temperament types, and LCA represents a less established approach, an additional aim of this study was to compare LCA and CA temperament types.

The literature is not consistent with respect to the number of clusters/classes; however, three or four have been typically reported for infants (Beekman et al., 2015; Komsi et al., 2006; Loken, 2004). On the basis of these findings, we tentatively anticipated identifying three to four profiles/clusters. Nevertheless, in light of the limited existing infant studies, more differentiated profiles/clusters identified among older children (e.g., Caspi & Silva, 1995; Prokasky et al., 2017), and our consideration of 14 IBQ-R Scales, up to eight profiles/clusters were evaluated. We also hypothesized differences among solutions derived for younger and older infants, because of considerable changes between early/mid infancy and the end of the first year. Moreover, solutions for older infants were expected to be more complex in nature, given that in prior work three classes were identified for 4-month-olds (Loken, 2004), with a four-profile solution deemed optimal for older infants (Beekman et al., 2015). In regards to the specific nature of hypothesized profiles/clusters, most relevant studies with infants (i.e., Beekman et al., 2015; Loken, 2004) suggest that high reactive (high distress, activity level, low smiling), low reactive (demonstrating an opposite pattern), and a aroused (low distress/high activity) types could be expected earlier in the first year. For older infants, types consistent with typical/low expressive, typical/expressive, negative reactive, positive reactive, active reactive, and/or fearful (Beekman et al., 2015) were expected.

Finally, with respect to the optimal solutions, we generally anticipated consistency across the two person-centered approaches. Nevertheless, limited evidence based on direct comparisons of CA and LCA techniques in nontemperament contexts suggests the possibility of some differences among types derived by these approaches, although sufficient specificity for a priori hypotheses is currently lacking. Thus, probabilities of participants' assignment to parallel profiles/clusters were compared via a chi-square test, and within-group homogeneity differences were considered in direct comparisons of LPA and CA solutions.

Method

Sample

Data sets were acquired by emailing researchers who had requested the IBQ-R or published research using the instrument between 2006 and 2011 (see Table 1 for additional demographic information). Only families with healthy infants were eligible to participate in the projects (samples of origin) providing IBQ-R data.

1. The first infant temperament data set (*n* = 410) was provided by the third and fourth authors. These data were collected in the context of a longitudinal study examining individual differences in cognition-emotion integration (Gartstein, Bell, & Calkins, 2014).

2. The second data set (*n* = 158), provided by the fifth author, included information collected when infants were 6 (*n* = 114), 8 (*n* = 95), 10 (*n* = 87), and 12 months of age (*n* = 79). This study addressed temperament development, parenting, and emerging behavior problems (Bridgett et al., 2009).

3. The third data set was contributed by the 6th author and included temperament ratings at 3 (*n* = 135), 5 (*n* = 127), 7 (*n* = 116), and 12 (*n* = 116) months of age, as described by Braungart-Rieker et al. (2014). This work focused on temperament, parent–child interactions, and attachment.

4. The fourth data set (*n* = 118) was provided by the seventh author, with the IBQ-R collected at 6 months for a study of temperament and mother-infant interactions (Parade & Leerkes, 2008).

5. The fifth data set (*n* = 86) was contributed by the eighth author, who obtained temperament ratings at 6 months for a study addressing nutrition and cognitive development (Cheatham & Sheppard, 2015).

6. The sixth data set, containing IBQ-R assessments when children were 9 months of age, was collected by the ninth author for a study examining the effects of prenatal tobacco exposure (see Eiden et al., 2015 for full sample description) on infant functioning. Only control group infants (*n* = 75) not exposed to tobacco in utero were included in the current study.

7. The seventh data set was contributed by the tenth and eleventh authors, who obtained infant temperament ratings (*n* = 85) when children were 7 to 12 months of age as part of ongoing research on the loss of maternal attention to a social-rival (Mize & Jones, 2012; Mize, Pineda, Blau, Marsh, & Jones, 2014).

8. The eighth data set, provided by the twelfth author, included monthly longitudinal data on thirty 3-month-old infants collected through 6 months, and again at 12 months of age (Mireault et al., 2012). This research examined infant humor perception.

9. The final three samples, contributed by the first author, were recruited for several studies addressing temperament development. The first sample of 147 children was assessed at 4 (*n* = 114), 6 (*n* = 114), 8 (*n* = 114), 10 (*n* = 102), and 12 (*n* = 101) months of age, with portions of this dataset described in Gartstein et al. (2010) and Gartstein et al. (2013). The second sample (*N* = 140) was equally divided across four age groups: 3 months (*n* = 35), 6 months (*n* = 35), 9 months (*n* = 35), and 12 months (*n* = 35; Gartstein & Bateman, 2008). The third sample (*n* = 9) participated in a parental guidance temperament intervention, wherein caregivers were provided with information based on the psychobiological model (Iverson et al., 2014).

These data sets obtained by multiple laboratories were collectively utilized in the present study (*N* = 1,356). All infants were between 3 and 12 months of age (*M* = 7.85; *SD* = 3.00), and were equally distributed across sex (females: *n* = 672; males: *n* = 677). A number of studies relied on longitudinal evaluations. In these instances, in order to maintain independence of observations, only one assessment point per child was included in the combined data set. About an equal number of cases were selected from each of the different phases of the longitudinal studies. For example, for the first dataset (*n* = 410), 205 infants contributed 5-month temperament scores, whereas the remainder (*n* = 205) of the sample contributed 10-month data. To use all of the available data, if a participant completed only a portion of the longitudinal assessments, their data were selected from a completed assessment (i.e., not from one of the missing evaluations).

Measure: Infant Behavior Questionnaire-Revised (IBQ-R; Gartstein & Rothbart, 2003)

The IBQ-R is a parent-report measure of infant temperament for use between 3 and 12 months of age. The 191 items (rated on a 7-point Likert-type scale) represent 14 subscales, which in turn form three overarching factors. The *surgency* factor consists of approach (app), vocal reactivity (vr), high intensity pleasure (hp), smiling and laughter (sl), activity level (act), and perceptual sensitivity (ps) subscales. The *negative emotionality* factor consists of sadness (sad), distress to limitations (dl), fear, and falling reactivity (fall) subscales. Finally, the *regulatory capacity/orienting* factor includes low intensity pleasure (lp), cuddliness/affiliation (cud), duration of orienting (do), and soothability (sooth) subscales. Each item reflects the frequency of occurrence of reactivity/regulation during the prior week (most items), or 2 weeks, for less common events. The IBQ-R has consistently demonstrated good psychometric properties with mothers, fathers, and international samples, with Cronbach's alphas ranging from .77 to .96 (Gartstein & Rothbart, 2003; Parade & Leerkes, 2008). Interrater reliability, concurrent/predictive and construct validity, have been demonstrated for IBQ-R Scales (Gartstein & Bateman, 2008; Gartstein, Knyazev, & Slobodskaya, 2005; Gartstein & Marmion, 2008; Parade & Leerkes, 2008).

Table 1. *Demographic Metrics as a Function of Sample*

Demographic Metric	Sample								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Size	410	158	135	118	86	75	48	30	296[1]
Ages represented	5,10 months	4, 6, 8, 10, 12, 14 months	3, 5, 7, 12, 14 months	6 months	2, 9 months	6 months	7–12 months, mean age = 9.6 months	3, 4, 5, 6, 12 months	4, 6, 8, 10, 12 months 3, 6, 9, 12 months 3–12 months (mean age = 7.2 months)
Girls/boy	209/201	69/89	58/77	53/65	44/42	41/34	26/22	16/14	147/149
Ethnicity	6.3% Hispanic[2]	4.4% Hispanic	1.5% Hispanic	1% Hispanic	1% Hispanic	18% Hispanic	7.3% Hispanic	0% Hispanic	2% Hispanic
Race	77% Caucasian, 13.7% African American, 1% Asian, 7.8% Multiracial, .5% Other	92.2% Caucasian	85.9% Caucasian, 2.2% African American, .7% Asian, 8.9% Multiracial	77% Caucasian	89% Caucasian, 6% African American, 1% Native American, 3% Multiracial	30% Caucasian, 52% African American; of these, 8% reported > single race	72% Caucasian, 1.2% African American, 6.1% Asian, 13.4% Multiracial	100% Caucasian	89% Caucasian, 2% African American, 4% Asian, 3% Multiracial
Education	99% completed high school (HS[3]), 6% technical degree, 42% bachelor's degree, 22% graduate degree	*M* = 15.1 range = 8 to 25 years	95% completed HS, 59.3% completed college	67% had college degrees	4% HS only, 8% some college, 88% earned bachelor's degree or higher	26% below HS; 60% had HS, 10% had some college, 4% with a vocational or technical training degree	70.6% earned a college or graduate degree	*M* = 15.5 range = 12 to 19 years	97% completed HS, 80% earned a bachelor's degree
Income	—[4]	mean family income = $60,859.58	$10,000 to $150,000; median = $45,000	$6,000 to $190,000; *M* = $70,000	$25,000 to >$100,000	—	—	annual household income $78,000/year (*SD* = $51,400)	annual household income $7,000 to >$75,000; 80% > $16,000; 33% > $50,000
Parental age	Mothers' mean age = 29	Mothers' mean age = 33	Mothers' mean age = 29	Mothers' mean age = 28	Mothers' mean age = 45	Mothers' mean age = 25	Mothers' mean age = 31.5	Mothers' mean age = 32	Mothers' mean age = 33

