Development and Psychopathology (2022), **34**, 69–84
doi:10.1017/S0954579420000991

CAMBRIDGE UNIVERSITY PRESS

Regular Article

Latent class trajectories of infant temperament and associations with problem behavior at two years of age

Gerald F. Giesbrecht[1,2,3], Nicole Letourneau[1,4], Deborah Dewey[1,2] and the APrON Study Team

[1]Cumming School of Medicine (Pediatrics & Community Health Sciences), University of Calgary, Calgary, AB, Canada; [2]Alberta Children's Hospital Research Institute, Calgary, AB, Canada; [3]Department of Psychology, University of Calgary, Calgary, AB, Canada and [4]Faculty of Nursing, University of Calgary, Calgary, AB, Canada

Abstract

Individual differences in temperament have been well-described, but individual differences in temperament trajectories require elaboration. Specifically, it is unknown if subgroups of infants display different developmental patterns and if these patterns relate to later behavioral problems. The aims were to identify distinct developmental patterns in broad dimensions of temperament among typically developing infants, to determine whether these developmental patterns differ by sex, to evaluate how developmental patterns within each dimension of temperament relate to developmental patterns within other dimensions of temperament, and to determine whether developmental patterns of infant temperament are associated with internalizing and externalizing behavior at 2 years of age. Data from the longitudinal Alberta Pregnancy Outcomes and Nutrition study (n = 1,819) were used to model latent class trajectories of parent-reported infant temperament at 3, 6, and 12 months. Four to five unique latent trajectories were identified within each temperament dimension. Sex was not associated with trajectory groups. Developmental coordination was observed between trajectories of negative emotionality and regulatory capacity, and between regulatory capacity and positive affect, but not between positive affect and negative emotionality. Negative emotionality and regulatory capacity predicted internalizing and externalizing behavior. Patterns of development in infant temperament, and not just intensity of temperament, contribute toward later problem behavior.

Keywords: Alberta pregnancy outcomes and nutrition study, developmental trajectories, infant temperament, internalizing and externalizing behavior, latent class growth modeling

(Received 29 July 2019; revised 2 June 2020; accepted 3 June 2020)

Introduction

Classic work by Cicchetti, Sroufe, and others on the origins of developmental psychopathology has highlighted both the potential for diversity of outcomes given initial adaptations, and the coherence between earlier adaptations and later functioning (Cicchetti & Rogosch, 1996; Sroufe & Rutter, 1984). A major concern of the developmental approach to psychopathology has been to identify the proximal and distal links between early adaptation and later disorder (Rutter, 1984), including the contributions of temperament. Delineating longitudinal patterns of behavior over time has been a key method for linking the organization of behavior at one timepoint to outcomes at a later timepoint. Here, we take a longitudinal latent class approach to identify developmental patterns of temperament during the first year of life as they relate to behavioral problems in toddlerhood.

Author for Correspondence: Gerald F. Giesbrecht, Department of Pediatrics, University of Calgary, Child Development Center, #355, 3820–24 Ave, NW, Calgary, AB, Canada, T3B 2X9; E-mail: ggiesbre@ucalgary.ca.

Cite this article: Giesbrecht GF, Letourneau N, Dewey D, the APrON Study Team (2022). Latent class trajectories of infant temperament and associations with problem behavior at two years of age. *Development and Psychopathology* **34**, 69–84. https://doi.org/10.1017/S0954579420000991

Dimensions of temperament

Temperament, according to Rothbart's psychobiological model, refers to individual differences in the reactivity and regulation of emotional, motor, and attentional systems (Rothbart & Bates, 2006; Rothbart & Derryberry, 1981). Reactivity, which can be observed in the expression of positive and negative affect, entails processes such as the latency, rise time, intensity, and duration of response to stimulation. Regulation includes processes that serve to modulate reactivity. Individual differences in temperament constitute the early developmental substrate from which individuals organize their behaviors, and therefore interrogating the development of temperament can illuminate our understanding of the risk for developmental psychopathology (Rothbart & Posner, 2006).

Although there is disagreement on the definition of some fine-grained aspects of temperament, there is broad agreement regarding the fundamental dimensions (Rothbart, 1989). On the Infant Behavior Questionnaire—Revised (IBQ-R), which we focus on here, three overarching dimensions of temperament have been identified: (a) negative emotionality (NE), which refers to the expression of negative emotion such as fear, anger, irritability, sadness, or distress to limitation; (b) regulatory capacity/orienting (RCO), referring to processes that function to enhance or inhibit reactivity; and (c) positive affect/surgency (PAS), which refers to positive aspects of emotional reactivity that result in positive

emotion and engagement with the environment. Together, these three aspects of temperament describe the psychobiological systems that animate the characteristic behavioral and emotional expressions of infants.

Stability and change in infant temperament

To some extent, the utility and meaningfulness of temperament as a developmental construct depends on its individual and developmental continuity over time (Bornstein et al., 2015). Not surprisingly, questions about individual and temporal stability have been a major theme in temperament research. This body of work has shown moderate to strong stability in temperament during infancy (Bornstein et al., 2015), with a clear developmental sequence showing, for example, that individual differences in smiling are more fully present in early infancy than individual differences in fear (Brooker et al., 2013; Gartstein et al., 2010; Rothbart, 1986). Nevertheless, modest between-person stability over time implies that some individuals experience greater change than others. This raises the possibility of heterogeneity within normative patterns of stability and change such that subgroups of infants may follow distinct developmental pathways. Two observations strengthen this proposal. First, several studies using parent-report measures of infant temperament have observed overall modest age-related increases in both negative and positive emotional reactivity, with accompanying decreases in RCO ability over the course of infancy (Braungart-Rieker, Hill-Soderlund, & Karrass, 2010; Costa & Figueiredo, 2011; Gartstein & Rothbart, 2003; Holmboe, Nemoda, Fearon, Sasvari-Szekely, & Johnson, 2011; Mink, Henning, & Aschersleben, 2013). Second, previous longitudinal studies that fit linear growth models to parent-report temperament data observed significant residual variability in both intercept and slope (e.g., Bridgett et al., 2009; Erickson, Gartstein, & Beauchaine, 2017; Gartstein et al., 2010), suggesting that developmental patterns of temperament could be better modeled by allowing for more than one trajectory. Because change and stability are both expected developmental patterns, we anticipated that the majority of infants follow pathways characterized by modest change, whereas smaller subsets of infants follow trajectories with more dramatic change over time (Janson & Mathiesen, 2008).

