Developmental Psychology

Categorical and Latent Profile Approaches to Temperamental Infant Reactivity and Early Trajectories of Socioemotional Adjustment

Berenice Anaya, Laura Bierstedt, Nora Tucker, Kristin A. Buss, Vanessa LoBue, and Koraly Pérez-Edgar

Online First Publication, January 8, 2024. https://dx.doi.org/10.1037/dev0001679

CITATION

Anaya, B., Bierstedt, L., Tucker, N., Buss, K. A., LoBue, V., & Pérez-Edgar, K. (2024, January 8). Categorical and Latent Profile Approaches to Temperamental Infant Reactivity and Early Trajectories of Socioemotional Adjustment. *Developmental Psychology*. Advance online publication. https://dx.doi.org/10.1037/dev0001679

AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

https://doi.org/10.1037/dev0001679

Categorical and Latent Profile Approaches to Temperamental Infant Reactivity and Early Trajectories of Socioemotional Adjustment

Berenice Anaya[1], Laura Bierstedt[2], Nora Tucker[3], Kristin A. Buss[3], Vanessa LoBue[2], and Koraly Pérez-Edgar[3]
[1] Department of Psychiatry, School of Medicine, Washington University in St. Louis
[2] Department of Psychology, Rutgers University
[3] Department of Psychology, The Pennsylvania State University

This article examines the patterns, and consequences, of infant temperamental reactivity to novel sensory input in a large ($N = 357$; 271 in current analysis) and diverse longitudinal sample through two approaches. First, we examined profiles of reactivity in 4-month-old infants using the traditional theory-driven analytic approach laid out by Jerome Kagan and colleagues, and derived groups characterized by extreme patterns of negative reactivity and positive reactivity. We then used a theory-neutral, data-driven approach to create latent profiles of reactivity from the same infants. Despite differences in sample characteristics and recruitment strategy, we noted similar reactivity groups relative to prior cohorts. The current data-driven approach found four profiles: high positive, high negative, high motor, and low reactive. Follow-up analyses found differential predictions of internalizing, externalizing, dysregulation, and competence trajectories across 12, 18, and 24 months of life based on 4-month reactivity profiles. Findings are discussed in light of the initial formulation of early reactivity by Kagan and the four decades of research that has followed to refine, bolster, and expand on this approach to child-centered individual differences.

Public Significance Statement
This study provides insights into the patterns and consequences of infant temperamental reactivity to novel sensory input, using both theory-driven and data-driven approaches. The findings demonstrate the predictive power of infant reactivity profiles in understanding socioemotional trajectories across the first 2 years of life. These results could inform early identification and intervention efforts aimed at optimizing child well-being in diverse populations.

Keywords: temperamental reactivity, internalizing problems, infancy, latent profile analysis

Supplemental materials: https://doi.org/10.1037/dev0001679.supp

In the mid-1980s, Kagan and colleagues (García-Coll et al., 1984; Kagan et al., 1984) published a series of studies that proved to be a launching point for four decades of research into the early emergence of temperamental differences in social and emotional functioning. This work helped the emerging consensus pushing the field away from the duality of genetic versus environmental influences on development. The introduction of a new temperament type—behavioral inhibition (BI)—provided a developmental

Sara Rimm-Kaufman served as action editor.
Koraly Pérez-Edgar https://orcid.org/0000-0003-4051-9563
Support for this work was provided by a National Institute of Mental Health Grant (R01 MH109692 to Koraly Pérez-Edgar, Kristin A. Buss, and Vanessa LoBue) and a National Institute of Neurological Disorders and Stroke Grant (F99K00 NS120411 to Berenice Anaya). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was also supported by endowments through the Tracy Winfree and Ted H. McCourtney Professorship in Children, Work, and Families (Kristin A. Buss) and the McCourtney Professorship of Child Studies (Koraly Pérez-Edgar). Kristin A. Buss and Koraly Pérez-Edgar's Psychology Professorships are also supported by the Social Science Research Institute of The Pennsylvania State University. The authors thank the families who participated in this study and the research team members who collected and coded the data.

Berenice Anaya served as lead for formal analysis, visualization, and writing–original draft and contributed equally to funding acquisition. Laura Bierstedt served as lead for conceptualization and contributed equally to project administration and supervision. Nora Tucker contributed equally to data curation and formal analysis. Kristin A. Buss served as lead for conceptualization, funding acquisition, and supervision and served in a supporting role for writing–original draft. Vanessa LoBue served as lead for conceptualization, funding acquisition, and supervision and served in a supporting role for writing–original draft. Koraly Pérez-Edgar served as lead for conceptualization, funding acquisition, supervision, and writing–original draft. Berenice Anaya and Nora Tucker contributed equally to conceptualization. Laura Bierstedt and Nora Tucker contributed equally to writing–original draft. Berenice Anaya, Laura Bierstedt, Nora Tucker, Kristin A. Buss, Vanessa LoBue, and Koraly Pérez-Edgar contributed equally to writing–review and editing.

Correspondence concerning this article should be addressed to Koraly Pérez-Edgar, Department of Psychology, The Pennsylvania State University, 270 Moore Building, University Park, PA 16802, United States. Email: kxp24@psu.edu

construct that integrated a multifaceted array of biological, behavioral, and experiential processes (Pérez-Edgar & Fox, 2018). BI is characterized as an early appearing temperament trait marked by distinct responses to sensory and social novelty (Kagan et al., 1984). These responses often take the form of observed social withdrawal, unique cognitive and attentional patterns, a sensitivity to threat, and a distinct neural and psychophysiological profile marked by both hyperreactivity and overcontrol (Fox et al., 2005, 2022). While biologically based the trajectory and impact of BI is sensitive to environmental input (Anaya et al., 2023). Importantly, BI has proven to be our most powerful individual difference predictor for the emergence of social anxiety in late childhood and adolescence (Clauss & Blackford, 2012; Sandstrom et al., 2020). In addition to examining the prospective consequences of BI, Kagan's work also outlined an early infant antecedent of BI, negative reactivity, that was evident as early as 4 months of age and increased the probability of later presenting the BI phenotype (Kagan & Snidman, 1991; Kagan et al., 1994).

In the current article, we first focus on observed patterns of infant reactivity using the initial Kagan typology. We then compare the initial approach to data-driven groupings that emerge from a latent profile analysis (LPA). Our final analyses compared how Kagan's reactivity categories and our data-driven profiles predicted trajectories of socioemotional adaptation in the first 2 years of life. This work probes the utility of typological approaches to temperament and asks whether data-driven, but theory-informed, reactivity profiles suggest continued predictive validity of socioemotional outcomes even when applied in a more ethnically and economically diverse sample of infants than originally examined.

In contrast to other temperament approaches that focus on individual traits captured on a continuum (e.g., Rothbart, 1989), Kagan's approach was categorical. He argued that a unique constellation of traits within a child generated the qualitatively distinct phenotypic profile of BI (Kagan, 1997, 2018). Kagan also emphasized the need to hew closely to definitional boundaries across constructs, such that BI was defined by a clearly delineated set of observed behaviors in a specific context. Other measures could be considered correlates of, antecedents to, or consequences of BI—but they were not to be confused with the construct in and of itself (Kagan et al., 2002).

To better understand and predict the emergence of BI in toddlerhood, Kagan looked to see if early patterns of behavior could be reliably captured in the first months of life. In examining infant reactivity, Kagan drew on both empirical and theoretical work to characterize a specific profile that would reflect both the ontogeny and phylogeny of early development. Although more recent research has pointed to a very distributed neural model (Clauss & Blackford, 2012; Filippi et al., 2022), early work in this area focused on an amygdalar model of sensory sensitivity (Morgan, 2006; Schwartz et al., 2003). Given the early maturation of the amygdala and the limbic system, stable individual differences should therefore be evident early in life. Based on his early work in rodent models (Kagan, 1955), Kagan turned to the animal literature (Capitanio, 2018; Cavigelli, 2018; Cavigelli & McClintock, 2003).

There, he noted a distinct pattern of amygdala-driven responses to novelty that included initial freezing, arching of the back, negative vocalizations, and vigorous limb movements. One can note here that these are all behaviors well within the behavioral repertoire of the young infant. Initial empirical tests found that 2-month-old infants did not provide adequate data (Kagan & Snidman, 1991). It appeared that stable, individual variation in reactivity can only emerge when infants have settled into a semiregular behavioral and affective routine. The often chaotic first 100 days of life are both too variable within individuals (e.g., unsettled diurnal rhythms) and too homogenous across individuals (e.g., near-universal long stretches of sleep) to capture individual differences. Thus, reactivity testing typically begins at 4 months of age based on the initial empirical data.