[1]Gartstein sample included three separate data collections, reported together in this table. [2]Hispanic ethnic category is separate from racial categories, thus sum > 100%. [3]High School (HS). [4]Income data not available for every included sample.

Analytic Strategy

Latent profile analysis

LPA was accomplished using Mplus Version 7 (Muthén & Muthén, 2012), with full information maximum likelihood estimation employed to accommodate missing data (Enders, 2013). LPA provides indices to discern the optimal number of subsets of infants who share similar patterns of maternal ratings concerning fine-grained temperament attributes. As recommended, a number of indices were taken into account simultaneously in making decisions about the optimal number of profiles (Lanza & Cooper, 2016). We considered the Akaike Information Criteria (AIC), Sample-Size Adjusted Bayesian Information Criteria (BIC), and Entropy measures in comparing models, attempting to minimize the AIC and BIC, and producing a strong entropy measure (approaching 1.00). The Entropy index reflects effectiveness of categorization based on posterior probabilities, which were also examined in this study. The Lo, Mendell, Rubin Likelihood Ratio Test (Lo, Mendell, & Rubin, 2001) was considered in determining if an additional profile improved the overall model fit (e.g., comparing two-profile to a three-profile model). Infant age and sex, and sample of origin, were considered as covariates and retained in the final models if they were associated with significant paths to the latent variable reflecting profile membership. Multiple solutions (up to eight profiles) were considered.

Cluster analysis

In line with previous studies relying on clustering techniques in discerning temperament types (e.g., Caspi & Silva, 1995; Sanson et al., 2009), a two-step clustering procedure was employed. If the number of underlying clusters within the data is unknown—a circumstance encountered in the current investigation, a hierarchical cluster analysis is typically performed as an initial step. Therefore, an exploratory hierarchical cluster analysis was initially conducted on a random sample of 200 cases. This preliminary analysis is performed to help guide decisions about the number of clusters within a dataset, and it is also recommended to examine multiple cluster solutions with the entire sample (Hair, Anderson, Tatham, & Black, 1995; Mooi & Sarstedt, 2011). For this reason, and to parallel the LPA procedures, a series of k-means cluster analyses (two through eight cluster solutions) were performed using the entire sample.

Comparison of LPA and CA

Optimal LPA and CA solutions were plotted to enable interpretation and comparison across these techniques. ANOVAs with IBQ-R Scales as dependent variables were conducted to compare types resulting from the optimal solutions, with profiles and clusters compared in turn, to further characterize and contrast these approaches. Follow-up pairwise tests—independent-group *t* tests with Bonferroni corrections—were subsequently performed and used to inform decisions regarding profile labels. Specifically, attributes associated with statistically significant differences across all profiles/clusters, and differentiating types as highest/lowest relative to other groups, were prioritized in naming profiles. In addition, we compared LPA and CA classification outcomes in terms of: (a) agreement between LPA and CA with respect to participant assignment to parallel profiles/clusters, and (b) within-group homogeneity differences. Chi-square tests were conducted to discern agreement between LPA and CA solutions in terms of case assignment. Average Euclidean distances for LPA and CA solutions were also computed and compared via matched-pair *t* tests to determine which method resulted in greater homogeneity (Eshghi et al., 2011).

Accounting for development

As infant development was expected to play a role in differentiation between temperament types, the sample was divided by age, with parallel LPA and CA analyses conducted separately for younger and older infants. Specifically, the median split (8 months) was used to divide the sample into groups of younger (*n*= 731; 3 to 8 months of age) and older infants (*n* = 625; 9 to 12 months of age).

Results

Descriptive statistics and Cronbach's alpha for IBQ-R Scales were computed using the entire sample, and separately for younger and older infants, using SPSS Version 23 (see Table 2).

Table 2. *IBQ-R Scale Means, Standard Deviations, and Cronbach's Alphas for the Entire Sample (N = 1,356), Younger (n = 731), and Older (n = 625) Infants*

IBQ-R Scale(Skewness)[1]	Entire sample Mean (*SD*)	Cronbach's alpha	Younger infants Mean (*SD*)	Younger infants Cronbach's alpha	Older infants Mean (*SD*)	Older infants Chronbach's alpha
Activity	4.43 (.85)	.76	4.13 (.83)	.76	4.69 (.78)	.72
Distress to limitations	3.70 (.87)	.79	3.12 (.74)	.75	4.06 (.83)	.76
Fear	2.64 (1.00)	.88	2.21 (.80)	.86	3.03 (1.00)	.87
Duration of orienting	3.99 (1.05)	.82	4.04 (1.04)	.82	3.94 (1.06)	.83
Smiling and laughter	4.99 (1.00)	.79	4.70 (.83)	.82	5.07 (.92)	.78
High intensity pleasure	5.83 (.80)	.84	5.57 (.83)	.82	6.07 (.69)	.84
Low intensity pleasure	5.14 (.91)	.83	5.24 (.88)	.83	5.03 (.93)	.84
Soothability	4.64 (.93)	.82	4.56 (.89)	.80	4.72 (.97)	.83
Falling reactivity	4.78 (.97)	.81	4.81 (1.02)	.83	4.76 (.92)	.79
Cuddliness	5.24 (.94)	.87	5.46 (.97)	.87	5.03 (.87)	.85
Perceptual sensitivity	4.12 (1.18)	.85	3.72 (1.16)	.84	4.48 (1.06)	.84
Sadness	3.46 (.89)	.80	3.34 (.87)	.82	3.56 (.89)	.79
Approach	5.04 (1.18)	.87	4.50 (1.25)	.88	5.56 (.84)	.83
Vocal reactivity	4.88 (1.05)	.84	4.44 (1.05)	.85	5.28 (.88)	.82

Note. Overall sample: range –.95 to .70, $M = –.08$; Younger age group: range –1.09 to .93, $M = –.12$; Older age group: range –1.23 to .48, $M = –.07$. Infant Behavior Questionnaire (IBQ-R).

Latent Profile Analysis (LPA)

All LPA models were initially evaluated with the sample of origin, and infant age and sex as covariates. Infant sex was not retained, with results indicating that a three-profile model was optimal for younger infants (Figures 1a). According to the Lo, Mendell, Rubin Likelihood Ratio Test, an additional profile improved the overall model fit from a two-profile to a three-profile model, but not from a three-profile to a four-profile solution (see Table 3). BIC and AIC were lowered with additional profiles, however the Entropy and posterior probabilities were optimized in the three-profile model. Although more complex solutions (up to eight profiles) were examined, these did not result in improved fit. All path coefficients from infant age to the latent profile variables were significant for the younger age group.

For older infants, the five-profile solution was deemed optimal based on a number of indicators (Figure 1b). Specifically, the five-profile solution resulted in the minimum BIC and AIC values, while also maximizing Entropy. None of the five profiles was associated with an *n* < 10% of the total *N*, further supporting this model as superior overall, despite a nonsignificant Lo, Mendell, and Rubin Likelihood Ratio Test. In contrast to our findings with the younger age group, path coefficients from infant age to the latent profiles were not significant for the older infants.