Surprisingly few sex differences have been observed in either mean levels or temporal changes in fine-grained aspects of infant temperament; however, meta-analyses of the broad dimensions of NE, PAS, and RCO have revealed significant mean level sex differences in regulatory ability (higher in girls) and positive affect (higher in boys), with no sex differences in NE, although most of these differences emerge only after infancy (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006). Furthermore, similar temporal stability coefficients were observed for girls and boys during infancy, with girls showing somewhat higher stability for positive affect (Bornstein et al., 2015). Nevertheless, sex differences have been observed in older children, suggesting that sex differences in temperament may increase over time. Despite the few sex differences observed in previous infant work, it has not yet been determined if some developmental patterns of temperament are more likely for girls or boys. Furthermore, there are well-known sex differences in the occurrence of developmental psychopathology (Zahn-Waxler, Crick, Shirtcliff, & Woods, 2015), and to the extent that temperament contributes to later occurrence of developmental psychopathology, it is important to evaluate the role of sex in temperament trajectories.

Developmental patterns of infant temperament

Recent advances in longitudinal modeling make it possible to identify multiple developmental trajectories within a population and the approximate proportion of the population that follows each trajectory (Nagin & Tremblay, 2005). This approach is especially useful for identifying developmental patterns that unfold over time when the trajectory and distribution of these patterns in the population is not known, as is the case for temperament. Whereas traditional growth curve models assume that one trajectory (and associated residuals) adequately represents the underlying individual growth trajectories in the population, latent class growth modeling (LCGM) allows for the possibility that the population is composed of more than one trajectory, each of which describes the developmental pattern in a subset of individuals within the population (Nagin, 2005). LCGM is a specialized application of a finite mixture model designed to identify a finite number of distinctive developmental trajectories within the population. The LCGM deals with heterogeneity in development by identifying latent classes (unobserved groups) of individuals with similar patterns of development. We are aware of only one previous application of LCGM to temperament, in which Brooker and colleagues (Brooker et al., 2013) identified four latent class trajectories of stranger fear in children (6–36 months of age), which were meaningfully related to later behavioral inhibition.

Another approach that could advance our understanding of developmental patterns within infant temperament research is the simultaneous consideration of multiple aspects of temperament. The need to consider temperament as a functioning whole is implied by the notion that temperament entails interactive processes of both reactivity and regulation (Rothbart & Bates, 2006). There is a long tradition in temperament research for wholistic approaches, as exemplified in the classic work by Thomas and Chess classified children as "difficult," "easy," or "slow to warm up" based upon multiple dimensions of temperament (Thomas & Chess, 1977). More recent work has employed modern typological approaches, such as cluster analysis, to determine how aspects of temperament coalesce within individuals and their implications for development (Gartstein et al., 2017; Janson & Mathiesen, 2008). Although these modern approaches have the advantage of providing holistic analyses of temperament, studies that have used them typically conducted cross-sectional analyses and then addressed questions of development by examining differences in findings between younger and older infants. Few studies to date have examined longitudinally the co-development of broad temperament dimensions, including NE, RCO, and PAS in the same children.

There are a few notable examples in which analyses were conducted both longitudinally and with multiple dimensions of temperament, including linear growth modeling approaches to assess time-varying covariation between individual dimensions of temperament (e.g., Bridgett et al., 2009; Planalp & Braungart-Rieker, 2015), and longitudinal cluster analyses to model changes in multidimensional temperament profiles over time (Janson & Mathiesen, 2008). The later approach yielded a five-cluster classification of children from 18 months to 9 years of age in which the primary finding was significant individual stability in temperament profiles over time, supporting our proposal that most infants follow pathways characterized by modest change, whereas smaller subsets of infants follow trajectories with more dramatic change over time. Furthermore, patterns of association between the observed temperament profiles with concurrent internalizing and externalizing behaviors suggested that these profiles have utility

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

for understanding the connections between temperament and developmental psychopathology.

Temperament and developmental psychopathology

Previous work assessing the links between early life temperament and developmental psychopathology has produced valuable insight into the developmental origins of psychopathology (De Pauw & Mervielde, 2010; Gartstein et al., 2010; Gartstein, Putnam, & Rothbart, 2012; Janson & Mathiesen, 2008; Nigg, 2006; Rettew & McKee, 2005; Rothbart & Posner, 2006). Specifically, this work suggests that regulatory control may be a key dimension that distinguishes children with internalizing problems from children with externalizing problems, such that children with externalizing problems have less control than average whereas children with internalizing problems are over-controlled (Nigg, 2006). In contrast, temperament may also be a key resilience resource allowing children to overcome adversity, perhaps through physiological processes that increase the positive and regulatory aspects of temperament (Degnan & Fox, 2007). Thus, temperament may serve as either a risk or resilience factor for developmental psychopathology.

Nigg's (2006) excellent discussion of temperament as a risk or resilience factor for developmental psychopathology highlights difficulties inherent to studying their relations. Specifically, both are hypothetical constructs derived from similar types of data, which means that at least some of the observed empirical associations may derive from conceptual and empirical overlap between measures of temperament and psychopathology. Lemery and colleagues conducted both empirical and expert consensus analysis of overlap between measures of temperament and psychopathology, identifying content overlap in 9–38% of the items, depending on the subscale, but found that removing items with potential content overlap did not affect the relation between temperament and psychopathology (Lemery, Essex, & Smider, 2002). Furthermore, the modest correlations between temperament and psychopathology, even in studies relying upon the same rater for both measures, suggest that the constructs are related but not coterminous. Nigg (2006) estimates that no more than half of the reliable variance is shared between temperament and psychopathology measures, a value which suggests that these measures assess appreciably different constructs. In addition, there is considerable variability in psychopathology outcomes given initial temperament, and different aspects of temperament may nevertheless result in similar forms of psychopathology (Janson & Mathiesen, 2008). For example, the majority of infants with significant regulatory problems in early life (e.g., persistent and extreme crying) do not go on to develop psychopathology (Akhnikh, Engelberts, van Sleuwen, L'Hoir, & Benninga, 2014), although the rate of developmental psychopathology in this group is higher than children without these regulatory problems (Hemmi, Wolke, & Schneider, 2011). Thus, despite some conceptual overlap between temperament and developmental psychopathology, temperament theory and research have important potential to illuminate the developmental origins of psychopathology.