While the developmental window for capturing infant reactivity is driven by both temperament theory and methodological considerations, Kagan solidified his approach by directly observing the infants who came to his laboratory. In his final article, published posthumously (Kagan, 2022), he wrote

> I took the films of several dozen infants to a room to get a feeling for the range of variation in behavior. The eighteenth infant I looked at supplied a first hypothesis. … I saw a female infant become increasingly aroused by the mobiles moving across her face, displaying vigorous limb movements and crying to these innocent stimuli. (p. 5)

In order to systematically elicit these responses, reactivity batteries present infants with an array of sensory stimuli that gradually increase in intensity. These include mobiles (visual), overlapping voices (auditory), and, at times, cotton swabs of diluted butyl alcohol (olfactory). Infants are then coded for positive and negative vocalizations and motoric behavior. Each category is scored continuously. Kagan and colleagues (Kagan, 1997; Kagan & Snidman, 1991) then created a simple, but powerful, formulation for classifying three distinct reactivity groups: high negative reactive infants showed elevated negative affect and motoric behavior, high positive reactive infants displayed elevated positive affect and motoric behavior, and low motor infants were typically, but not always, defined as scoring low across all three categories. Follow-up work in independent cohorts employed slightly different formulations within this same protocol (Fox et al., 2001; Hane et al., 2008).

Despite analytic variation, in each case, a distinct pattern emerges. Infants in the high negative reactivity group were more likely to display BI in toddlerhood (Fox et al., 2015; Kagan et al., 1998), and, in turn, greater levels of social anxiety in adolescence (Fox et al., 2022). Still, questions remain regarding how to best capture the full range of potential reactivity profiles. First, while the original formulation was an empirical and intuitive approach to categorization, it may reflect only the most concrete and phenotypically extreme profiles. Rarer, or more subtle, combinations of behavior could be overlooked (Woodward et al., 2000). Second, the original cohorts were overwhelmingly demographically homogeneous, White, upper middle class, and well-educated. Thus, it is not clear if the groupings represent universal reactivity typologies evident across diverse populations or are the product of the narrow slice of humanity living in and around Cambridge, Massachusetts. These concerns can be addressed by applying (a) data-driven approaches to (b) larger and more diverse samples. These are the aims of the current study.

With respect to data-driven approaches, one study (Loken, 2004) directly examined the use of latent class analysis with the infant data first published by Kagan (Kagan, 1997; Kagan & Snidman, 1991). Although data-driven and theoretically neutral, this approach found groupings quite similar in nature. In addition to the high and low negative reactivity groups of the original article, Loken (2004) found a third class characterized by high motor activity and high positive affect, a high positive reactivity group. This third group was also identified in the work of other laboratories (Fox et al., 2001; Hane et al., 2008), and is associated with later temperamental exuberance (Degnan et al., 2011; Putnam & Stifter, 2005). In a prior study (Woodward et al., 2000), Kagan and colleagues applied a maximum covariance analysis (MAXCOV; Meehl, 1995) to the same data as Loken (2004). They found that approximately 10% of infants fell into a latent high negative reactivity taxon. At 4.5 years of age, these children were more reticent and withdrawn in a social interaction.

Although not the same in focus, work by Beekman et al. (2015) presents one of the few other studies to take a data-driven, person-centered approach in the context of early temperament. This analysis leveraged a unique prospective adoption study (Leve et al., 2019). Adoptive parents completed the Infant Behavior Questionnaire (IBQ; Putnam et al., 2014) and the Toddler Behavior Assessment Questionnaire (TBAQ; Goldsmith, 1996) at 9, 18, and 27 months. These questionnaires provide continuous scores of individual traits that can be combined to create scores for high-order temperament traits. Here, the authors used dimensional indicators to create individual typologies through LPA. This approach can create unique patterns in the data as variable- and person-centered analyses can generate distinct constellations of relations even when using the same underlying data (Vallorani et al., 2021). At 9 months of age, they noted that 17% of infants fell into a high negative reactivity group, while 13% of infants were high positive. Over time, infants in these two groups (two other "typical" groups were also noted), consistently showed the highest levels of stability at 18 and then 27 months. Work within the BI literature suggests that extreme high, stable traits are most likely to predict the later emergence of maladjustment or social anxiety (Chronis-Tuscano et al., 2009; Clauss & Blackford, 2012).

On the second point of sample diversity, the field of developmental science has recently focused much-needed attention on the relative lack of diversity within the samples used to capture human development (Nielsen et al., 2017; Stein et al., 2023). This homogeneity has limited the breadth and width of developmental experience that can be captured (García-Coll, 2020). To take one example from the BI literature, Chen et al. (1995) found that despite similar initial base rates, the proportion of Chinese children categorized as BI rose over time. In addition, unlike in the United States, behaviorally inhibited children were seen as leaders among peers, academic stars, and presented with fewer symptoms of anxiety and depression (Chen et al., 2009). The implication is that while BI is a biologically based and relatively stable trait, the environmental responses experienced by the child will shift developmental trajectories within the probabilistic windows provided by temperament (Kagan, 1994). Indeed, years later, Liu et al. (2012) found that while the positive pattern previously noted in Chinese rural communities held true, children in urban and more westernized cities now displayed outcome profiles similar to those noted by researchers in the United States. Diversity of experience is central to our ability to expand theory, understand interwoven mechanisms, and as needed, implement interventions.

The current analyses examine the patterns, and consequences, of infant reactivity by leveraging the Longitudinal Attention and Temperament Study (LAnTs; Pérez-Edgar et al., 2021). LAnTs was designed as a large ($N = 357$) community sample of typically developing infants recruited in the first months of life. The sample was diverse: families were recruited from three distinct communities that varied in size, ethnic and racial distribution, socioeconomic status, and rurality. Starting at 4 months of age, children and families completed a comprehensive battery of tasks that included direct observation of temperament, eye-tracking, electrophysiology, psychophysiology, parent–child interaction, and questionnaire measures of socioemotional functioning. They returned to the laboratory at 8, 12, 18, and 24 months of age. For the current study, we focused on reactivity patterns at 4 months of age.

We first categorized infants using the foundational Kagan approach (Kagan, 1997; Kagan & Snidman, 1991). We then applied an LPA to examine data-driven, person-centered profiles. Our first aim was to compare the groupings that emerged from the two approaches. We then created trajectories of adaptive functioning at 12, 18, and 24 months of age based on maternal reports. Our second aim was to examine how early reactivity profiles predicted the emergence of internalizing, externalizing, dysregulation, and competence levels. Based on prior work, we expected to find a relatively small group of infants marked by high negative reactivity to sensory stimuli. In turn, these infants would show developmental trajectories marked by elevated symptoms of internalizing problems and dysregulated behavior. Thus, the current article builds on the prior work to examine the robustness and predictive power of early temperamental reactivity in a more diverse sample, over 40 years after its initial conceptualization.

Method

Participants

Participants were recruited through local baby registries (40% families) and university-sponsored participant databases (13% families). In addition, we used a variety of community-level recruitment strategies, such as visiting local lactation/parenting classes, communicating with families at local community events, and talking to parents at local hospitals, health care centers, and Women's and Infant Centers. Community recruiting identified 38% of our families. The remaining 10% of families were recruited by word-of-mouth. Prospective families were contacted by letter, email, or phone explaining the motivations and methods of the study. The Institutional Review Boards at the Pennsylvania State University and Rutgers University approved all procedures and parents provided written consent and were compensated for their participation. This study was not preregistered. Data are accessible through Databrary (LoBue et al., 2021) for those participants who consented to data sharing.

The larger cohort encompasses 357 infants (176 males, 181 females). Participants were recruited from areas surrounding three sites: State College, PA ($N = 167$), Harrisburg, PA ($N = 81$), and Newark, NJ ($N = 109$). Infants and their caregivers were enrolled when the infants were 4 months of age and completed the standard reactivity protocol ($N = 298$; 151 males, 147 females; $M_{\text{age}} = 4.80$ months; $SD_{\text{age}} = 0.80$). An additional 59 participants enrolled at older ages and were not included in the current manuscript.

Procedure

We collected data longitudinally at 4, 8, 12, 18, and 24 months. Infants and their parents came into the lab at all five assessments. At these visits, children completed eye-tracking tasks and a behavioral temperament battery. Parents also completed eye-tracking tasks and questionnaires assessing infant temperament, their own psychological state and traits, and the sociodemographic features of their environment.

Data collection was generally completed in two, 2-hr visits to the lab for the first four assessments, although some families completed all tasks in a single visit, and a subset of families required three visits. Most caregivers completed the online questionnaires at home prior to the visit, but in some cases, they were completed in the lab or over the phone. If questionnaires had to be completed in the lab, primary caregivers would do so while the infant was completing the eye-tracking tasks or after data collection was completed.

Measure

Demographics

Parents reported on the demographic characteristics of their family. Descriptive statistics for demographic variables can be found in Table 1.

Four-Month Reactivity

At the 4-month visit, infant temperamental reactivity was assessed behaviorally with a validated reactivity battery (Fox et al., 2001; Kagan & Snidman, 1991). The infant was seated in a car seat with a primary caregiver seated nearby out of the infant's line of sight. The experimenter then played the infant two audio tracks, one consisting of a series of sentences with overlapping voices and a second containing three groups of 10 repeated syllables. Interwoven with the audio tracks, the experimenter presented the infant with a series of mobiles. The mobiles consisted of plush neutral figures (bears or jungle animals, with order counterbalanced). The figures were presented in phases, going from one to three to five figures. Each phase lasted 20 s.