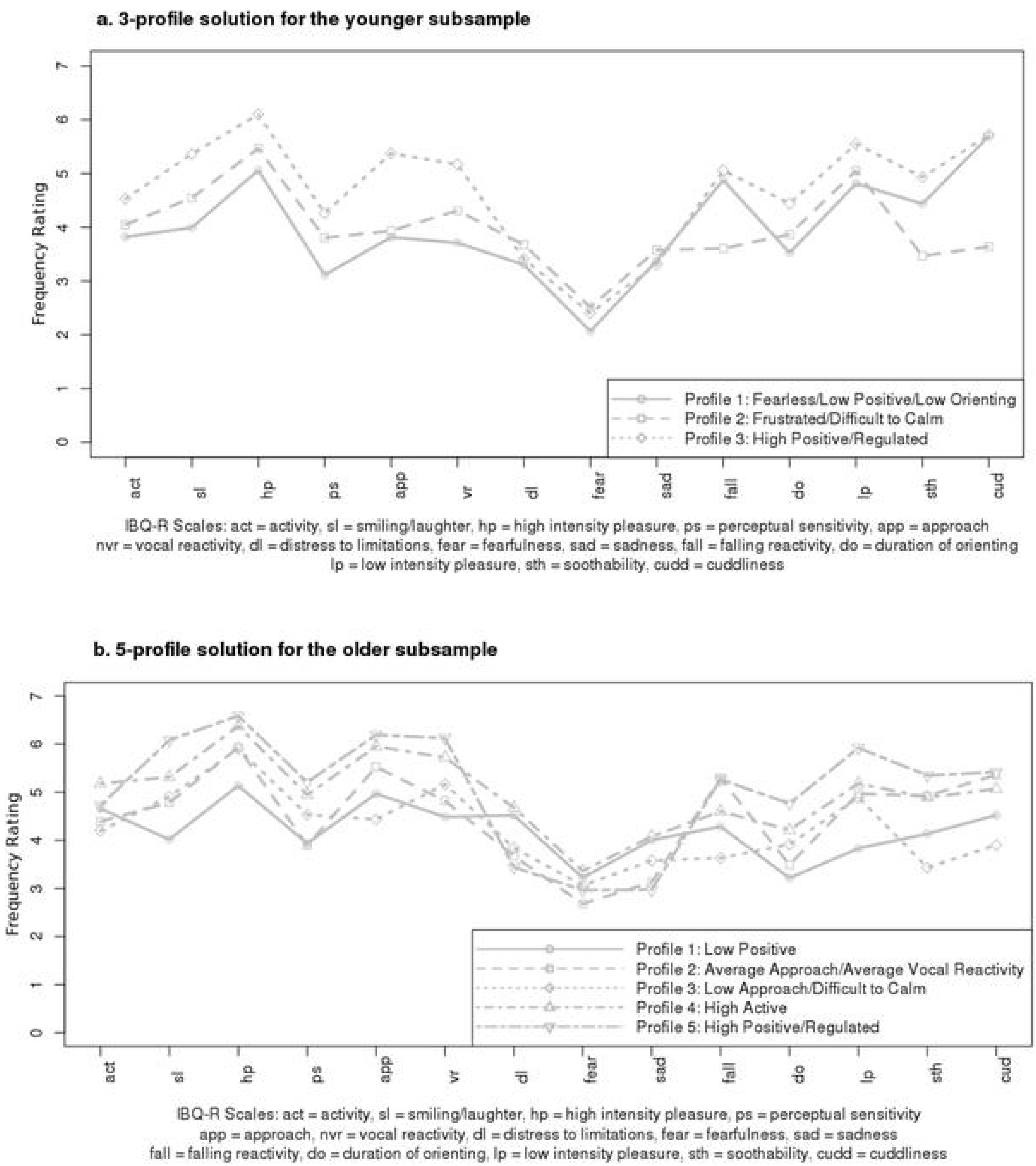

Figure 1. Three-profile solution for younger and five-profile solution for older infants.

Table 3. *Latent Profile Analysis: Assessing Model Fit for Older and Younger Infants*

	1 Class	2 Classes	3 Classes	4 Classes	5 Classes
Younger infants					
AIC	32710.97	26710.99	26148.24	25805.95	25513.03
BIC	32931.50	26954.49	26506.61	26279.17	26101.12
Sample adjusted BIC	32779.08	26786.19	26258.93	25952.11	25694.67
Entropy	na	.834	.887	.853	.851
Lo, Mendell, Rubin Test	na	2 vs. 1	3 vs. 2	4 vs. 3	5 vs. 4
		Value *p* = .00	Value *p* = .04	Value *p* = .31	Value *p* = .12
N for each class	C1 = 731	C1 = 325 C2 = 406 P1 = –.57 *p* = .00	C1 = 282 C2 = 92 C3 = 357 P1 = –.60 *p* = .00 P2 = .07 *p* = .79	C1 = 253 C2 = 238 C3 = 81 P1 = .30 *p* = .11 P2 = .31 *p* = .46 P3 = .92 *p* = .00	C1 = 171 C2 = 232 C3 = 76 C4 = 127 C5 = 125 P1 = .30 *p* = .43 P2 = .44 *p* = .51 P3 = .86 *p* = .00 P4 = .93 *p* = .03
Average posterior probability		.948	.953	.923	.916
Older infants					
AIC[1]	27367.02	21731.75	21263.29	21003.17	20757.72
BIC[2]	27571.15	21962.52	21600.56	21446.94	21308.00
Adjusted BIC	27425.11	21797.42	21359.27	21129.46	20914.32
Entropy	na	.832	.845	.829	.849
Lo, Mendell, Rubin Test	na	2 vs. 1	3 vs. 2	4 vs. 3	5 vs. 4
		Value *p* = .00	Value *p* = .07	Value *p* = .24	Value *p* = .15
N for each class	C1[3] = 625	C1 = 232 C2 = 393	C1 = 98 C2 = 243 C3 = 284	C1 = 208 C2 = 183 C3 = 92 C4 = 142	C1 = 78 C2 = 173 C3 = 79 C4 = 186 C5 = 109
Covariate paths/effects	na	P1 = .04 *p* = .81	P1 = –.12 *p* = .72 P2 = –.16 *p* = .60	P1 = –.14 *p* = .54 P2 = .07 *p* = .85 P3 = –.03 *p* = .89	P1 = –.08 *p* = .82 P2 = .19 *p* = .49 P3 = .48 *p* = .31 P4 = –.03 *p* =.93
Average posterior probability		.954	.929	.916	.912

[1] Alaike Information Criteria (AIC). [2] Bayesian Information Criteria (BIC). [3] C1-C5 values represent sizes of classes/profiles. [4] P1-P4 path coefficients for infant age (covariate) as a predictor of profile membership (i.e., latent profile variable).

Consistent with prior investigations (e.g., Beekman et al., 2015), interpretation of types was guided by visual inspection, in conjunction with statistical tests. Resulting types were compared via ANOVAs, for younger and older age groups, respectively. All IBQ-R dimensions reliably differentiated between the three temperament types derived on the basis of LPA for younger infants (mean η^2 = .23; range .01–.52). Cuddliness (η^2 = .41) and vocal reactivity (η^2 = .42) were associated with the largest effects sizes. In the older age group, all comparisons indicated significant profile differences (mean η^2 = .28; range .07–.41). Vocal reactivity (η^2 = .41) and approach (η^2 = .39) produced the largest effects.

Follow-up *t* tests with Bonferroni corrections (see online supplementary tables) and Figure 1a indicated that in the younger subsample the lowest scores on all but one of the positive affectivity scales marked Profile 1. That is, Profile 1 was significantly different from Profiles 2 and 3 on smiling and laughter, activity level, high intensity pleasure, perceptual sensitivity and vocal reactivity, with approach scores significantly different from Profile 3 only. Significantly lower levels of fear, duration of orienting, and low intensity pleasure were also observed for Profile 1, thus labeled *fearless/low positive/low orienting*. According to the follow-up tests and Figure 1a, Profile 2 was characterized by a pattern of average positive affectivity, with approach significantly different from Profile 3 only. This type was also associated with the highest levels of distress to limitations, as well as the lowest falling reactivity, soothability, and cuddliness. Given the salience of frustration, coupled with a lack of responsiveness to caregivers' attempts to calm, and an inability to lower one's own level of arousal, Profile 2 was labeled *frustrated/difficult to calm*. Profile 3 was described as *high positive/regulated*, as infants in this group received significantly higher scores for positive affectivity and all but one scale related to regulation (cuddliness was significantly different from Profile 2 ratings only), coupled with average or low negative emotionality dimension ratings.