In keeping with our focus on broad dimensions of temperament, we focus on two higher-order dimensions of developmental psychopathology: internalizing and externalizing behavior. Internalizing behavior refers to processes within the self, such as anxiety, somatization, and depression. Children high in internalizing behaviors tend to respond to stressors by withdrawing, becoming anxious or depressed, or experiencing psychosomatic symptoms, such as stomachache without known medical cause. Externalizing behavior is characterized by actions directed toward the external world, such as antisocial, aggressive, or acting out behavior. Although children may present with behavior that is predominantly internalizing or externalizing, the moderate to strong positive association between them ($r \sim 0.50$) suggests the constructs are separable but partially overlapping (Achenbach, Ivanova, Rescorla, Turner, & Althoff, 2016).

Present study

In summary, there are gaps in our knowledge of the development and co-development of broad temperament dimensions during infancy and the implications of different developmental pathways for internalizing and externalizing behavior in early childhood. Furthermore, there is a need for studies that examine the range of developmental pathways in broad dimensions of infant temperament to better understand the processes of stability and change that may amplify or reduce risk for developmental psychopathology within typically developing infants. This is important because a better understanding of temperament development may help to improve assessment of early life risk for later psychopathology. Accordingly, the objective of the current study was to identify latent temperament trajectories that differentiate the developmental course of parent-reported temperament during the first year of life. Specifically, the aims of this analyses were to identify distinct developmental patterns in broad dimensions of temperament among typically developing infants, to determine whether these developmental patterns differ by sex, to evaluate how developmental patterns within each dimension of temperament relate to developmental patterns within the other dimensions of temperament, and to determine whether developmental patterns of infant temperament are associated with internalizing and externalizing behavior at 2 years of age.

Given previous longitudinal and cluster profile work, we had the following hypotheses. First, we expected that models with more than one latent trajectory would have better fit to the data than a model with only one latent trajectory. Based on previous work using latent classification techniques with cross-sectional data (Gartstein et al., 2017) and longitudinal cluster analyses (Janson & Mathiesen, 2008), it is reasonable to expect 2–5 latent subpopulations (i.e., unique trajectories) within each temperament dimension. Second, we expected moderate longitudinal stability with modest increases/decreases for trajectories that describe most infants, with some more sharply increasing or decreasing trajectories that describe developmental patterns for smaller subgroups of infants. Because previous studies suggest normative age-related increases in NE and PAS, and decreases in RCO (Gartstein et al., 2017; Gartstein & Rothbart, 2003; Mink et al., 2013), we expected to observe these overall patterns in our data for the majority of infants. Third, given evidence that infant temperament trajectories may follow a nonlinear pattern (e.g., Brooker et al., 2013; Gartstein et al., 2006, 2010), we expected that the majority of temperament trajectories would follow curvilinear paths (i.e., more rapid changes during some periods than others). Fourth, we did not expect to observe sex differences in temperament trajectories because most infant temperament studies have observed few differences between males and females. Fifth, we expected some trajectories within each dimension of temperament to have higher cross classification to specific trajectories within other dimensions of temperament. Nevertheless, given the lack of previous studies using this

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

approach, we only had a general expectation that trajectories with greater negativity would have higher cross classification to trajectories with lesser positive affect and regulation/orienting ability. Finally, we expected that children classified to different trajectories would differ in their levels of internalizing and externalizing behaviors in early childhood.

Method

Participants

Participants were 1,819 mother–infant pairs recruited to the Alberta Pregnancy Outcomes and Nutrition (APrON) study (Kaplan et al., 2014). APrON is a community sample of women recruited during pregnancy between June 2009 and July 2012 from two metropolitan areas in Alberta, Canada through advertisements in local media and by stationing research assistants in waiting rooms of high-volume maternity care and ultrasound clinics. Ethics approval for this study was obtained from the Health Research Ethics Boards at the University of Calgary (E22101) and the University of Alberta (pro00002954). Written consent was obtained from all women prior to data collection.

As shown in Table 1, the sample had low sociodemographic risk with the majority reporting a university education, household income above CAD $100,000, and cohabitating or married. The majority of mothers were mature (mean age 31.8 years), White, and had low levels of depression and anxiety. Mean gestational age at birth and birthweight were in the normal range (Kramer et al., 2001) and the sex ratio was consistent with the worldwide birth ratio (Hesketh & Xing, 2006).

Table 1. Sociodemographic characteristics for the study sample (n = 1,819)

	M (*SD*)	Range
Maternal age (years)	31.8 (4.2)	16.5–44.4
Infant gestational age at birth (weeks)	39.1 (1.7)	24.1–42.4
Birthweight (grams)	3,343 (523)	550–5,370
	%	
Infant sex (male)	53.4	
Married or cohabitating	96.9	
Maternal education		
University degree or higher	71.1	
Completed Trade or Technical Degree	18.7	
High school diploma or less	10.1	
Ethnicity		
Caucasian/White	80.8	
Asian	11.7	
Latin American	2.9	
Aboriginal	0.7	
Arab	0.5	
African Canadian	1.3	
Other	1.6	
Missing	0.5	
Annual household income		
More than $100,000/year	56.8	
$70,000–$100,000/year	22.8	
$40,000–$70,000/year	13.3	
Less than $40,000/year	7.1	

Procedure

A complete description of the recruitment procedures and methods for the APrON study are available elsewhere (Kaplan et al., 2014). Briefly, measures of maternal sociodemographics were collected during pregnancy (at the first study visit), along with measures of maternal nutrition and mental health, which are not reported here. Measures of birthweight and gestational age at birth were obtained from the medical birth record. Maternal report of infant temperament was collected at 3 months (T1), 6 months (T2), and 12 months (T3) of infant age. Completed questionnaires were returned by prepaid mailing envelope. Likewise, at 2 years (T4) of child age, mothers completed a measure of internalizing and externalizing behavior.