Behavioral coding of reactivity focused on the infants' affective and motoric responses to the auditory and visual stimuli (Fox et al., 2001, 2015). Coding captured the frequency of arm waves, arm wave bursts, leg kicks, leg kick bursts, back arches, hyperextensions, and smiles. The amount of time engaged in vocalizations, fusses, and crying was also coded. The presence (vs. absence) of hand clasping, finger/foot sucking, and feet rubbing were also noted. Videos were coded by four coders trained versus a master coder on a set of training videos until reaching reliability levels of at least 0.75. In the end, reliability was high with strong values across individual training tapes ($M = 0.929$, range = 0.789–0.977), and for individual coders ($M = 0.936$, range = 0.925–0.956).

Data were prorated for infants who could not complete all segments of the protocol. All analyses and classifications rely on prorated data. Reactivity data were available for 271 infants. The remaining 27 infants did not provide sufficient data for coding or prorating.

A positive affect score was calculated as the sum of vocalizations and smiles. A negative affect score was calculated as the sum of crying and fusses. Finally, a motor behavior score was calculated as the sum of arm waves, arm wave bursts, leg kicks, leg kick bursts, back arches, and hyperextensions. Descriptive statistics for the total motor, positive, and negative scores can be found in Table 2.

Child Behavior

Parents reported on their child's behavior using the Infant-Toddler Socioemotional Assessment (ITSEA; Carter et al., 2003). The ITSEA is a 200-item survey designed to assess multiple dimensions

Table 1
Demographic Characteristics for the Sample and as a Function of the Kagan Reactivity Groups

Sample characteristic	Sample level	High negative reactive	High positive reactive	Low reactive
N (%)	271	77 (28%)	79 (29%)	157 (42%)
Child biological sex (M/F)	135/136	36/41	38/41	84/73
Race/ethnicity				
African American/Black	58	9	4	45
Asian	9	3	2	4
Latinx	78	8	12	58
White	180	51	57	72
Mixed race	27	5	4	18
Household income				
$15,000 or less	49	9	7	33
$16,000–20,000	20	4	2	14
$21,000–30,000	22	5	4	13
$31,000–40,000	16	8	5	3
$41,000–50,000	22	7	7	8
$51,000–60,000	29	5	8	16
Above $60,000	113	39	46	28
Childcare				
Attended childcare	114	33	29	52
Did not attend childcare	157	42	49	66
Siblings (*Mdn*)	1	1	1	1
4-month age (months)	4.76 (0.81)	4.82 (0.93)	4.80 (0.74)	4.74 (0.76)
Mother education (years)	15.92 (3.28)	15.71(3.41)	16.12 (2.99)	16.11 (3.29)
Father education (years)	15.53 (3.31)	15.54(2.86)	15.64 (3.21)	15.71 (3.41)
Mother BAI	6.64 (7.73)	6.94 (9.30)	7.53 (8.80)	6.18 (6.58)
Mother BDI	5.69 (6.30)	6.29 (8.00)	5.75 (7.23)	5.56 (5.64)

Note. M = male; F = female; BAI = Beck anxiety inventory; BDI = Beck depression inventory.

Table 2
Descriptive Statistics for Main Study Variables

Reactivity measures	M (SD)	Range
Reactivity codes		
Positive affect	14.96 (20.55)	0–119
Negative affect	19.89 (27.91)	0–127
Motor activity	23.60 (19.90)	0–97

	12 months		18 months		24 months	
	M (SD)	Range	M (SD)	Range	M (SD)	Range
Socioemotional outcomes						
Internalizing	0.58 (0.30)	0–1.35	0.69 (0.31)	0–1.60	0.66 (0.33)	0–1.67
Externalizing	0.61 (0.34)	0–1.83	0.66 (0.35)	0–2.00	0.62 (0.37)	0–2.00
Dysregulation	0.44 (0.28)	0–1.19	0.41 (0.27)	0–1.58	0.42 (0.30)	0–1.89
Competence	0.86 (0.33)	0–1.74	1.21 (0.31)	0–1.88	1.32 (0.31)	0.17–1.95

Note. Reactivity scores presented are continuous values for each coding category separately. Reactivity groups are derived from combining rank positions across codes (e.g., median split), as described in the text.

of social-emotional problems and competencies in 1- to 3-year-old children 34. It was collected at the 12-, 18-, and 24-month assessments. Parents described their child on a set of behaviors or attributes for their child (e.g., Sleeps through the night; Is stubborn) in the past month on a 3-point scale (0 = *not true/rarely*, 1 = *somewhat true/sometimes*, 2 = *very true/often*). A "No opportunity" code allowed parents to indicate that they have not had the opportunity to observe certain behaviors (e.g., peer interactions). Each item taps into one of three problem domains (internalizing, externalizing, dysregulation) or a competence domain. The externalizing problems factor comprises three subscales (*activity/impulsivity*, *aggression/defiance*, and *peer aggression*). The internalizing problems factor comprises four subscales (*depression/withdrawal*, *general anxiety*, *separation distress*, *inhibition to novelty*). The dysregulation factor consists of four subscales (*sleep*, *negative emotionality*, *eating*, *sensory sensitivity*). The competence factor comprises six subscales (*compliance*, *attention*, *imitation/play*, *mastery motivation*, *empathy*, *prosocial peer relations*). The reliability and validity of the ITSEA have been examined in several prior studies (Briggs-Gowan & Carter, 1998, 2007). Descriptive statistics for the externalizing, internalizing, dysregulation, and competence scales at 12, 18, and 24 months are presented in Table 2.

Analytic Strategy

Traditional Analysis of Infant Reactivity

High and low groupings were created for motor behavior, positive affect, and negative affect scores by assigning children to groups based on median split. We then combined the affect and motor scores based on prior work (Fox et al., 2015). High positive reactivity infants scored above the median for both motor and positive affect. High negative reactivity infants scored above the median for both motor and negative affect. The number of infants assigned to each group can be found in Table 1.

LPA of Infant Reactivity

We used LPA to identify classes or subgroups of infants with similar behavioral patterns of motor, positive, and negative reactivity. LPA is a special case of mixture modeling where categorical latent classes are estimated to explain relations among the observed dependent variables or class indicators, and individuals are then classified to a specific subgroup based on person-specific similarities (Spurk et al., 2020). We carried out LPA in Mplus Version 8.8 (Muthén & Muthén, 2017) using behavioral codes of infant motor activity, positive affect, and negative affect as continuous class indicators. Descriptive statistics for the sample-level observed indicators are presented in Table 2.

The means and variances of the class indicators and the categorical latent classes were freely estimated. The variances of the class indicators were held equal across classes and the covariances were fixed at zero to estimate the most parsimonious measurement model. The optimal number of profiles was selected based on a set of indices, including Akaike information criteria (AIC), sample-size-adjusted Bayesian information criteria (BIC), and entropy. Based on previous recommendations, lower AIC and BIC, and an entropy value approaching 1.0 indicate a better fit (Lanza & Cooper, 2016). After selecting the optimal profile solution, infant age and sex, childcare outside the home, parent education, marital status, and Kagan's original median-split categories of positive and negative reactivity were considered as covariates. Covariates were examined using the automatic three-step approach (Asparouhov & Muthén, 2014; Vermunt, 2017), which carries out multinomial logistic regressions to test whether each covariate is associated with significant paths to the latent class while holding other covariates constant.

Trajectories of ITSEA Domain Scales

We used a series of linear growth models (LGMs) to estimate individual trajectories across the internalizing, externalizing, dysregulation, and competence scales of the ITSEA across 12, 18, and 24 months. We then examined whether reactivity class membership from our optimal LPA solution predicted the growth factors (i.e., intercept and slope) of the ITSEA scales. In separate analyses, we also tested whether Kagan's reactivity categories (i.e., negative and positive reactivities) predicted the growth factors. LGMs were carried out in Mplus, allowing for random slopes. The coefficients of the intercept were fixed at 1, and the coefficients for the slope were fixed at 0, 1, and 2 for the 12-, 18-, and 24-month assessments, respectively, to indicate linear and equidistant time intervals. Residual variances were freely estimated and allowed to vary across time. Missing data in the ITSEA scales were handled using multiple imputation via Bayesian analysis to estimate plausible values (Rubin, 1987; Schafer, 1997). Specifically, 20 imputations were used to compute the plausible value distribution for the random intercept and slope factors of each ITSEA domain scale, and these values were then exported as the individual-level data. We then used the individual-level Bayesian estimated intercept and slopes as the distal outcomes in subsequent confirmatory LPAs, separately for the internalizing, externalizing, dysregulation, and competence ITSEA scales. Distal outcomes were evaluated using the modified bounded cumulative hazard approach (Bakk et al., 2017) to test equality of means across the latent classes, which is the preferred method for continuous distal outcomes (Bakk et al., 2013). To test differences in trajectories between high and low negative and positive reactivity groups, the growth factors were regressed on the reactivity categories.