In the older age group, Profile 1 was best described as *low positive,* because of significantly lower levels of smiling and laughter, high and low intensity pleasure, as well as vocal reactivity. Profile 2 scores were generally unremarkable, and this type was referred to as *average approach/average vocal reactivity*, as these scales significantly differentiated Profile 2 infants from all remaining groups, placing them in the middle. Profile 3 was characterized by low levels of falling reactivity, soothability, cuddliness, and approach, and was thus referred to as *low approach/difficult to calm*. Profile 4 can be described as *active*, as activity level was the only scale associated with an extreme (highest) value for this type and statistically significant differences relative to the other four profiles. Follow-up *t* tests and Figure 1b indicated that the fifth profile was distinguished by high scores on a number of scales addressing positive affectivity and regulation: smiling and laughter, high intensity pleasure, perceptual sensitivity, vocal reactivity, duration of orienting, low intensity pleasure, and soothability, and was thus labeled *high positive/regulated*.

Cluster Analysis

Hierarchical cluster analysis, linking pairs of cases with the smallest distance between them until all cases are linked into one cluster composed of all cases (agglomerative clustering), was performed first, consistent with previous studies (e.g., Caspi & Silva, 1995; Prokasky et al., 2017; Sanson et al., 2009). Visual inspection of the resulting dendogram indicated that a three-cluster solution provided optimal fit for the younger infants. Consistent with findings from LPA, a five-cluster solution fit well for the older subsample. Model fit indices are not available with CA techniques, thus, K-means clustering analyses results were evaluated conceptually, providing an interpretable three-cluster solution for the younger, as well as a five-cluster solution for the older age group. Decisions concerning the number of clusters are generally based on “practical judgment or theoretical foundations” (Hair, Anderson, Tatham, & Black, 1995, p. 443), and these solutions were deemed optimal in light of the available literature supporting both three- and five-cluster typologies. These findings parallel reported LPA results with respect to the number of types. The clustering results were also graphed (Figures 2a and 2b), and examined for content similarity to the LPA solutions.

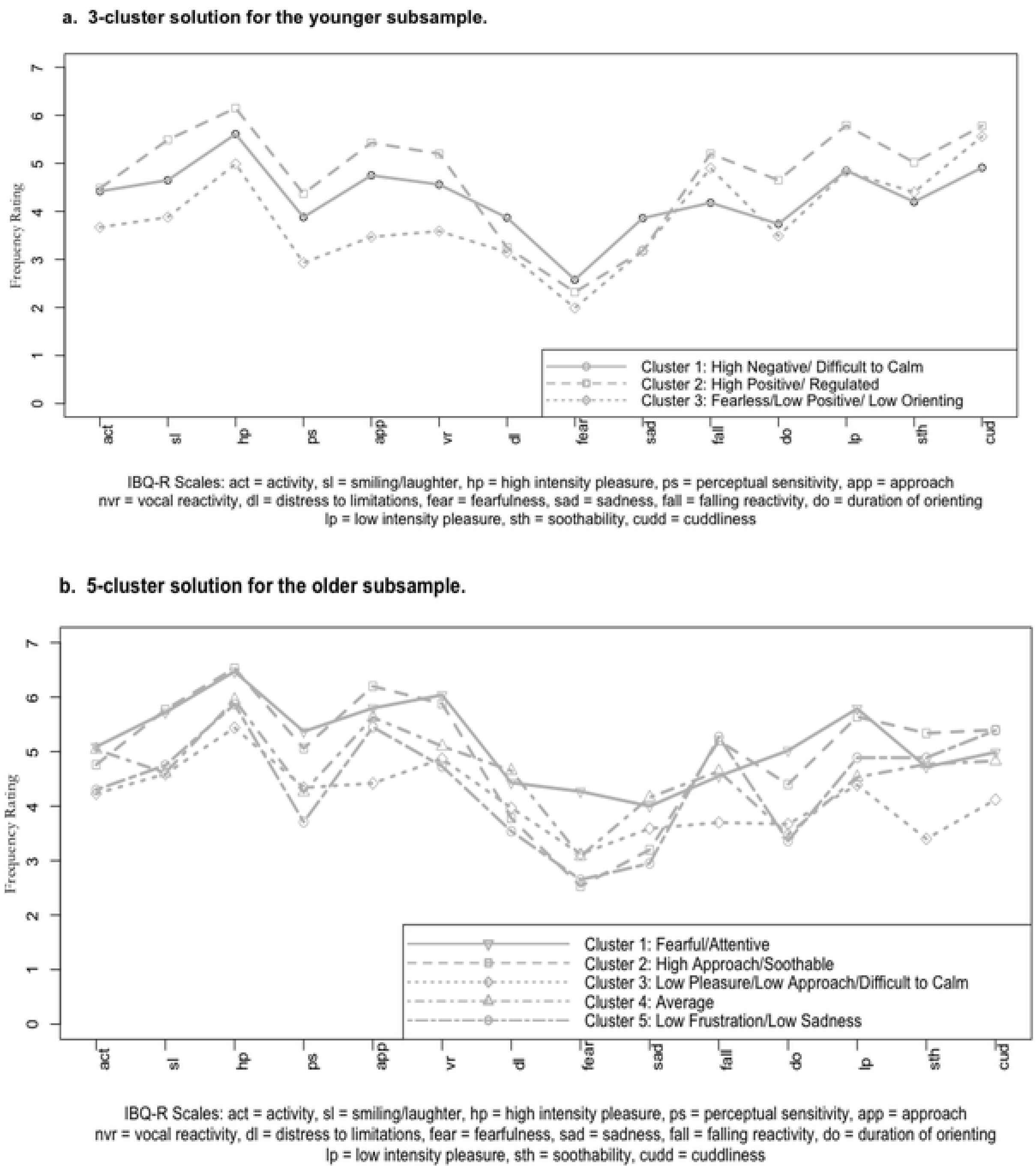

Figure 2. Three-cluster solution younger and five-cluster solution for older infants.

Resulting types were compared via ANOVAs, conducted separately for younger and older age groups. All 14 IBQ-R dimensions were able to reliably differentiate between three temperament types derived on the basis of CA for younger infants (mean $\eta^2 = .25$; range .07–.45). Approach ($\eta^2 = .45$), vocal reactivity ($\eta^2 = .42$), and smiling and laughter ($\eta^2 = .42$) were associated with the largest effect sizes. Similarly, all IBQ-R dimensions reliability differentiated between five temperament clusters for the older infants (mean $\eta^2 = .33$; range .20–.45). The largest effects sizes were noted for approach ($\eta^2 = .45$) and soothability ($\eta^2 = .41$).

The three-cluster solution for younger infants paralleled LPA types (Figure 2a). For younger infants, Cluster 1 was labeled *high negative/difficult to calm*, because of highest scores on all negative emotionality scales (distress to limitations, fear, and sadness), along with the lowest levels of falling reactivity, soothability, and cuddliness; significantly different from the other two profiles according to follow-up *t* tests with Bonferroni corrections (see online supplementary tables). Cluster 2 was referred to as *high positive/regulated*, because infants in this group consistently received the highest positive affectivity and regulation-related ratings, including: smiling and laughter, high intensity pleasure, perceptual sensitivity, approach, vocal reactivity, falling reactivity, duration of orienting, low intensity pleasure, soothability and cuddliness. Cluster 3 was marked by low positive affectivity/surgency scores, including: activity level, smiling and laughter, high intensity pleasure, perceptual sensitivity, approach, and vocal reactivity, as well as low levels of fear and duration of orienting, and was thus labeled *fearless/low positive/low orienting*.

There were also notable similarities to the LPA solution for the older subsample (Figure 2b). Cluster 1 had the highest scores on fear and duration of orienting, significantly different from the remaining four clusters, and was thus labeled as the *fearful/attentive* type. Cluster 2 infants had the highest scores on approach and soothability, also statistically significant in differentiating this cluster from all others, with this type referred to as *high approach/soothable* as a result. Cluster 3 was associated with significantly different and lowest scores on high intensity pleasure, approach, falling reactivity, soothability, and cuddliness, consistent with the assigned *low pleasure/low approach/difficult to calm* label. Cluster 4 was average on all dimensions, as none of the scales were associated with statically significant differences relative to all remaining clusters, so was labeled *average*. Finally, Cluster 5 scores represented the lowest levels of distress to limitations and sadness, and this group was accordingly referred to as the *low frustration/low sadness* type. Perceptual sensitivity and low intensity pleasure also significantly differentiated this cluster from all others; however, these scores were midrange relative to the other clusters, and thus not referenced in the label.