Measures

Temperament

The IBQ-R is a widely used multidimensional parent-report measure of infant temperament (Gartstein & Rothbart, 2003). The 36-item very short form of the IBQ-R (Putnam, Helbig, Gartstein, Rothbart, & Leerkes, 2013) was used to assess three broad dimensions of temperament: NE, RCO, and PAS, which have consistently emerged in infant studies (Gartstein, Knyazev, & Slobodskaya,2005). As per the standard scoring protocol, a mean score was calculated from the 12 items in each temperament dimension. Scores ranged from 1 (never) to 7 (always) with higher scores reflecting stronger evidence of each dimension. Cronbach alphas in the current study for NE at T1, T2, and T3 were 0.76, 0.78, and 0.76, respectively, for RCO 0.75, 0.76, and 0.75, and for PAS 0.84, 0.77, and 0.70.

Internalizing and externalizing behavior

Maternal report of child behavior problems was assessed at two years of age using the Child Behavior Checklist (CBCL) (Achenbach & Rescorla, 2010). The CBCL examines internalizing problems (e.g., anxiety/depression, somatic complaints, withdrawal—36 items), and externalizing problems (e.g., aggressive behavior, attention problems—24 items). The validity (Achenbach & Rescorla, 2010) and reliability (Kristensen, Henriksen, & Bilenberg, 2010) of the CBCL have been well supported. In our sample, Cronbach alphas were .78 and .88 for internalizing and externalizing problems, respectively. Because there tends to be sex differences in the frequency and severity of internalizing and externalizing behavior, raw scores were standardized separately for males and females prior to analyses. The resulting analysis makes it less likely that any associations between temperament trajectory and behavior problems could result from mean sex differences in behavior problems but preserves the possibility that associations between temperament trajectories and behavior problems may differ by sex.

Data analysis

Analysis proceeded through five steps. First, the number of latent trajectories that best described the data was determined separately for NE, RCO, and PAS (i.e., univariate trajectories). We used both

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

theory and previous studies to inform our modeling of the trajectories. We tested the assumption that more than one latent trajectory was needed to model the data by comparing the fit of a single class model to the fit of multiple class models. Selection of the best model was based on the smallest absolute Bayesian information criteria (BIC) value and the BIC log Bayes factor approximation. Our objective was to identify a finite set of latent trajectories that describe the distinct features of stability and change in infant temperament in order to gain new insight into the range of developmental patterns in a community sample of infants. Second, each infant was assigned to one trajectory group based on highest posterior probability of the latent trajectory. For clarity, we refer to the output of the LGCM model itself as latent trajectories or latent classes, whereas we use the term "trajectory group" to indicate the manifest assignment of infants based on the latent trajectories. Third, we used chi-square analysis to determine if trajectory group membership differed as a function of sex. Fourth, to examine how the development of each dimension of temperament evolves contemporaneously with other aspects of temperament, we assessed dual trajectory models in which trajectories for NE, RCO, and PAS identified in Step 1 (above) were simultaneously fit to the data. Finally, we evaluated the associations between the trajectory groups and measures of internalizing and externalizing behaviors using a two-way analysis of covariance (ANCOVA) with trajectory group and sex as fixed factors and child age at T4 (24 months) as a covariate.

Temperament trajectories across T1–T3 were estimated using the Proc Traj procedure (Jones, Nagin, & Roeder, 2001; Nagin, 1999) in SAS 9.4. We chose Proc Traj because it makes simplifying assumptions, including homogeneity of individual growth trajectories within each class, which sets the within-class variance and covariance to zero. There is a pragmatic advantage to this approach, especially when previous data and theory are insufficient to specify, a priori, the within-class variance–covariance structure, as is the case for developmental patterns of infant temperament. All models were initially fit with quadratic trajectories and nonsignificant parameters were trimmed in the final parsimonious models (linear models were also tested but had inferior fit). An extension of the Proc Traj procedure allows for modeling the contemporaneous linkages between the latent trajectories for different dimensions of temperament, which we used to probe how development within one dimension of temperament relates to development within other dimensions of temperament. Like the univariate trajectories, in the dual trajectory analysis, the model generates the parameters defining the latent trajectories as well as the probability of membership in each trajectory. Additionally, this model estimates the probabilities linking membership in latent trajectories across the two dimensions of temperament. The probabilities describing the linkage across temperament dimensions provide a metric for describing the degree of overlap in the developmental course of the two dimensions of temperament (Nagin, 2005).

Our analyses and presentation of the results were guided by the Guidelines for Reporting on Latent Trajectory Studies (GRoLTS) (van de Schoot, Sijbrandij, Winter, Depaoli, & Vermunt, 2017). To ensure that the final models were stable and not based on local maxima, we estimated models using 50 different start values (Jones et al., 2001). We centered time at 3 months, with time as a continuous measure of infant age, so that intercepts in the models refer to infant temperament at 3 months.

Results

Descriptive findings

Sample sizes at 3, 6, 12, and 24 months were 1,731, 1,525, 1,442, and 1,131, respectively, with a total of 1,819 different infants participating across the three waves of temperament data (i.e., some infants not assessed in earlier waves were present in later waves, and vice versa). Findings for temperament trajectories (Aims 1–3) were based on $n = 1{,}819$, whereas analyses of internalizing and externalizing (Aim 4) were based on an $n = 1{,}131$. Descriptive statistics for study variables are displayed in Table 2; means and standard deviations are presented separately for boys and girls. Bivariate correlations are reported in Supplementary Table S1. Girls had higher NE at 6 months and lower PAS at 6 and 12 months. RCO did not differ by sex at any time point. At 24 months, girls had higher internalizing scores compared to boys, $p = .02$, and boys had higher externalizing scores, although not statistically different from girls, $p = .08$.

Missing data

In total, 19.9% of values were missing. Participants with complete data did not differ from those with incomplete data on infant sex or birth characteristics (gestational age, birth order, delivery mode, being admitted to a neonatal intensive care unit), but those with missing data were more likely to be immigrants to Canada, to be a minority (non-White), have lower household income and lower education, and mothers who were older, $p < .05$. Given these findings, missing data were estimated using a maximum likelihood estimator in SAS 9.4.