Supplemental Analysis of Fear Trajectories

The initial study aims included examining laboratory-based BI as an outcome measure. However, due to COVID-19 restrictions on in-person research imposed in March 2020 (Weiner et al., 2020), we were only able to complete observed BI protocols with 40 participants at age 24 months. In order to supplement our main analyses, we therefore also examined trajectories in maternal report of fear using the IBQ (Putnam et al., 2014) and TBAQ (Goldsmith, 1996). To do so, we replicated the analysis of the ITSEA scales, implementing a LGM to capture general trajectory. We then entered the individual mean and slope parameters as distal outcomes in our LPA model, examining starting levels (intercepts at 8 months) and change over time.

Results

Kagan's Reactivity Groups

Of children who completed the 4-month reactivity task, 192 (70.8%) children were classified as low positive reactivity, and 79 (29.2%) children were classified as high positive reactivity. Similarly, 194 (71.6%) children were classified as low negative reactivity and 77 (28.4%) children were classified as high negative reactivity. Demographic characteristics by group can be found in Table 1.

LPA of Infant Reactivity

We examined up to five-class solutions of infant reactivity. A four-class solution was retained as the best fit based on lower AIC and BIC, a high entropy value (0.933, indicating low classification error), and high posterior probabilities (Class 1 = 0.962, Class 2 = 0.866, Class 3 = 0.970, Class 4 = 0.995) which indicate low error of individual classifications into a given profile. While AIC and BIC were lower in the five-class solution, entropy decreased (0.915) and class proportions became more disparate ($n < 5$). Fit indices for all tested solutions are reported in Table S1 in the online supplemental materials. The four-class solution identified a high negative group ($n = 26$), a high positive group ($n = 19$), and a high motor group of infants ($n = 15$), representing roughly 10%, 7%, and 5% of the general sample, respectively. This solution also identified a much larger low reactive group of 211 infants, representing 78% of the sample. Means for each reactivity class across the observed indicators and Kagan's variable-centered reactivity groups are presented in Table 3.

We then compared the four classes to the Kagan-based classifications. The high negative class was characterized by more infants who were classified as high reactive according to Kagan's original approach (log odds = 1.661, $p = .001$) and more male infants (log odds = −1.386, $p = .008$) compared to the low reactive class. Additionally, the high motor and high positive classes were both characterized by more infants who were classified as high positive according to Kagan's original approach (log odds = 27.148 and 2.217, respectively, $p = .001$) compared to the low reactive class. Distinctively, the high positive class was characterized by older infants (log odds = 0.932, $p = .002$) while the high motor class was characterized by younger infants (log odds = −1.291, $p = .008$).

Growth Models of ITSEA Domain Scales

Model fit statistics for all retained models and growth parameters are presented in Table 4. The LGMs adequately fit the data for the internalizing, externalizing, and dysregulation scales. In contrast, the competence scale exhibited poor fit (posterior predictive p-value = 0; comparative fit index [CFI] = 0.769; Tucker–Lewis index [TLI] = 0.732; root-mean-square error of approximation [RMSEA] = 0.061). We then tested a fixed quadratic trend, which substantially improved model fit (posterior predictive p-value =

Table 3
Reactivity Behavioral Codes Across Kagan's Categories and LPA Classes

Temperament category or class	Reactivity coding: Positive affect	Reactivity coding: Negative affect	Reactivity coding: Motor activity
Kagan's categories			
High positive reactivity (79)	31.15	26.65	47.41
High negative reactivity (77)	20.07	53.06	45.06
Low reactivity (157)	11.96	21.06	13.82
LPA			
Low reactive (211)	10.71	13.88	21.46
High negative (26)	7.47	121.40	27.83
High positive (19)	84.17	38.50	35.47
High motor (15)	23.60	24.40	79.32

Note. LPA = latent profile analysis.

Table 4
Growth Models for ITSEA Domain Scales

	ITSEA domain scales				
Fit indices	Internalizing	Externalizing	Dysregulation	Competence LGM	Competence QGM
Model estimates					
I_{mean}	0.469*	0.410*	0.430*	0.928*	0.841*
S_{mean}	0.041*	0.047*	−0.008	0.218*	0.500*
I_{variance}	0.030*	0.037*	0.040*	0.044*	0.074*
S_{variance}	0.006*	0.009*	0.007*	0.007*	0.142*
$I \sim S$	−0.003	−0.002	−0.002	0.003	0.023*
Q_{mean}					−0.124*
Q_{variance}					0.026*
$I \sim Q$					−0.059*
Fit measures					
PP p-value	.073	.090	.156	.000	.527
CFI/TLI	0.880/0.876	0.940/0.940	0.962/0.965	0.769/0.732	1.00/1.00
RMSEA	0.057	0.053	0.025	0.061	0.000

Note. ITSEA = Infant-Toddler Socioemotional Assessment; LGM = linear growth model; QGM = quadratic growth model; PP = posterior probability; I = intercept; S = slope; Q = quadratic term; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root-mean-square error of approximation.
* $p < .05$.

5.27; CFI = 1.000; TLI = 1.000; RMSEA = 0.001), and this model was retained for subsequent analysis. Linear growth parameters indicated that internalizing and externalizing levels significantly increased from 12 to 24 months ($S_{\text{Int}} = 0.041$, $p = .001$; $S_{\text{Ext}} = 0.047$, $p = .001$), while dysregulation remained stable during this period ($S_{\text{Dys}} = -0.008$, $p = .213$). The quadratic growth parameter for the competence scale indicated that competence levels increased from 12 to 18 months ($S_{\text{Comp}} = 0.500$, $p = .001$) and then decreased from 18 to 24 months ($S_{\text{Comp}} = -0.124$, $p = .001$). Importantly, intercept and slope variance parameters were significant across all scales, indicating substantial variability in individual trajectories. Raw socioemotional trajectories are presented in Figure 1 and model-estimated average trajectories are reported in Figures S4.1–S4.4 in the online supplemental materials.

Predicting ITSEA Trajectories From Class Membership and Categorical Groups

Confirmatory LPAs using the four-class solution and growth parameters of the ITSEA scales as distal outcomes indicated significant differences in internalizing, externalizing, and dysregulation levels as a function of class membership (Tables 5–7).

Within the Internalizing domain, infants in the low reactive ($M_I = 0.472$, $SE = 0.008$, $\chi^2\Delta = 4.125$, $p = .042$) and high positive ($M_I = 0.499$, $SE = 0.030$, $\chi^2\Delta = 4.500$, $p = .034$) classes exhibited a significantly higher intercept (i.e., higher internalizing levels at 12 months) compared to the high motor class ($M_I = 0.415$, $SE = 0.026$). There were no significant differences in Internalizing trajectories across reactivity classes.

Within the externalizing domain, infants in the high negative class ($M_S = 0.020$, $SE = 0.009$, $\chi^2\Delta = 4.414$, $p = .036$) had significantly flatter slopes or less pronounced increases in externalizing levels compared to the low reactive class ($M_S = 0.042$, $SE = 0.004$).

Within the dysregulation domain, the high negative class exhibited significantly lower dysregulation levels at 12 months ($M_I = 0.395$, $SE = 0.017$, $\chi^2\Delta = 4.977$, $p = .026$) and slightly decreased over time at trend ($M_S = -0.015$, $SE = 0.004$, $\chi^2\Delta = 3.042$, $p = .081$) compared to the low reactive class ($M_I = 0.439$, $SE = 0.010$; $M_S = -0.006$, $SE = 0.002$). Similarly, infants in the high positive class ($M_S = -0.017$, $SE = 0.009$, $\chi^2\Delta = 2.930$, $p = .087$) slightly decreased in dysregulation levels compared to the low reactive class.

There were no significant differences between reactivity classes within the competence domain.

We then tested how Kagan's negative and positive reactivity groups predicted the growth factors of the ITSEA scales. Neither high nor low groups of negative and positive reactivity were significantly associated with any of the growth factors across internalizing, externalizing, dysregulation, or competence scales (all $ps > .078$).