In summary, there are important similarities between LPA and CA solutions, wherein both approaches indicated potentially key differences in types that coalesce in early/mid versus late infancy. Parallel profiles/clusters could be identified (see Table 4), and despite some notable differences, both person-centered techniques point to the importance of surgency/positive affectivity and regulation-related dimensions in discriminating among temperament types. For example, the combination of low falling reactivity, soothability and cuddliness, referred to as *difficult to calm*, often contributed to differentiation among types across approaches. Although cluster and profile solutions for younger infants involved fearlessness, distinctions based on high levels of fear for the older age group emerged only in the context of CA follow-up *t* tests. LPA and CA results indicated that approach gained importance in type differentiation among older infants.

Table 4. *Matching Profiles and Clusters for Younger and Older Age Groups*

Profiles	Clusters
Younger	
Profile 1: Fearless/low positive/low orienting (*n* = 282)[1]	Cluster 3: Fearless/low positive/low orienting (*n* = 234)[1]
Profile 2: Frustrated/difficult to calm (*n* = 92)[2]	Cluster 1: High negative/difficult to calm (*n* = 207)[2]
Profile 3: High positive/regulated (*n* = 357)[3]	Cluster 2: High positive/regulated (*n* = 282)[3]
Older	
Profile 1: Low positive (*n* = 78)[1]	Cluster 4: Average (*n* = 142)[1]
Profile 2: Average approach/average vocal reactivity (*n* = 173)[2]	Cluster 5: Low frustration/low sadness (*n* = 134)[2]
Profile 3: Low approach/difficult to calm (*n* = 79)[3]	Cluster 3: Low pleasure/low approach/difficult to calm (*n* = 100)[3]
Profile 4: High active (*n* = 186)[4]	Cluster 1: Fearful/attentive (*n* = 97)[4]
Profile 5: High positive/regulated (*n* = 109)[5]	Cluster 2: High approach/soothable (*n* = 151)[5]

Note. Superscript 1-5 represent matching clusters.

Direct Comparison of LPA and CA Results

Younger infants

The chi-square test comparing distributions of cases assigned to matched types (see Table 4) based on the LPA versus the CA classification of the younger subsample was significant (χ^2 = 60.49; $p < .001$). Thus, the LPA and CA resulted in relatively comparable profiles/clusters, yet each method classified some children differently, producing a significant χ^2 test and different sizes for profiles/clusters. Specifically, 30% (*n* = 113) of children classified into the *frustrated/difficult to calm* profile were also classified into the *high negative/difficult to calm* cluster, whereas 50.7% (*n* = 71) of children categorized into the *high positive/regulated* profile were also assigned to the *high positive/regulated* type based on CA. Finally, 49.3% (*n* = 102) of children classified into the *fearless/low positive/low orienting* profile were also classified into the *fearless/low positive/low orienting* cluster. Thus, despite apparent similarity in the nature of profiles/clusters, younger infants were often not classified in the same manner across parallel LPA and CA types.

Distances from individual cases to their profile/cluster center were calculated using the Euclidean distance as an index of within-group homogeneity, and compared via matched-pair *t* tests. The average distance between cases in the *fearless/low positive/low orienting* (Profile 1) and their profile center was 3.16 (*SD* = .80), and between cases in the *fearless/low positive/low orienting* (Cluster 3) and their cluster center was 3.20 (*SD* = .80), with these means not statistically different from one another: $t(514) = -.591$ $p = .55$. For the *frustrated/difficult to calm* (Profile 2), the average distance between cases and their profile center was 3.17 (*SD* = .87), whereas an average distance of 2.96 (*SD* = .81) separated *high negative/difficult to calm* (Cluster 1) cases and their cluster center. *Frustrated/difficult to calm* and *high negative/difficult to calm* Euclidean distances were significantly different from one another: $t(297) = 2.05$, $p = .041$. For the *high positive/regulated* (Profile 3) cases, the average distance from the profile center was 2.95 (*SD* = .75), with the average distance to center of 2.91 (*SD* = .69) for its matched Cluster 2 (*high positive/regulated*). The average distances to center were not significantly different for this profile/cluster matched pair: $t(637) = 0.798$, $p = .425$. Thus, in the younger subsample, LPA and

CA performed equally well in creating homogeneous groups, with the exception of the *difficult to calm* type comparison between matched Profile 2 and Cluster 1, wherein CA outperformed LPA in terms of minimizing homogeneity.

Older infants

The chi-square test comparing distributions of cases resulting from the LPA versus the CA classification for the older subsample into matching profiles/clusters (see Table 4) was significant ($\chi^2 = 60.81$; $p < .001$), indicating some children were classified differently. Specifically, 52.6% ($n = 41$) of children who were assigned to the *low positive* (Profile 1) were classified into its closest match, the *average* (Cluster 4). For the *average approach/average vocal reactivity* LPA type (Profile 2), 74% ($n = 128$) of children were classified into the corresponding *low frustration/low sadness* (Cluster 5). Next, 81% ($n = 64$) of children classified into the *low approach/difficult to calm* (Profile 3) were also classified into the *low pleasure/low approach/difficult to calm* (Cluster 3). For the *high active* (Profile 4) children, 35.6% ($n = 66$) were classified into the most closely matching *fearful/attentive* (Cluster 1). Finally, 78.9% ($n = 86$) of the *High Positive/Regulated* (Profile 5) infants were assigned to the corresponding *high approach/soothable* (Cluster 2). Notably, despite an apparent divergence of profile/cluster labels reflecting most highly discriminating scales, a greater consistency in classification across LPA and CA techniques was observed for older infants.

Distances from individual cases to their profile/cluster center were calculated using the Euclidean distance and compared. The average distance between cases in the *low positive* (Profile 1) and the profile center was 2.80 ($SD = .83$), and between cases in the *average* (Cluster 4) and the cluster center was 2.65 ($SD = .62$). These distances were not statistically different from one another: $t(218) = 1.507$ $p = .113$. The mean distance between *average approach/average vocal reactivity* (Profile 2) cases and their center was 2.68 ($SD = .71$), whereas for *low frustration/low sadness* (Cluster 5) cases the average distance was 2.64 ($SD = .70$), which were not statistically different from one another: $t(305) = .456$, $p = .679$. The *low approach/difficult to calm* profile (Profile 3) average Euclidean distance was 2.41 ($SD = 0.63$), with the average distance of 2.67 ($SD = 0.84$) between cases in the *low pleasure/low approach/difficult to calm* (Cluster 3) and the cluster center. These distances were significantly different from one another: $t(177) = -2.26$, $p = .025$. The average distance between cases in the *high active* (Profile 4) and their profile center was 2.77 ($SD = .67$), with the average distance of 2.79 ($SD = .60$) between cases in the matched *fearful/attentive* (Cluster 1) and the cluster center. These distances were not significantly different from each other: $t(281) = -.211$, $p = .83$. Finally, the average distance between cases and the *high positive/regulated* (Profile 5) center was 2.69 ($SD = .64$), whereas the average distance between cases and the *high approach/soothable* (Cluster 2) center was 2.64 ($SD = .62$). These distances were not significantly different: $t(258) = .593$, $p = .554$. Thus, in the older subsample, LPA and CA performed equally well in creating homogeneous groups, except in the case of the *low approach/difficult to calm* (Profile 3) and *low pleasure/low approach/difficult to calm* (Cluster 3) comparison, wherein LPA outperformed CA.

Discussion

Understanding temperament types, particularly during infancy, informs other areas of developmental science (e.g., developmental psychopathology), given links between early temperament and diverse outcomes later in life. To that end, this study aimed to discern the nature of temperament typologies across the first year of life, via a fine-grained approach, on the basis of the psychobiological model of temperament (Rothbart & Bates, 2006; Rothbart & Derryberry, 1981). We used LPA and CA to characterize optimal solutions, directly comparing these two person-centered techniques. To achieve our goals, analyses were performed separately for younger and older infants due to considerable developmental changes that occur over the first year of life.