Univariate LCGM trajectory models

We report all models tested, beginning with a one trajectory model and proceeding until model fit deteriorated. Six models were tested for NE and RCO, and five models for PAS. As shown in Table 3, models with only one trajectory had the poorest fit for each dimension of temperament, suggesting that multiple latent trajectories better represent the data than a single overall trajectory. Model performance and distribution of the sample into the trajectories are reported in Table 4. Mean posterior probabilities (entropy) for the NE, RCO, and PAS trajectories were 0.74, 0.74, and 0.76, respectively, indicating adequate classification. Plots of individual trajectories showed that the final LCGM models were reasonable representations of the individual trajectories (see Supplementary Figures S1–S3). Parameter estimates for the final univariate trajectory models are presented in Table 5. Overall, the models demonstrated very good fit to the data.

NE

A five-trajectory model had best fit for NE (Table 3). As shown in Figure 1, three of the five NE trajectories suggest mostly stable negativity over infancy at relatively low (trajectory 1), moderate (trajectory 2) and high (trajectory 5) levels of negativity. The model estimated that 91.1% of infants in the population follow one of these three relatively stable trajectories, with 480, 1,040, and 214 infants in our data classified to trajectories 1, 2, and 5, respectively. As expected, the model classified most infants (78.7%) to latent trajectories (1 and 2) with modestly increasing negativity during infancy. A small proportion (2.5%) were classified to a latent trajectory with sharply decreasing negativity (trajectory 3, $n = 21$ in our sample) and another small proportion

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

Table 2. Descriptive statistics for temperament and behavior variables

Timepoint	Infant age (months) mean, range (*SD*)	Scale	Mean (*SD*) Males	*n*	Females	*n*	*F* test (*p* value)
T1	2.7, 1.15–5.7 (0.51)	NE	3.5 (0.97)	928	3.6 (0.96)	803	1.2 (0.28)
		RCO	5.3 (0.72)		5.3 (0.71)		0.03 (0.86)
		PAS	3.3 (1.1)		3.3 (1.0)		0.0 (0.99)
T2	6.2, 4.2–11.7 (0.71)	NE	3.5 (0.91)	815	3.6 (0.93)	710	4.6 (0.03)
		RCO	5.3 (0.69)		5.3 (0.68)		0.0 (0.99)
		PAS	4.6 (0.87)		4.5 (0.90)		6.9 (0.01)
T3	12.3, 10.2–17.1 (0.86)	NE	4.0 (0.94)	781	4.0 (0.93)	661	0.02 (0.96)
		RCO	5.0 (0.73)		5.0 (0.70)		2.5 (0.12)
		PAS	5.4 (0.69)		5.3 (0.66)		11.6 (0.001)
T4	24.3, 20.1–34.3 (1.1)	Internalizing behavior	4.7 (4.1)	605	5.3 4.4	526	5.3 (0.02)
		Externalizing behavior	10.1 (6.7)		9.4 6.4		3.1 (0.08)

Note: NE = negative emotionality; RCO = regulatory capacity/orienting; PAS = positive affectivity/surgency (PAS). *F* tests are for differences between males and females.

Table 3. Model fit results for univariate unconditional trajectory models with different number of trajectory classes

	BIC	$2Log_e(B_{10})$
Negative emotionality (NE)		
1	−6,170.3	
2	−5,890.1	560.4
3	−5,838.1	103.9
4	−5,834.1	8.1
5*	−5,830.2	7.8
6	−5,835.8	−11.2
Regulatory capacity/orienting (RCO)		
1	−5,001.7	
2	−4,737.4	528.58
3	−4,669.0	136.76
4	−4,661.9	14.34
5*	−4,650.3	23.2
6	−4,657.6	−14.74
Positive affect/surgency (PAS)		
1	−5,494.6	
2	−5,134.3	720.5
3	−5,035.4	197.8
4*	−5,018.9	33.04
5	−5,019.0	−0.18

Note: *Best fitting models are indicated with an asterisk. BIC = Bayesian information criterion; $2Log_e(B_{10})$ = BIC log Bayes factor approximation comparing current model to the model with one less class. Best fitting models for NE, RCO, and PAS converged in 132, 95, and 101 iterations, respectively.

Table 4. Trajectory class proportions and model adequacy indicators for final univariate trajectory models

Trajectory class	Proportion	Number of infants	Entropy (mean posterior probability)
Negative emotionality (NE)			
1. Low	0.281	480	0.80
2. Moderate	0.506	1,040	0.76
3. Decreasing	0.025	21	0.67
4. Increasing	0.064	64	0.73
5. High	0.124	214	0.75
Regulatory capacity/orienting (RCO)			
1. Low	0.023	36	0.80
2. Moderate	0.223	380	0.73
3. Increasing	0.050	55	0.65
4. High decreasing	0.491	1,009	0.74
5. High stable	0.214	339	0.79
Positive affect/surgency (PAS)			
1. Low	0.096	129	0.78
2. Low-moderate	0.384	765	0.72
3. High-moderate	0.372	697	0.72
4. High	0.148	228	0.82

Note: Proportion refers to the estimated proportion of the population that belong to each trajectory class. Number refers to the number of individuals in the sample assigned to each trajectory (based on highest posterior probability).

(6.4%) to a trajectory with strongly increasing negativity (trajectory 4, $n = 64$ in our sample). As suggested by Figure 1, intercepts of the five latent trajectories differed significantly from each other, except for trajectories 2 and 4, and 3 and 5 (see Supplementary Table S2). All NE trajectory slopes differed significantly from each other, except trajectories 1 and 2.