Supplemental Fear Trajectories

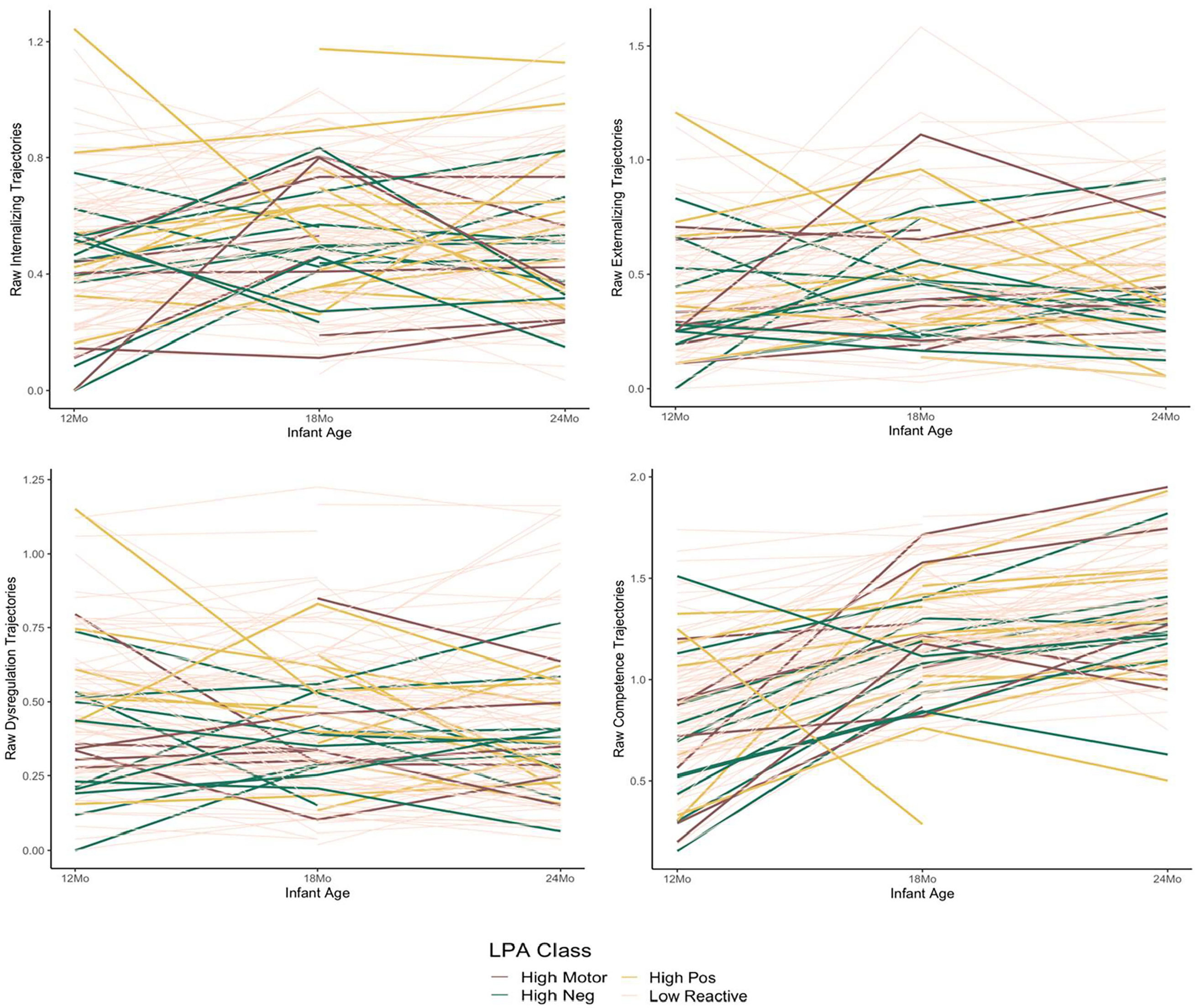
The LGM indicated that fear generally increased across the sample from 8 to 24 months ($I = 2.952$, $p = .001$; $S = 0.248$, $p = .001$). When we entered the individual mean and slope parameters as distal outcomes in our LPA model, all four classes started out with similar fear levels at 8 months (i.e., intercepts were not significantly different from each other). However, mean slopes between the high motor and high negative classes were significantly different ($\chi^2 = 4.535$, $p = .033$), such that the high motor class showed exacerbated increases in fear ($M = 0.067$, $SE = 0.012$) relative to the high negative class ($M = 0.034$, $SE = 0.010$).

Discussion

Kagan viewed infant reactivity as a typology rather than a continuum (Kagan et al., 1994), and he argued that categorical distinctions between high- and low-reactive infants were supported by distinct developmental profiles of behavior, brain mechanisms, and fearful tendencies (Kagan et al., 1998). Indeed, one of his greatest scientific contributions was to document how these early occurring reactivity types predicted inhibited and uninhibited temperament during early childhood and then their long-term socioemotional outcomes, including shyness and anxiety (Kagan, 2018). In the present study, we used a person-centered analytic approach to model latent classes

Figure 1

Raw Trajectories of the ITSEA Scales as a Function of LPA Class Membership



Note. ITSEA = Infant-Toddler Socioemotional Assessment; LPA = latent profile analysis; Mo = months; Pos = positive; Neg = negative. See the online article for the color version of this figure.

of infant reactivity from Kagan's typological framework in a large and more diverse sample of 4-month-old infants. We then compared these latent classes to the median-split categories based on Kagan's original work (Kagan, 1994), contextualizing each type with demographic variables of interest. Finally, we examined the validity of reactivity latent classes to predict developmental differences in internalizing, externalizing, dysregulation, and competence levels from 12 to 24 months compared to Kagan's original categories.

We found four latent classes of reactivity that were comparable to Kagan's original reactivity groups, supporting his conceptualization of infant reactivity as a typology based on meaningful individual differences in motor and affective expressions. However, class proportions were different from the group sizes based on Kagan's coding, and from the proportions identified in his previous work. Contrary to our expectations, the high negative class was not associated with higher or increasing internalizing trajectories. Instead, we found more nuanced associations between reactivity classes and externalizing and dysregulation trajectories that may reflect deviations from prototypical development and fundamental differences between the reactivity classes in the context of a demographically diverse sample. No predictive relations were found with the original Kagan categories.

Reactivity Types as a Robust Temperament Measure

Our LPA analysis indicated the presence of four reactivity classes. The high negative class corresponded closely with Kagan's originally negative reactivity group, scoring the highest on negative

Table 5

Differences in ITSEA Internalizing Trajectories as a Function of Class Membership

Class comparison	*M*	*SE*	χ^2	*p*
Intercept				
High motor versus high negative	0.415 versus 0.460	0.026/0.014	2.307	.129
High motor versus low reactive	**0.415 versus 0.472**	**0.026/0.008**	**4.125**	**.042**
High motor versus high positive	**0.415 versus 0.499**	**0.026/0.030**	**4.500**	**.034**
High negative versus low reactive	0.460 versus 0.472	0.014/0.008	0.551	.458
High negative versus high positive	0.460 versus 0.499	0.014/0.030	1.461	.227
Low reactive versus high positive	0.472 versus 0.499	0.008/0.030	0.806	.369
Slope (no significant differences)				
High negative	0.033	0.006		
High motor	0.042	0.007		
Low reactive	0.037	0.002		
High positive	0.048	0.009		

Note. Significant and trending comparisons are noted in bold. ITSEA = Infant-Toddler Socioemotional Assessment.

affect, lowest in positive affect, and average in motor activity. Importantly, when we included Kagan's categorical grouping into the LPA, only the high negative class was characterized by significantly more infants who would have been categorized as "negative reactive" by Kagan. The low reactive class corresponded closely with Kagan's low reactivity type, scoring the lowest in negative and positive affect, as well as motor activity.

Our reactivity classes also align with results from Loken (2004), who used a person-centered approach to examine reactivity latent classes in Kagan's original data. Beyond two classes that corresponded with Kagan's high and low reactivity groups, Loken (2004) found a third class characterized by similar motor activity as the high reactive group, but lower in negative affect and higher in positive vocalization. Our high positive class seems to directly replicate this third class. While Loken (2004) selected a three-class over a four-class solution (the latter presenting multimodality issues and worse fit indices), that article briefly mentions that the fourth class in a four-class solution aligned with a profile that Kagan (1994) labeled as "aroused." This class was characterized by low distress and high motor activity, similar to the behavioral pattern we see in our high motor class.

Some differences emerged in our LPA results that should be highlighted. For instance, the class proportions for our high negative (10% of our sample) and low reactive classes (78%) were lower and higher, respectively, compared to the proportions previously reported in Kagan's sample (high reactive group ~ 20% and low reactive group ~ 40% of the sample; Kagan, 1994, 1997). However, it should be noted that our proportion of high negative infants is remarkably similar to the 10% found by Woodward and colleagues (Woodward et al., 2000) using a MAXCOV approach. In addition, the class proportions for our high negative and high positive classes were comparable to the proportions reported by Fox and colleagues (Fox et al., 2015) when a more racially and ethnically diverse sample was examined.

Predicting Socioemotional Outcomes

When we examined the validity of our reactivity latent classes to predict socioemotional development, we found that infants in the high negative class did not significantly differ in their internalizing trajectories compared to other classes. Instead, children in this group exhibited flatter increases in externalizing levels between 12 and 24 months compared to the low reactive class. We note a similar pattern in our supplemental fear trajectory analysis. This was a puzzling finding, because in both Kagan's original work (Kagan et al., 1994) and follow-up studies (Fox et al., 2015; Hane et al., 2008; Kagan et al., 1998), this high motor and high

Table 6

Differences in ITSEA Externalizing Trajectories as a Function of Class Membership

Class comparison	*M*	*SE*	χ^2	*p*
Intercept (no significant differences)				
High negative	0.401	0.015		
High motor	0.413	0.029		
Low reactive	0.421	0.008		
High positive	0.437	0.028		
Slope				
High motor versus high negative	0.045 versus 0.020	0.012/0.009	2.623	.105
High motor versus low reactive	0.045 versus 0.042	0.012/0.004	0.058	.810
High motor versus high positive	0.045 versus 0.037	0.012/0.015	0.177	.674
High negative versus low reactive	**0.020 versus 0.042**	**0.009/0.004**	**4.414**	**.036**
High negative versus high positive	0.020 versus 0.037	0.009/0.015	0.819	.365
Low reactive versus high positive	0.042 versus 0.037	0.004/0.015	0.106	.745

Note. Significant and trending comparisons are noted in bold. ITSEA = Infant-Toddler Socioemotional Assessment.