LPA and CA person-centered approaches produced results often consistent with our expectations, based on existing theoretical and empirical models of temperament types. For younger infants, we anticipated three- or four-profile/cluster models, consistent with prior research conducted early in the first year of life (e.g., Loken, 2004). LPA and CA results obtained for younger infants indicated three-profile/cluster solutions provided the best fit, supporting this hypothesis. However, the observed types were more complex than those in earlier studies, likely due to our consideration of 14 fine-grained temperament domains. The *frustrated/difficult to calm* and *high negative/difficult to calm* types resemble a number of previously identified categories, including high reactive (Loken, 2004), negative reactive (Beekman et al., 2015), dysregulated/negative reactive (Scott et al., 2016), and "difficult" (Thomas & Chess, 1977; Thomas, Chess, Birch, Hertzig, & Korn, 1963). The *high positive/regulated* type was consistent with the "easy" Thomas and Chess (1977) classification, and reminiscent of the resilient cluster (above-average activity level and positive affectivity, below-average negative emotionality) identified by Komsi et al. (2006) and others (e.g., Robins, John, Caspi, Moffitt, & Stouthamer-Loeber, 1996). Types reflecting low levels of fear have been previously reported, typically in the context of an exuberant constellation (e.g., Degnan et al., 2011), yet our LPA- and CA-based types marked by fearlessness also involve low levels of positive affectivity, so are different in content.

As anticipated, the picture emerging for older infants was more complex, with five-profile/cluster solutions based on LPA and CA results. The resulting types were differentiated largely by positive affectivity and regulation-related dimensions—a pattern distinct from prior investigations. The *low positive* type, demonstrating low smiling and laughter, high and low intensity pleasure, and vocal reactivity, bears similarity to the typical/low expressive profile, characterized by somewhat below average reactivity scores (Beekman et al., 2015); although, the *low positive* designation was less common in this study. The *average approach/average vocal reactivity* profile and the *average* cluster seem consistent with previously identified typical/average types (Martin et al., 2000; Prokasky et al., 2017; van den Akker et al., 2010), as both present with largely unremarkable scores (i.e., lacking any extreme values), and contain sizable proportions of the older subsample. The *low approach/difficult to calm* profile and *low pleasure/low approach/difficult to calm* cluster resemble the "difficult temperament" constellation (Thomas & Chess, 1977; Thomas et al., 1963), due to low levels of positive emotionality and regulation-related attributes. The *high active* profile has its closest counterpart in the "aroused" (low distress/high activity) type, previously derived with younger infants (Loken, 2004). This profile is also somewhat consistent with the typical types, as it reflects the most commonly assigned category, according to LPA results.

It remains to be seen if the *high positive/regulated* type confers protection with respect to adjustment/psychopathology in a manner similar to the resilient classification, and whether or not its CA counterpart for the older age group (*high approach/soothable*, Cluster 2) is associated with similar effects. The *fearful/attentive* cluster appears somewhat consistent with previously derived categories reflecting high levels of fear (Beekman et al., 2015; van den Akker et al., 2010). Although only observed in the context of CA, this type suggests fearfulness and orienting attention work in tandem to differentiate infants later in the first year of life, consistent with conceptual and empirical models that link behavioral inhibition and attention/information processing (Pine & Fox, 2015). The *low frustration/low sadne*ss cluster does not appear to have a direct parallel in the literature; however, seems reminiscent of the typical category (Martin et al., 2000; van den Akker et al., 2010), as this group presents with only two extreme scores (both speaking to low levels of negative emotions), and reflects a sizable portion of the older subsample.

There was considerable agreement for LPA and CA solutions concerning the important role of surgency-related scales, and age-related distinctions in their respective contributions to differences between types. According to both person-centered techniques, approach gained importance in type differentiation for older infants. This pattern of results could be linked to developmental changes in behavioral inhibition (hesitation in approach to novelty) at the end of the first year of life (e.g., Rothbart, 1988), serving to make individual differences in approach more salient. Alternatively, the increasing importance of approach could be a product of environmental demands, as positive affectivity has been shown to be context dependent (Goldsmith, Lemery, Buss, & Campos, 1999). Environmental demands on positive affectivity change at the end of the first year of life as infants are expected to become more effective interactional partners, likely making approach tendencies (which include excitement, positive anticipation, and enthusiasm) more critical in distinguishing among temperament types. Approach and positive affect/joy have been associated with the development of behavior problems. Whereas greater capacity to experience positive affectivity appears protective with respect to internalizing problems, depression in particular (Dougherty, Klein, Durbin, Hayden, & Olino, 2010), the risk for externalizing problems increases with higher levels of surgency, especially associated impulsivity (Frick & Morris, 2004). There are developmental considerations as well, insofar as greater expression of positive emotionality/joy and approach in the first year of life can be described as largely protective (e.g., facilitating the development of self-regulation and school readiness; Gartstein, Putnam, & Kliewer, 2016; Gartstein, Slobodskaya, Putnam, & Kinsht, 2009). Additional research is required to determine if types high in positive affectivity convey protection, or potential risk, with impulsive tendencies potentially becoming problematic during the toddler period.

Developmental differences between younger and older infants, apparent using both LPA and CA approaches, are notable. First, more complex five-profile/cluster solutions were deemed optimal for the older infants, compared with the three-profile/cluster typology derived for the younger subsample. The *high positive/regulated* type was the only one to emerge across the first year of life, via LPA only for older infants. Of interest, CA produced a fearfulness-based type at the end of the first year of life, which coincides in timing with rapid developmental shifts in this domain of temperament (Gartstein et al., 2017). Infant age made more notable contributions in the context of younger infants' solutions, suggesting greater developmental heterogeneity between 3 and 8 months of age, relative to the 9- to 12-month period. The latter may simply be a function of the wider age range in the younger subsample, yet developmental factors could also play a role. Early/mid infancy marks the emergence of most reactive tendencies, with consolidation of traits beginning at the end of the first year of life (Gartstein et al., 2016).

Our results further illustrate the importance of considering fine-grained temperament indicators. An exclusive emphasis on overarching factors obscures potentially critical distinctions among more narrowly specified attributes, and this study provided several illustrations. Falling reactivity, soothability, and cuddliness, emerged as critical to discrimination among types across both person-centered techniques. Although soothability and cuddliness are components of the IBQ-R regulatory capacity/orienting factor, falling reactivity loaded negatively with negative emotionality dimensions (Gartstein & Rothbart, 2003). Falling reactivity reflects infants' ability to lower their level of arousal, and has been shown to function as a marker of early self-regulation (Bridgett, Burt, Laake, & Oddi, 2013; Gartstein, Slobodskaya, Putnam, & Kirchhoff, 2013). Reducing temperament scores to overarching factors would have prevented falling reactivity from contributing to group membership in a synergistic manner with soothability and cuddliness, forming the *difficult to calm* constellation. Importantly, the fine-grained level of analysis resulted in a more prominent role played by temperament attributes associated with surgency/positive affectivity, frequently understudied relative to negative emotionality and regulation, and likely facilitated the emergence of relatively complex types, especially for older infants.

Inconsistencies among types derived based on LPA versus CA person-centered techniques were noted with respect to most discriminating combinations of scales, case assignment to parallel types, and within-group homogeneity. Although profile/cluster labels varied more widely for the older subsample, case assignment differences were more pronounced in the younger age group. The latter may be a function of developmental factors, for example, with more coalesced types observed later in infancy. This pattern of results may have also occurred because infant age, included as a covariate in LPA (but not CA), played a more notable role during early/midinfancy. Discrepancies in the obtained results could also be attributed to the fact that CA is completely "bottom-up" or data-driven, whereas LPA represents a model-based technique. The main goal of CA is to derive the most homogenous subgroups possible, regardless of any underlying assumptions about the data. LPA on the other hand, assumes that class/profile membership is the driving force behind individuals' scores; therefore, the latent class model plays a critical role in subgroup formation. Overall, our pattern of results echoed findings reported by DiStefano and Kamphaus (2006) and Eshghi et al. (2011) wherein different methodologies (LPA vs. CA) provided some similar types, but also discrepant groupings, likely leading to different interpretations of the underlying structure of the data. LCA appears to offer an advantage in providing fit indices lacking in CA, wherein decisions concerning the number of types are largely conceptually driven. However, future work should compare these approaches with respect to their ability to predict distal outcomes, critical to determining their relative advantages.