RCO

A five-trajectory model had the best fit for RCO (Table 3). Four (trajectories 1, 2, 4, and 5) out of the five trajectories were flat

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

Table 5. Parameter estimates for final latent class growth models for NE, RCO, and PAS

Trajectory class	Parameter	Estimate	95% CI
Negative emotionality (NE)			
1.	Intercept	2.724	2.62–2.82
	Linear	−0.029	−0.07–0.008
	Quadratic	0.008	0.006–0.10
2.	Intercept	3.542	3.45–3.64
	Linear	−0.028	−0.06–0.007
	Quadratic	0.009	0.007–0.01
3.	Intercept	4.715	4.03–5.40
	Linear	−0.531	−0.82–−0.24
	Quadratic	0.037	0.017–0.06
4.	Intercept	3.778	3.43–4.12
	Linear	0.387	0.22–0.56
	Quadratic	−0.026	−0.04–−0.01
5.	Intercept	4.928	4.76–5.10
	Linear	−0.149	−0.24–−0.06
	Quadratic	0.016	0.01–0.02
Regulatory capacity/orienting (RCO)			
1.	Intercept	3.770	3.60 −3.94
2.	Intercept	4.796	4.67–4.93
	Linear	−0.044	−0.06–−0.03
3.	Intercept	4.405	4.18–4.63
	Linear	0.274	0.18–0.37
	Quadratic	−0.017	−0.02–0.01
4.	Intercept	5.427	5.35–5.50
	Linear	−0.011	−0.04–0.02
	Quadratic	−0.003	−0.005–−0.001
5.	Intercept	5.922	5.84–6.01
	Linear	0.037	0.004–0.07
	Quadratic	−0.005	−0.007–−0.003
Positive affect/surgency (PAS)			
1.	Intercept	2.322	2.17–2.48
	Linear	0.310	0.23–0.39
	Quadratic	−0.009	−0.02–−0.001
2.	Intercept	2.798	2.65–2.95
	Linear	0.521	0.48–0.56
	Quadratic	−0.028	−0.03–−0.02
3.	Intercept	3.761	3.57–3.95
	Linear	0.418	0.38–0.46
	Quadratic	−0.023	−0.03–−0.02
4.	Intercept	4.958	4.80–5.11
	Linear	0.240	0.19–0.29
	Quadratic	−0.014	−0.02–−0.01

Note: The estimates reported here are the basis of Figures 1–3 for NE, RCO, and PAS, respectively.

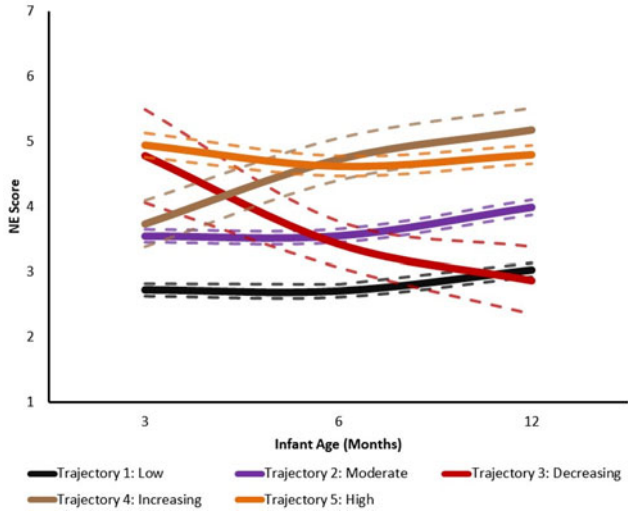

Figure 1. Predicted trajectories of negative emotionality for infants at 3, 6, and 12 months. Lines represent the means; dashed lines indicate the 95% confidence intervals. Predicted proportion of the population classified to trajectory is as follows: 1–28.1%; 2–50.6%; 3–2.5%; 4–6.4%; 5–12.4%. NE = negative emotionality.

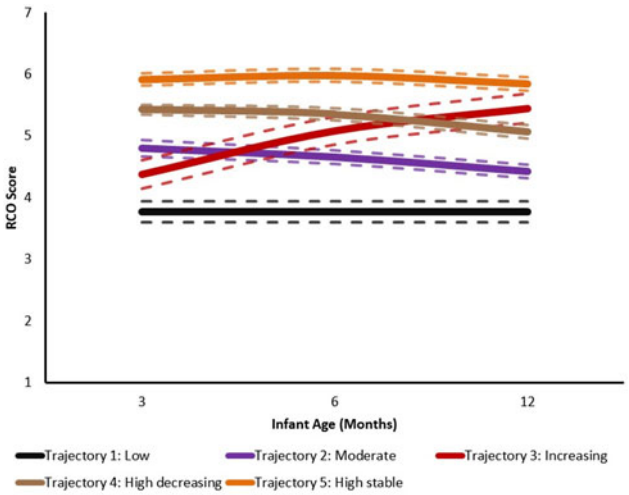

Figure 2. Predicted trajectories of regulatory capacity/orienting (RCO) for infants at 3, 6, and 12 months. Lines represent the means; dashed lines indicate the 95% confidence intervals. Predicted proportion of the population classified to each trajectory is as follows: 1–2.3%; 2–22.3%; 3–5.0%; 4–49.1%; 5–21.4%.

or slightly decreasing (see Figure 2), suggesting overall rank-order stability within most RCO trajectories. The model estimated that 95% of infants in the population, $n = 36$, 380, 1,009, and 339 for trajectories 1, 2, 4, and 5, respectively in our sample. The exception was latent trajectory 3 (estimated at 5% of the population; $n = 55$ in our sample), which displayed clearly increasing RCO. Intercepts and slopes for each latent trajectory differed significantly from the intercepts and slopes of each other trajectory (see Supplementary Table S2).