Table 7
Differences in ITSEA Dysregulation Trajectories as a Function of Class Membership

Class comparison	*M*	*SE*	χ^2	*p*
Intercept				
High motor versus high negative	0.424 versus 0.395	0.031/0.017	0.314	.575
High motor versus low reactive	0.424 versus 0.439	0.031/0.010	0.196	.658
High motor versus high positive	0.424 versus 0.450	0.031/0.030	0.363	.547
High negative versus low reactive	**0.395 versus 0.439**	**0.017/0.010**	**4.977**	**.026**
High negative versus high positive	0.395 versus 0.450	0.017/0.030	0.052	.820
Low reactive versus high positive	0.439 versus 0.450	0.010/0.017	0.127	.721
Slope				
High motor versus high negative	−0.011 versus −0.015	0.005/0.005	0.314	.575
High motor versus low reactive	−0.011 versus −0.006	0.005/0.002	0.629	.428
High motor versus high positive	−0.011 versus −0.017	0.005/0.006	0.501	.479
High negative versus low reactive	**−0.015 versus −0.006**	**0.005/0.002**	**3.042**	**.081**
High negative versus high positive	−0.015 versus −0.017	0.005/0.006	0.052	.820
Low reactive versus high positive	**−0.006 versus −0.017**	**0.002/0.006**	**2.930**	**.087**

Note. Significant and trending comparisons are noted in bold. ITSEA = Infant-Toddler Socioemotional Assessment.

distress reactivity group has been consistently associated with inhibited and avoidant tendencies across late infancy (14 and 21 months) and early childhood.

There are three potential reasons for the lack of a significant relation in our study. First, we recruited a community sample of families who may not have presented with the range or concentration of symptoms necessary. This may reflect a low sensitivity of the ITSEA to probe mothers' reports of children's internalizing levels in a nonclinical sample. Second, the relatively low level of risk based on sample and age may have been compounded by the fact that our high negative class only represented 10% of the sample. Thus, we may have been underpowered to detect such differences. Third, infants in the study are fairly young relative to the work showing a link between early temperament and internalizing problems. Although symptoms can be detected in the first and second years of life (Whalen et al., 2017), many children who go on to have socioemotional difficulties may not display steady and stable symptoms until later. Indeed, infant reactivity typically predicts BI, which in turn later predicts the robust emergence of symptoms.

The question of power is a common concern with typological or person-centered approaches to research since a sufficient number of group members are needed in order to then examine potential relations with other variables of interest (Tein et al., 2013). Indeed, most longitudinal studies focused on infant reactivity and BI (Fox et al., 2001, 2015) oversample for high reactive profiles at the very outset of the study. An enriched sample will increase the probability of having a greater proportion of behaviorally inhibited children in the sample, and then, over time, provide a foundation for the emergence of symptoms in late childhood and adolescence.

In our data, the high negative class was instead associated with less increases in externalizing levels between 12 and 24 months. In examining early emerging symptoms, the behaviors associated with externalizing problems are often more salient (e.g., "Acts aggressive when frustrated") than behaviors used to assess internalizing difficulties (e.g., "Looks unhappy or sad without any reason"). It is possible that in the context of a community sample, mothers' reports of externalizing levels reflect some of the challenging behaviors that increase during this period as infants gain vocabulary and demand more autonomy (Aktar & Pérez-Edgar, 2020). Indeed, Kagan (2018) noted that "parents typically award greater weight to behaviors that are both more intense and less frequent…" (p. 3). Thus, this deviation of the high negative class from externalizing increases that are normative during this developmental period (Kjeldsen et al., 2021) may reflect a low probability that these infants' trajectories will change toward uninhibited or exuberant behavioral profiles later in life. Kagan and others (Fox et al., 2001; Kagan et al., 1998) suggest that while change in trajectory remains possible throughout development given diverse environmental exposure and selection (Pérez-Edgar, 2018). Indeed, recent work in an adoption cohort suggests that BI trajectories are sensitive to both underlying genetic predisposition and the specific characteristics of the rearing environment (Anaya et al., 2023). Thus, while temperament is not impervious to the experience, drastic change from one extreme to the other is unlikely. For example, while some high negative reactive infants may decrease in fear and inhibition over time and resemble their low reactive peers by age five, they are unlikely to become exuberant adolescents or extroverted adults.

Finally, we note that there were no significant differences across internalizing, externalizing, dysregulation, or competence trajectories as a function of Kagan's reactivity groups. This is an interesting finding, considering that creating groups based on a median split of behavioral codes naturally resulted in larger groups of high positive ($n = 79$) and negative reactivity ($n = 77$) infants compared to the analogous LPA classes. Thus, these categorical groups should have provided enough power to detect developmental differences that emerged as a function of LPA classes. However, a closer look at the average behavioral codes for each reactivity group compared to the LPA classes (Table 3) suggests stark differences in how the groups and the classes were defined by behavioral patterns.

For example, the average positive affect code was substantially higher for the high positive class ($M = 84.17$) than for the high positive reactivity group ($M = 31.15$). A similar pattern emerged for the high negative class ($M = 121.40$) compared to the high negative reactivity group ($M = 53.06$). Thus, it is possible that while high and low reactivity groups were different from each other, the median-split categorization allowed enough overlap in behavior between the groups, decreasing sensitivity to predict socioemotional trajectories in our sample.

In contrast, our LPA classes were characterized by unique patterns of behavior, only assigning infants to each class based on similar highest or lowest behavioral expressions. Furthermore, the high entropy in our model (0.933) also indicates high within-class homogeneity and between-class differences. This, of course, is an evident utility of person-centered approaches, which maximize relationships among individuals rather than variables, based on individual differences that emerge through logical behavioral patterns to explain meaningful differences in the population (Ferguson et al., 2020).

Limitations and Future Directions

Our current findings set the stage for future research but should be assessed with limitations in mind. For example, in the initial launch of the LAnTs study, we incorporated an in-laboratory BI assessment at age two (Pérez-Edgar et al., 2021). This would have allowed us to directly examine the previously noted relation between 4-month reactivity and BI in toddlerhood (Fox et al., 2015). However, with the implementation of COVID-19 mitigation efforts (Weiner et al., 2020), we terminated in-laboratory data collection early. As a result, only 40 children were able to complete the BI protocol in person. The analyses presented here leverage the questionnaire data that we continued to collect remotely throughout the full term of the study.

Our study included 273 infants who provided reactivity data. While this sample is larger than usual for this type of infant study, LPA solutions are generally more likely to converge and be properly replicated when using larger sample sizes ($N > 500$) and a greater number of class indicators (Swanson et al., 2012; Wurpts & Geiser, 2014). Thus, the LPA solution we report here should be replicated with a larger sample. Alternatively, researchers may also wish to harmonize available data sets of infant reactivity from different cohorts and use LPA to examine reactivity profiles in a larger and more nationally representative sample of infants.

Relatedly, we retained the most parsimonious LPA solution, assuming homogeneity of variance between classes and local independence (i.e., any relation between indicators is explained by the latent classes). However, it is possible that classes with more extreme temperament reactivity may also exhibit higher class-specific variance. In the current analyses, specifying the variance for the extreme reactivity classes worsened the model fit, and the constrained model was retained. Nonetheless, this could also be the result of increasing model parameters with a limited sample size. Thus, we also urge future studies using larger sample sizes to further explore equality constraints between reactivity classes.

Future work should also normalize the use of data-driven approaches to supplement traditional groupings. Beekman et al. (2015) employed an LPA approach with maternal report of temperament to characterize initial profiles and change in profile membership at 9, 18, and 27 months of age. They found greater stability in temperament profiles than in membership over time. Infants in the current study provided laboratory observations of temperament via the Laboratory Temperament Assessment Battery (Planalp et al., 2017), at each visit from 8 to 24 months, until interrupted by COVID-19. Depending on available sample sizes, an interesting next step would be to examine whether similar temperamental profiles emerge in our data, which covers overlapping infant ages, and whether similar patterns of continuity and change emerge at overlapping times.

Finally, socioemotional outcomes were examined via maternal reports on the ITSEA, a clinical screening instrument to identify significant symptom-level problems and delays that may warrant pediatric follow-up and intervention (Carter et al., 2003). Thus, it is possible that our community sample of infants had a low rate of clinical symptoms, truncating variability of socioemotional outcomes and hindering our ability to link early profiles of reactivity to more robust differences in socioemotional trajectories. Following enriched samples later into childhood may help spotlight a wider range of outcomes.