Limitations

Limitations have to be taken into account when interpreting the results of the present investigation, and should be addressed by future research. Most importantly, the data set aggregated for this investigation did not allow us to compare the LPA and CA derived solutions with respect to their ability to predict distal outcomes, such as emerging behavior problems and/or school readiness. This limitation is a function of the fact that the data were obtained from multiple projects conducted to answer a variety of research questions, and large-scale longitudinal investigations are required to address this issue. Replication and extension is also necessary to ensure that the reported types are more than products of the derivation algorithms, and represent biologically based groupings. This investigation relied exclusively on parent-report of temperament, thus questions remain regarding profiles that would emerge on the basis of observational data and/or physiological indicators. Moreover, our conclusions regarding developmental effects are admittedly tentative due to the cross-sectional nature of this study, and future research should employ multiple observation points across infancy using the same sample. Finally, with any typology there is a risk of labeling misuse, albeit seemingly minimal for the profiles/clusters derived in this study. To be fair, the terms "difficult to calm" and "frustrated" are negative in connotation, yet these labels are not diagnostic in nature. Moreover, types derived in this study were largely reflective of distinctions based in positive affect and regulation, not excessive distress or difficultness.

Conclusion

In conclusion, results of this study suggest potentially important differences in the temperament types based on scores obtained in early/mid versus late infancy, and indicate that different person-centered computational approaches yield somewhat parallel temperament types. Among LPA advantages, the ability to include covariates is likely most significant to similar investigations, including data from multiple sources and/or developmentally heterogeneous groups of children. Although CA is primarily concerned with increasing within-group homogeneity, our results indicated it does not uniformly outperform LPA in this regard. Importantly, results of this study have implications for developmental pathways that involve temperament, such as risk/protection with respect to developmental psychopathology. Types defined by higher levels of attributes linked with positive affectivity and regulation could exhibit more optimal outcomes across different areas of functioning (e.g., achievement, peer relationships), to be determined by future research. Similarly, being a member of the fear-related cluster emerging at the end of the first year of life may bode risk for internalizing symptoms, although this type was not derived in LPA analyses. Our findings provide additional evidence supporting the importance of fine-grained temperament distinctions, and further underscore the importance of considering developmental shifts in temperament typologies.

Footnotes

1. Results available from the first author upon request.

References

Beekman, C., Neiderhiser, J. M., Buss, K. A., Loken, E., Moore, G. A., Leve, L. D., . . .Reiss, D. (2015). The development of early profiles of temperament: Characterization, continuity, and etiology. *Child Development*, *86*, 1794–1811. 10.1111/cdev.12417

Braungart-Rieker, J. M., Zentall, S., Lickenbrock, D. M., Ekas, N. V., Oshio, T., & Planalp, E. (2014). Attachment in the making: Mother and father sensitivity and infants' responses during the Still-Face Paradigm. *Journal of Experimental Child Psychology*, *125*, 63–84. 10.1016/j.jecp.2014.02.007

Bridgett, D. J., Burt, N. M., Edwards, E. S., & Deater-Deckard, K. (2015). Intergenerational transmission of self-regulation: A multidisciplinary review and integrative conceptual framework. *Psychological Bulletin*, *141*, 602–654. 10.1037/a0038662

Bridgett, D. J., Burt, N. M., Laake, L. M., & Oddi, K. B. (2013). Maternal self-regulation, relationship adjustment, and home chaos: Contributions to infant negative emotionality. *Infant Behavior and Development*, *36*, 534–547. 10.1016/j.infbeh.2013.04.004

Bridgett, D. J., Gartstein, M. A., Putnam, S. P., McKay, T., Iddins, E., Robertson, C., . . .Rittmueller, A. (2009). Maternal and contextual influences and the effect of temperament development during infancy on parenting in toddlerhood. *Infant Behavior and Development*, *32*, 103–116. 10.1016/j.infbeh.2008.10.007

Carranza, J. A., Perez-Lopez, J., Gonzalez, C., & Martinez-Fuentes, M. T. (2000). A longitudinal study of temperament in infancy: Stability and convergence of measures. *European Journal of Personality*, *14*, 21–37. 10.1002/(SICI)1099-0984(200001/02)14:1<21::AID-PER367>3.0.CO;2-A

Caspi, A., & Silva, P. A. (1995). Temperamental qualities at age three predict personality traits in young adulthood: Longitudinal evidence from a birth cohort. *Child Development*, *66*, 486–498. 10.2307/1131592

Cheatham, C. L., & Sheppard, K. W. (2015). Synergistic effects of human milk nutrients in the support of infant recognition memory: An observational study. *Nutrients*, *7*, 9079–9095.

Degnan, K. A., Hane, A. A., Henderson, H. A., Moas, O. L., Reeb-Sutherland, B. C., & Fox, N. A. (2011). Longitudinal stability of temperamental exuberance and social-emotional outcomes in early childhood. ***Developmental Psychology***, *47*, 765–780. 10.1037/a0021316

DiStefano, C., & Kamphaus, R. W. (2006). Investigating subtypes of child development a comparison of cluster analysis and latent class cluster analysis in typology creation. *Educational and Psychological Measurement*, *66*, 778–794. 10.1177/0013164405284033

Dougherty, L. R., Klein, D. N., Durbin, C. E., Hayden, E. P., & Olino, T. M. (2010). Temperamental positive and negative emotionality and children's depressive symptoms: A longitudinal prospective study from age three to age ten. *Journal of Social and Clinical Psychology*, *29*, 464–490. 10.1521/jscp.2010.29.4.462

Eiden, R. D., Molnar, D. S., Granger, D. A., Colder, C. R., Schuetze, P., & Huestis, M. A. (2015). Prenatal tobacco exposure and infant stress reactivity: Role of child sex and maternal behavior. *Developmental Psychobiology*, *57*, 212–225. 10.1002/dev.21284

Enders, C. K. (2013). Dealing with missing data in developmental research. *Child Development Perspectives*, *7*, 27–31. 10.1111/cdep.12008

Eshghi, A., Haughton, D., Legrand, P., Skaletsky, M., & Woolford, S. (2011). Identifying groups: A comparison of methodologies. *Journal of Data Science*, *9*, 271–291.

Frick, P. J., & Morris, A. S. (2004). Temperament and developmental pathways to conduct problems. *Journal of Clinical Child and Adolescent Psychology*, *33*, 54–68. 10.1207/S15374424JCCP3301_6

Gartstein, M. A., & Bateman, A. E. (2008). Early manifestations of childhood depression: Influences of infant temperament and parental depressive symptoms. *Infant and Child Development*, *17*, 223–248. 10.1002/icd.549

Gartstein, M. A., Bell, M. A., & Calkins, S. D. (2014). EEG asymmetry at 10 months of age: Are temperament trait predictors different for boys and girls?*Developmental Psychobiology*, *56*, 1327–1340.

Gartstein, M. A., Bridgett, D. J., Rothbart, M. K., Robertson, C., Iddins, E., Ramsay, K., & Schlect, S. (2010). A latent growth examination of fear development in infancy: Contributions of maternal depression and the risk for toddler anxiety. ***Developmental Psychology***, *46*, 651–668. 10.1037/a0018898

Gartstein, M. A., Bridgett, D. J., Young, B. N., Panksepp, J., & Power, T. (2013). Origins of effortful control: Infant and parent contributions. *Infancy*, *18*, 149–183. 10.1111/j.1532-7078.2012.00119.x

Gartstein, M. A., Hancock, G. R., & Iverson, S. L. (2017). Positive affectivity and fear trajectories in infancy: Contributions of mother-child interaction factors. *Child Development*. Advance online publication. 10.1111/cdev.12843

Gartstein, M. A., Knyazev, G. G., & Slobodskaya, H. R. (2005). Cross-cultural differences in the structure of infant temperament: United States of America (U.S.) and Russia. *Infant Behavior and Development*, *28*, 54–61.

Gartstein, M. A., & Marmion, J. (2008). Fear and positive affectivity in infancy: Convergence/discrepancy between parent-report and laboratory-based indicators. *Infant Behavior and Development*, *31*, 227–238. 10.1016/j.infbeh.2007.10.012

Gartstein, M. A., Potapova, N., & Hsu, A. (2014). Infant temperament: Implications for later sleep and eating/feeding. *Journal of Reproductive and Infant Psychology*, *32*, 185–198. 10.1080/02646838.2013.853172

Gartstein, M. A., Putnam, S. P., Aaron, E., & Rothbart, M. K. (2016). Temperament and personality. In S.Matzman (Ed.), *Oxford handbook of treatment processes and outcomes in counseling psychology* (pp. 11–41). New York, NY: Oxford University Press.