PAS

A four-trajectory model had the best fit for PAS (Table 3). All four latent PAS trajectories increased over the first year of infancy (see Figure 3). Most infants, estimated as 75.6% of the population ($n = 1462$ in our sample), were represented by the middle two

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

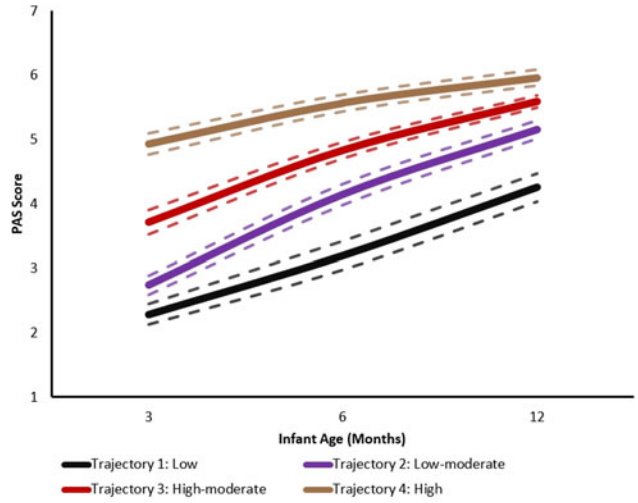

Figure 3. Trajectories of positive affectivity/surgency (PAS) for infants at 3, 6, and 12 months. Lines represent the means; dashed lines indicate the 95% confidence intervals. Predicted proportion of the population classified to each trajectory is as follows: 1–9.6%; 2–38.4%; 3–37.2%; 4–14.8%.

trajectories (trajectories 2 and 3), indicating overall moderate and increasing PAS during infancy. The Low PAS trajectory (trajectory 1) accounted for 9.6% of the population ($n = 129$ in our sample) and the High trajectory (Trajectory 4) for 14.8% of the population ($n = 228$ in our sample). Intercepts and slopes for each latent trajectory differed significantly from the intercepts and slopes of each other trajectory (see Supplementary Table S2). As shown in Figure 3, the trajectories tended to converge such that mean differences at 12 months were smaller than at 3 months, except for latent trajectory 1, which had a lower slope than the other trajectories.

Test of sex differences

Although we observed some sex difference in mean NE and PAS levels (see Table 2), we found no evidence that infant sex was related to trajectory group assignment in any of the temperament dimension, χ^2 (4, $N = 1,819$) = 3.54, $p = .47$ for NE; χ^2 (4, $N = 1,819$) = 1.03, $p = .91$ for RCO; and χ^2 (3, $N = 1,819$) = 1.75, $p = .63$ for PAS.

Dual trajectory models

We examined the developmental linkages between NE, RCO, and PAS trajectories by estimating the joint probability of trajectory classification across these dimensions of temperament. These analyses are intended to evaluate the pattern of coordination across two dimensions of temperament to determine whether some combinations of trajectories are more likely than others. One would expect all probabilities to be more or less equal if in fact there is no pattern of developmental coordination. In order to conduct all pairwise joint trajectory modeling for the three temperament dimensions, three separate analyses were conducted (i.e., NE–RCO, NE–PAS, and RCO–PAS). As recommended (Nagin & Tremblay, 2001), the dual trajectory models were estimated with the final models identified as optimal in the univariate trajectory analyses above. Results of this analysis are reported in Table 6. Table 6 should be read by first selecting a column, and then noting the distribution of probabilities within the column (which total to 1). For example, in Panel A, the column for the Low NE trajectory class indicates the probability for classification to each of the latent RCO trajectories for those individuals classified to the Low NE trajectory.

RCO–NE

Table 6, Panel A, reports the probability of membership in each of the five latent RCO trajectories as a function of membership in the five latent NE trajectories. Probabilities ranged from 0.01 to 0.61, indicating that some RCO trajectories are very unlikely and others very likely for each NE trajectory. Two overall patterns emerged. First, there was clear evidence of developmental coordination, such that each NE trajectory was probabilistically linked to one or two RCO trajectories, with low probabilities for other RCO trajectories. This suggests that each NE trajectory has relatively strong developmental linkage to one or two RCO trajectories and relatively low developmental linkage with other RCO trajectories. Classification to the Moderate NE trajectory, for example, had 0.61 probability of cross-classification to the High decreasing RCO trajectory and a 0.22 probability of cross-classification to the Moderate decreasing RCO trajectory. Thus, the probability of classification to a moderate or high RCO trajectory is 0.83 given classification to a moderate NE trajectory. The exception to the general pattern of developmental coordination was the Increasing NE trajectory class, which had relatively equal probabilities for four of the five RCO trajectories, suggesting poor specificity of RCO trajectories for the Increasing NE trajectory.

Second, we observed different patterns of developmental linkage at the lower and higher range of NE. For Low NE, there was clear evidence of negative coordination, as can be seen by the strong probabilities for being in one of the higher RCO trajectories. Likewise, classification to the Increasing NE trajectory had the highest probabilities for classification to either of the two decreasing RCO trajectories. These findings suggest a robust negative coordination between NE and RCO for trajectories at the low end of the NE scale such that being low in NE was strongly related to being high in RCO. In contrast, classification to one of the high NE trajectories was paired with cross-classification to one of the moderate or high RCO trajectories, suggesting positive coordination between NE and RCO such that trajectories displaying high levels of negativity during infancy were connected to trajectories with relatively high levels of regulation/orienting ability.

PAS–NE

There was a consistent pattern of probabilities between NE and PAS (see Table 6, Panel B) in which all NE trajectories had strong to very strong probabilities of cross classification to the two moderate PAS trajectories, suggesting little differentiation of PAS trajectories as a function of NE trajectories. The exception was the Increasing NE trajectory, for which low to moderate PAS trajectories had the highest probability. In addition, there was evidence of positive coordination between the NE and PAS such that trajectories with higher NE were cross classified to higher PAS trajectories, and vice versa.

PAS–RCO

Probabilities of membership in latent PAS trajectories as a function of latent RCO trajectories are reported in Table 6, Panel C. Here we see a clear example of positive coordination, as indicated by the strong probabilities along the diagonal from upper left to lower right. These findings suggest strong developmental

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

Table 6. Probabilities linking classification in dual trajectory models

Panel A. Probability of RCO class conditional on NE class					
	NE trajectory class				
RCO trajectory class	Low	Moderate	Decreasing	Increasing	High
1. Low	.01	.01	.04	.08	.04
2. Moderate	.09	.22	.09	.25	.35
3. Increasing	.03	.02	.44	.18	.02
4. High decreasing	.42	.61	.11	.31	.42
5. High stable	.44	.13	.32	.18	.17
Panel B. Probability of PAS class conditional on NE class					
	NE trajectory class				
PAS trajectory class	Low	Moderate	Decreasing	Increasing	High
1. Low	.16	.10	.04	.37	.00
2. Low-moderate	.39	.38	.46	.46	.36
3. High-moderate	.32	.41	.38	.00	.35
4. High	.12	.11	.12	.17	.29
Panel C. Probability of PAS class conditional on RCO class					
	RCO trajectory class				
PAS trajectory class	Low	Moderate	Increasing	High decreasing	High stable
1. Low	.60	.32	.16	.06	.04
2. Low-moderate	.40	.45	.69	.47	.06
3. High-moderate	.00	.21	.10	.39	.45
4. High	.00	.02	.05	.07	.45