Conclusion

Infant reactivity is one of the earliest observed temperamental indicators of anxiety risk (Fox et al., 2015). The current study looked to see if Kagan's initial approach could be applied to contemporary cohorts of infants. These reactivity groups were then compared to latent profiles derived by data-driven approaches. Altogether, our findings buttress Kagan's original empirical work (Fox et al., 2015; Kagan, 1994; Kagan et al., 1998) and more recent person-centered approaches (Loken, 2004). We advance this work by illustrating that similar core typologies emerged in an independent, large, and more diverse sample of infants. We also show that clusters of behavior, rather than continuous measures of a single trait, provide distinct predictions for early socioemotional trajectories. Much like the work of others (Fox et al., 2015; Miller et al., 2019), our data suggest that Kagan's reactivity categorizations are not arbitrary, but rather explain qualitatively different patterns across motor activity and positive and negative affects that are evident four decades after their initial introduction to the field.

References

Aktar, E., & Pérez-Edgar, K. (2020). Infant emotion development and temperament. In J. Lockman & C. Tamis-LeMonda (Eds.), *The Cambridge handbook of infant development: Brain, behavior, and cultural context* (pp. 715–741). Cambridge University Press.

Anaya, B., Neiderhiser, J. M., Pérez-Edgar, K., Leve, L. D., Ganiban, J. M., Reiss, D., Natsuaki, M. N., & Shaw, D. S. (2023). Developmental trajectories of behavioral inhibition from infancy to age seven: The role of genetic and environmental risk for psychopathology. *Child Development*, *94*(4), e231–e245. https://doi.org/10.1111/cdev.13924

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329–341. https://doi.org/10.1080/10705511.2014.915181

Bakk, Z., Oberski, D. L., & Vermunt, J. K. (2017). Relating latent class assignments to external variables: Standard errors for correct inference. *Political Analysis*, *22*(4), 520–540. https://doi.org/10.1093/pan/mpu003

Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociological Methodology*, *43*(1), 272–311. https://doi.org/10.1177/0081175012470644

Beekman, C., Neiderhiser, J. M., Buss, K. A., Loken, E., Moore, G. A., Leve, L. D., Ganiban, J. M., Shaw, D. S., & Reiss, D. (2015). The development of early profiles of temperament: Characterization, continuity, and etiology. *Child Development*, *86*(6), 1794–1811. https://doi.org/10.1111/cdev.12417

Briggs-Gowan, M. J., & Carter, A. S. (1998). Preliminary acceptability and psychometrics of the Infant–Toddler Social and Emotional Assessment (ITSEA): A new adult-report questionnaire. *Infant Mental Health Journal*, *19*(4), 422–445. https://doi.org/10.1002/(SICI)1097-0355(199824)19:4<422::AID-IMHJ5>3.0.CO;2-U

Briggs-Gowan, M. J., & Carter, A. S. (2007). Applying the Infant-Toddler Social & Emotional Assessment (ITSEA) and Brief-ITSEA in early intervention. *Infant Mental Health Journal*, *28*(6), 564–583. https://doi.org/10.1002/imhj.20154

Capitanio, J. P. (2018). Behavioral inhibition in nonhuman primates: The elephant in the room. In K. Pérez-Edgar & N. A. Fox (Eds.), *Behavioral inhibition: Integrating theory, research, and clinical perspectives* (pp. 17–33). Springer. https://doi.org/10.1007/978-3-319-98077-5_2

Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. D. (2003). The Infant–Toddler Social and Emotional Assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology*, *31*(5), 495–514. https://doi.org/10.1023/A:1025449031360

Cavigelli, S. A. (2018). Behavioral inhibition in rodents: A model to study causes and health consequences of temperament. In K. Pérez-Edgar & N. A. Fox (Eds.), *Behavioral inhibition: Integrating theory, research, and clinical perspectives* (pp. 35–58). Springer. https://doi.org/10.1007/978-3-319-98077-5_3

Cavigelli, S. A., & McClintock, M. K. (2003). Fear of novelty in infant rats predicts adult corticosterone dynamics and an early death. *Proceedings of the National Academy of Sciences*, *100*(26), 16131–16136. https://doi.org/10.1073/pnas.2535721100

Chen, X., Chen, H., Li, D., & Wang, L. (2009). Early childhood behavioral inhibition and social and school adjustment in Chinese children: A 5-year longitudinal study. *Child Development*, *80*(6), 1692–1704. https://doi.org/10.1111/j.1467-8624.2009.01362.x

Chen, X., Rubin, K. H., & Li, Z. (1995). Social functioning and adjustment in Chinese children: A longitudinal study. *Developmental Psychology*, *31*(4), 531–539. https://doi.org/10.1037/0012-1649.31.4.531

Chronis-Tuscano, A., Degnan, K. A., Pine, D. S., Pérez-Edgar, K., Henderson, H. A., Diaz, Y., Raggi, V. L., & Fox, N. A. (2009). Stable early maternal report of behavioral inhibition predicts lifetime social anxiety disorder in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(9), 928–935. https://doi.org/10.1097/CHI.0b013e3181ae09df

Clauss, J. A., & Blackford, J. U. (2012). Behavioral inhibition and risk for developing social anxiety disorder: A meta-analytic study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(10), 1066–1075.e1. https://doi.org/10.1016/j.jaac.2012.08.002

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*(3), 765–780. https://doi.org/10.1037/a0021316

Ferguson, S. L., Moore, E. W. G., & Hull, D. M. (2020). Finding latent groups in observed data: A primer on latent profile analysis in Mplus for applied researchers. *International Journal of Behavioral Development*, *44*(5), 458–468. https://doi.org/10.1177/0165025419881721

Filippi, C. A., Valadez, E. A., Fox, N. A., & Pine, D. S. (2022). Temperamental risk for anxiety: Emerging work on the infant brain and later neurocognitive development. *Current Opinion in Behavioral Sciences*, *44*, Article 101105. https://doi.org/10.1016/j.cobeha.2022.101105

Fox, N. A., Henderson, H. A., Marshall, P. J., Nichols, K. E., & Ghera, M. M. (2005). Behavioral inhibition: Linking biology and behavior within a developmental framework. *Annual Review of Psychology*, *56*(1), 235–262. https://doi.org/10.1146/annurev.psych.55.090902.141532

Fox, N. A., Henderson, H. A., Rubin, K. H., Calkins, S. D., & Schmidt, L. A. (2001). Continuity and discontinuity of behavioral inhibition and exuberance: Psychophysiological and behavioral influences across the first four years of life. *Child Development*, *72*(1), 1–21. https://doi.org/10.1111/1467-8624.00262

Fox, N. A., Snidman, N., Haas, S. A., Degnan, K. A., & Kagan, J. (2015). The relations between reactivity at 4 months and behavioral inhibition in the second year: Replication across three independent samples. *Infancy*, *20*(1), 98–114. https://doi.org/10.1111/infa.12063

Fox, N. A., Zeytinoglu, S., Valadez, E. A., Buzzell, G. A., Morales, S., & Henderson, H. A. (2022). Annual Research Review: Developmental pathways linking early behavioral inhibition to later anxiety. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *64*(4), 537–561. https://doi.org/10.1111/jcpp.13702

García-Coll, C. G. (2020). Globalization, culture, and development: Are we ready for a paradigm shift? *Human Development*, *64*(4–6), 245–249. https://doi.org/10.1159/000512903

García-Coll, C. G., Kagan, J., & Reznick, J. S. (1984). Behavioral inhibition in young children. *Child Development*, *55*(3), 1005–1019. https://doi.org/10.2307/1130152

Goldsmith, H. H. (1996). Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development*, *67*(1), 218–235. https://doi.org/10.2307/1131697

Hane, A. A., Fox, N. A., Henderson, H. A., & Marshall, P. J. (2008). Behavioral reactivity and approach-withdrawal bias in infancy. *Developmental Psychology*, *44*(5), 1491–1496. https://doi.org/10.1037/a0012855

Kagan, J. (1955). Differential reward value of incomplete and complete sexual behavior. *Journal of Comparative and Physiological Psychology*, *48*(1), 59–64. https://doi.org/10.1037/h0043461

Kagan, J. (1994). *Galen's prophecy: Temperament in human nature*. Routledge. https://doi.org/10.4324/9780429500282

Kagan, J. (1997). Temperament and the reactions to unfamiliarity. *Child Development*, *68*(1), 139–143. https://doi.org/10.2307/1131931

Kagan, J. (2018). The history and theory of behavioral inhibition. In K. Pérez-Edgar & N. A. Fox (Eds.), *Behavioral inhibition: Integrating theory, research, and clinical perspectives* (pp. 1–15). Springer. https://doi.org/10.1007/978-3-319-98077-5_1