Gartstein, M. A., Putnam, S., & Kliewer, R. (2016). Do infant temperament characteristics predict core academic abilities in preschool-aged children?*Learning and Individual Differences*, *45*, 299–306. 10.1016/j.lindif.2015.12.022

Gartstein, M. A., Putnam, S. P., & Rothbart, M. K. (2012). Etiology of preschool behavior problems: Contributions of temperament attributes in early childhood. *Infant Mental Health Journal*, *33*, 197–211. 10.1002/imhj.21312

Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development*, *26*, 64–86. 10.1016/S0163-6383(02)00169-8

Gartstein, M. A., Slobodskaya, H. R., Putnam, S. P., & Kinsht, I. A. (2009). A cross-cultural study of infant temperament: Predicting preschool effortful control in the United States of America (U.S.) and Russia. *European Journal of* ***Developmental Psychology***, *6*, 337–364. 10.1080/17405620701203846

Gartstein, M. A., Slobodskaya, H. R., Putnam, S. P., & Kirchhoff, C. (2013). Cross-cultural differences in the development of behavior problems: Contributions of infant temperament in Russia and U.S. *International Journal of Developmental Science*, *7*, 95–104.

Goldsmith, H. H. (1996). Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development*, *67*, 218–235. 10.2307/1131697

Goldsmith, H. H., Lemery, K. S., Buss, K. A., & Campos, J. J. (1999). Genetic analyses of focal aspects of infant temperament. ***Developmental Psychology***, *35*, 972–985. 10.1037/0012-1649.35.4.972

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice Hall.

Iverson, S., & Gartstein, M. A. (2014, October). *The development and piloting of a temperament-based intervention for infants: "Knowing Your Baby Matters"*. Presentation at the Occasional Temperament Conference, Lincoln, NE.

Komsi, N., Räikkönen, K., Pesonen, A. K., Heinonen, K., Keskivaara, P., Järvenpää, A. L., & Strandberg, T. E. (2006). Continuity of temperament from infancy to middle childhood. *Infant Behavior and Development*, *29*, 494–508. 10.1016/j.infbeh.2006.05.002

Lanza, S. T., & Cooper, B. R. (2016). Latent class analysis for developmental research. *Child Development Perspectives*, *10*, 59–64. 10.1111/cdep.12163

Lengua, L. J. (2006). Growth in temperament and parenting as predictors of adjustment during children's transition to adolescence. ***Developmental Psychology***, *42*, 819–832. 10.1037/0012-1649.42.5.819

Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*, 767–778. 10.1093/biomet/88.3.767

Loken, E. (2004). Using latent class analysis to model temperament types. *Multivariate Behavioral Research*, *39*, 625–652. 10.1207/s15327906mbr3904_3

Martin, R. P., Bridger, R., & Huttunen, M. (2000, July). *Prediction of mental health and somatic health hospital admissions from early childhood temperament types*. Paper presented at the International Congress of Psychology, Stockholm, Sweden.

Mireault, G., Sparrow, J., Poutre, M., Perdue, B., & Macke, L. (2012). Infant humor perception from 3- to 6-months and attachment at one year. *Infant Behavior and Development*, *35*, 797–802. 10.1016/j.infbeh.2012.07.018

Mize, K. D., & Jones, N. A. (2012). Infant physiological and behavioral responses to loss of maternal attention to a social-rival. *International Journal of Psychophysiology*, *83*, 16–23. 10.1016/j.ijpsycho.2011.09.018

Mize, K. D., Pineda, M., Blau, A. K., Marsh, K., & Jones, N. A. (2014). Infant physiological and behavioral responses to a jealousy provoking condition. *Infancy*, *19*, 338–438. 10.1111/infa.12046

Mooi, E., & Sarstedt, M. (2011). Cluster analysis. In *A concise guide to market research: The process, data, and methods using IBM SPSS statistics* (pp. 237–284). New York, NY: Springer.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus 7*. Los Angeles, CA: Author. Retrieved from http://statmodel.com

Nigg, J. T. (2006). Temperament and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *47*, 395–422. 10.1111/j.1469-7610.2006.01612.x

Oldehinkel, A. J., Hartman, C. A., DeWinter, A. F., Veenstra, R., & Ormel, J. (2004). Temperament profiles associated with internalizing and externalizing problems in preadolescence. *Development and Psychopathology*, *16*, 421–440.

Parade, S. H., & Leerkes, E. M. (2008). The reliability and validity of the infant behavior questionnaire-revised. *Infant Behavior and Development*, *31*, 637–646. 10.1016/j.infbeh.2008.07.009

Pine, D. S., & Fox, N. A. (2015). Childhood antecedents and risk for adult mental disorders. *Annual Review of Psychology*, *66*, 459–485. 10.1146/annurev-psych-010814-015038

Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2012). Control networks and neuromodulators of early development. ***Developmental Psychology***, *48*, 827–835. 10.1037/a0025530

Prokasky, A., Rudasill, K., Molfese, V. J., Putnam, S., Gartstein, M., & Rothbart, M. K. (2017). Identifying child temperament types using cluster analysis in three samples. *Journal of Research in Personality*, *67*, 190–201. 10.1016/j.jrp.2016.10.008

Robins, R. W., John, O. P., Caspi, A., Moffitt, T. E., & Stouthamer-Loeber, M. (1996). Resilient, overcontrolled, and undercontrolled boys: Three replicable personality types. *Journal of Personality and Social Psychology*, *70*, 157–171. 10.1037/0022-3514.70.1.157

Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, *52*, 569–578. 10.2307/1129176

Rothbart, M. K. (1988). Temperament and the development of inhibited approach. *Child Development*, *59*, 1241–1250. 10.2307/1130487

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development*, *72*, 1394–1408. 10.1111/1467-8624.00355

Rothbart, M. K., & Bates, J. E. (2006). Temperament. In W.Damon, R.Lerner, & N.Eisenberg (Eds.), Handbook of child psychology: Vol. *3*. *Social, emotional, and personality development* (6th ed., pp. 99–166). New York, NY: Wiley.

Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. L.Lamb & A. L.Brown (Eds.), *Advances in* ***developmental psychology*** (Vol. *1*, pp. 37–86). New York, NY: Guilford Press.

Sanson, A., Letcher, P., Smart, D., Prior, M., Toumbourou, J. W., & Oberklaid, F. (2009). Associations between early childhood temperament clusters and later psychosocial adjustment. *Merrill-Palmer Quarterly*, *55*, 26–54. 10.1353/mpq.0.0015

Scott, B. G., Lemery-Chalfant, K., Clifford, S., Tein, J. Y., Stoll, R., & Goldsmith, H. H. (2016). A twin factor mixture modeling approach to childhood temperament: Differential heritability. *Child Development*, *87*, 1940–1955. 10.1111/cdev.12561

Shonkoff, J. P., & Phillips, D. A. (Eds.). (2000). *From neurons to neighborhoods: The science of early child development*. Washington, DC: National Academy Press.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York, NY: Brunner/Mazel.

Thomas, A., Chess, S., Birch, H. G., Hertzig, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood*. New York, NY: New York University Press.

van den Akker, A. L., Deković, M., Prinzie, P., & Asscher, J. J. (2010). Toddlers' temperament profiles: Stability and relations to negative and positive parenting. *Journal of Abnormal Child Psychology*, *38*, 485–495. 10.1007/s10802-009-9379-0

von Eye, A., & Bergman, L. R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology*, *15*, 553–580. 10.1017/S0954579403000294

Willoughby, M. T., Waschbusch, D. A., Moore, G. A., & Propper, C. B. (2011). Using the ASEBA to screen for callous unemotional traits in early childhood: Factor structure, temporal stability, and utility. *Journal of Psychopathology and Behavioral Assessment*, *33*, 19–30. 10.1007/s10862-010-9195-4

Zentner, M., & Bates, J. (2008). Child temperament: An integrative review of concepts, research programs, and measures. *European Journal of Developmental Science*, *2*, 7–37.