Kagan, J. (2022). Temperamental and theoretical contributions to clinical psychology. *Annual Review of Clinical Psychology*, *18*(1), 1–18. https://doi.org/10.1146/annurev-clinpsy-071720-014404

Kagan, J., Arcus, D., Snidman, N., Feng, W. Y., Hendler, J., & Greene, S. (1994). Reactivity in infants: A cross-national comparison. *Developmental Psychology*, *30*(3), 342–345. https://doi.org/10.1037/0012-1649.30.3.342

Kagan, J., Reznick, J. S., Clarke, C., Snidman, N., & Garcia-Coll, C. (1984). Behavioral inhibition to the unfamiliar. *Child Development*, *55*(6), 2212–2225. https://doi.org/10.2307/1129793

Kagan, J., & Snidman, N. (1991). Temperamental factors in human development. *American Psychologist*, *46*(8), 856–862. https://doi.org/10.1037/0003-066X.46.8.856

Kagan, J., Snidman, N., & Arcus, D. (1998). Childhood derivatives of high and low reactivity in infancy. *Child Development*, *69*(6), 1483–1493. https://doi.org/10.1111/j.1467-8624.1998.tb06171.x

Kagan, J., Snidman, N., McManis, M., Woodward, S., & Hardway, C. (2002). One measure, one meaning: Multiple measures, clearer meaning. *Development and Psychopathology*, *14*(3), 463–475. https://doi.org/10.1017/S0954579402003048

Kjeldsen, A., Nes, R. B., Sanson, A., Ystrom, E., & Karevold, E. B. (2021). Understanding trajectories of externalizing problems: Stability and emergence of risk factors from infancy to middle adolescence. *Development and Psychopathology*, *33*(1), 264–283. https://doi.org/10.1017/S0954579419001755

Lanza, S. T., & Cooper, B. R. (2016). Latent class analysis for developmental research. *Child Development Perspectives*, *10*(1), 59–64. https://doi.org/10.1111/cdep.12163

Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics*, *22*(6), 716–727. https://doi.org/10.1017/thg.2019.66

Liu, J., Chen, X., Li, D., & French, D. (2012). Shyness-sensitivity, aggression, and adjustment in urban Chinese adolescents at different historical times. *Journal of Research on Adolescence*, *22*(3), 393–399. https://doi.org/10.1111/j.1532-7795.2012.00790.x

LoBue, V., Pérez-Edgar, K., & Buss, K. A. (2021). Publications from the Longitudinal Attention and Temperament Study (LANTS). *Databrary*. https://nyu.databrary.org/volume/1288

Loken, E. (2004). Using latent class analysis to model temperament types. *Multivariate Behavioral Research*, *39*(4), 625–652. https://doi.org/10.1207/s15327906mbr3904_3

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist*, *50*(4), 266–275. https://doi.org/10.1037/0003-066X.50.4.266

Miller, N. V., Degnan, K. A., Hane, A. A., Fox, N. A., & Chronis-Tuscano, A. (2019). Infant temperament reactivity and early maternal caregiving: Independent and interactive links to later childhood attention-deficit/hyperactivity disorder symptoms. *Journal of Child Psychology and Psychiatry*, *60*(1), 43–53. https://doi.org/10.1111/jcpp.12934

Morgan, B. E. (2006). Behavioral inhibition: A neurobiological perspective. *Current Psychiatry Reports*, *8*(4), 270–278. https://doi.org/10.1007/s11920-006-0062-7

Muthén, B., & Muthén, L. (2017). MPlus. In W. J. van der Linden (Ed.), *Handbook of item response theory* (pp. 507–518). Chapman & Hall/CRC.

Nielsen, M. W., Alegria, S., Börjeson, L., Etzkowitz, H., Falk-Krzesinski, H. J., Joshi, A., Leahey, E., Smith-Doerr, L., Woollet, A. W., & Schiebinger, L. (2017). Gender diversity leads to better science. *Proceedings of the National Academy of Sciences*, *114*(8), 1740–1742. https://doi.org/10.1073/pnas.1700616114

Pérez-Edgar, K. (2018). Attention mechanisms in behavioral inhibition: Exploring and exploiting the environment. In K. Pérez-Edgar & N. A. Fox (Eds.), *Behavioral inhibition: Integrating theory, research, and clinical perspectives* (pp. 237–261). Springer.

Pérez-Edgar, K., & Fox, N. A. (2018). *Behavioral inhibition: Integrating theory, research, and clinical perspectives*. Springer.

Pérez-Edgar, K., LoBue, V., Buss, K. A., Field, A. P., & LANTs Team. (2021). Study protocol: Longitudinal attention and temperament study. *Frontiers in Psychiatry*, *12*, Article 656958. https://doi.org/10.3389/fpsyt.2021.656958

Planalp, E. M., Van Hulle, C., Gagne, J. R., & Goldsmith, H. H. (2017). The infant version of the Laboratory Temperament Assessment Battery (Lab-TAB): Measurement properties and implications for concepts of temperament. *Frontiers in Psychology*, *8*, Article 846. https://doi.org/10.3389/fpsyg.2017.00846

Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and assessment of Short and Very Short Forms of the Infant Behavior Questionnaire–Revised. *Journal of Personality Assessment*, *96*(4), 445–458. https://doi.org/10.1080/00223891.2013.841171

Putnam, S. P., & Stifter, C. A. (2005). Behavioral approach–inhibition in toddlers: Prediction from infancy, positive and negative affective components, and relations with behavior problems. *Child Development*, *76*(1), 212–226. https://doi.org/10.1111/j.1467-8624.2005.00840.x

Rothbart, M. K. (1989). Temperament in childhood: A framework. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 59–73). John Wiley & Sons.

Rubin, D. B. (1987). The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to The Data Augmentation Algorithm for creating a few imputations when fractions of missing information are modest: The SIR Algorithm. *Journal of the American Statistical Association*, *82*(398), 543–546. https://doi.org/10.2307/2289460

Sandstrom, A., Uher, R., & Pavlova, B. (2020). Prospective association between childhood behavioral inhibition and anxiety: A meta-analysis. *Research on Child and Adolescent Psychopathology*, *48*(1), 57–66. https://doi.org/10.1007/s10802-019-00588-5

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Chapman & Hall.

Schwartz, C. E., Wright, C. I., Shin, L. M., Kagan, J., & Rauch, S. L. (2003). Inhibited and uninhibited infants "grown up": Adult amygdalar response to novelty. *Science*, *300*(5627), 1952–1953. https://doi.org/10.1126/science.1083703

Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and "how to" guide of its application within vocational behavior research. *Journal of Vocational Behavior*, *120*, Article 103445. https://doi.org/10.1016/j.jvb.2020.103445

Stein, G. L., Cheah, C. S., Oh, W., & Witherspoon, D. P. (2023). Developmental science in the twenty-first century: Eschewing segregated science and integrating cultural and racial processes into research. In D. P. Witherspoon & G. L. Stein (Eds.), *Diversity and developmental science: Bridging the gaps between research, practice, and policy* (pp. 1–18). Springer.

Swanson, S. A., Lindenberg, K., Bauer, S., & Crosby, R. D. (2012). A Monte Carlo investigation of factors influencing latent class analysis: An application to eating disorder research. *International Journal of Eating Disorders*, *45*(5), 677–684. https://doi.org/10.1002/eat.20958

Tein, J.-Y., Coxe, S., & Cham, H. (2013). Statistical power to detect the correct number of classes in latent profile analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(4), 640–657. https://doi.org/10.1080/10705511.2013.824781

Vallorani, A., Fu, X., Morales, S., LoBue, V., Buss, K. A., & Pérez-Edgar, K. (2021). Variable- and person-centered approaches to affect-biased attention in infancy reveal unique relations with infant negative affect and maternal anxiety. *Scientific Reports*, *11*(1), Article 1719. https://doi.org/10.1038/s41598-021-81119-5

Vermunt, J. K. (2017). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, *18*(4), 450–469. https://doi.org/10.1093/pan/mpq025

Weiner, D. L., Balasubramaniam, V., Shah, S. I., Javier, J. R., & Pediatric Policy Council. (2020). COVID-19 impact on research, lessons learned from COVID-19 research, implications for pediatric research. *Pediatric Research*, *88*(2), 148–150. https://doi.org/10.1038/s41390-020-1006-3

Whalen, D. J., Sylvester, C. M., & Luby, J. L. (2017). Depression and anxiety in preschoolers: A review of the past 7 years. *Child and Adolescent Psychiatric Clinics of North America*, *26*(3), 503–522. https://doi.org/10.1016/j.chc.2017.02.006

Woodward, S. A., Lenzenweger, M. F., Kagan, J., Snidman, N., & Arcus, D. (2000). Taxonic structure of infant reactivity: Evidence from a taxometric perspective. *Psychological Science*, *11*(4), 296–301. https://doi.org/10.1111/1467-9280.00259

Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Frontiers in Psychology*, *5*, Article 920. https://doi.org/10.3389/fpsyg.2014.00920

Received February 28, 2023
Revision received September 1, 2023
Accepted October 27, 2023 ■