Note: This table reports the probabilities linking each trajectory class within the dimension of temperament listed in the left most column (represented by rows) with the trajectory classes within the dimension of temperament listed in the right columns (represented by columns). The table should be read by column. Each column (within each panel) totals to 1, indicating that each temperament dimension within a column is completely distributed across each of the temperament dimension indicated by rows. That is, all infants assigned to a trajectory (column) are also assigned to a trajectory (row) on another dimension of temperament. Greyed cells highlight strongest probabilities. NE = negative emotionality; RCO = regulatory capacity/orienting; PAS = positive affectivity/surgency.

coordination between PAS and RCO such that trajectories with the lowest regulatory ability were developmental coordinated with trajectories that had the lowest PAS, and vice versa. What is particularly striking is that each RCO trajectory had a moderate to strong probability of connections with only one or two of the PAS trajectories, and low or very low probabilities for other PAS trajectories, indicating a high degree of specificity in the developmental coordination.

Temperament trajectories and internalizing and externalizing behavior

The final set of analyses examined the associations between the manifest temperament trajectory groups and internalizing and externalizing behavior. Despite the null effect for sex within the trajectory groups, we included sex in our test of association with behavior problems because there were sex differences in internalizing and externalizing behaviors (see Table 2). A two-way ANCOVA with trajectory group and sex as fixed factors and child age at T4 (24 months) as a covariate was used to assess trajectory group differences and interactions between sex and trajectory group on internalizing and externalizing behavior.

Overall models were significant for associations between NE and RCO with both internalizing and externalizing behavior, suggesting developmental linkages with these temperament trajectories, but not with PAS (see Supplementary Table S3). None of the group by sex interactions was significant, $p > .05$ (results not shown).

Estimated marginal means of internalizing and externalizing behavior for each trajectory group are displayed in Figure 4 (see also Supplementary Table S4). Note that scores on internalizing and externalizing behavior were separately centered at the mean for boys and girls to eliminate mean sex differences in behavior problem scores, thus scores below zero indicate below average problems and scores above zero indicate above average problems. Panel A shows that two NE trajectory groups (Low—trajectory 1 and Decreasing—trajectory 3) had below average levels of internalizing and externalizing behaviors, which may indicate a protective effect of these trajectories for later behavior problems. Two NE trajectory groups (Increasing—trajectory 4 and High—trajectory 5) had above average levels of internalizing and externalizing

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

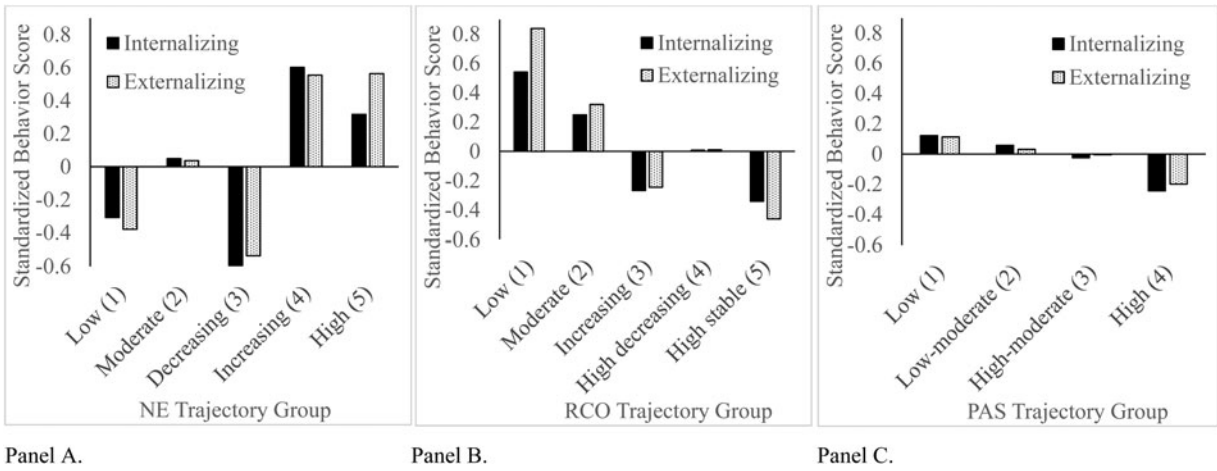

Figure 4. Estimated marginal means of internalizing and externalizing behavior by temperament trajectory group with sex and child age at T4 as covariates. Behavior scores were standardized separately by sex. Internalizing and externalizing behavior were mean centered so that bars extending above and below zero indicate above average and below average behavior problems, respectively, relative to the sample mean. Panel A—internalizing and externalizing behavior by NE trajectory groups. Panel B—internalizing and externalizing behavior by RCO trajectory groups. Panel C—internalizing and externalizing behavior by PAS trajectory groups.

behavior, suggesting that they may be linked to increased risk for behavior problems. Within each of the NE trajectory groups, internalizing and externalizing behavior were relatively equivalent, with the exception of the High trajectory group (trajectory 5) where externalizing behavior was relatively more elevated than internalizing, t (125) = 2.43, p = .02. This suggests that high negativity throughout infancy is a relatively greater risk for later externalizing compared to internalizing behavior. Infants in the Moderate group (trajectory 2) did not differ from the sample mean for either internalizing or externalizing behavior.

Among the RCO trajectory groups, Figure 4, Panel B, two groups were associated with high internalizing and externalizing behavior (Low—trajectory 1 and Moderate—trajectory 2) and two with low internalizing and externalizing behavior (Increasing—trajectory 3 and High stable—trajectory 5), suggesting that these temperament trajectories are risk or resilience factors, respectively, for later behavior problems. Externalizing was elevated relatively more than internalizing in the Low (trajectory 1) group, although this comparison had small sample size and the results did not reach statistical significance, t (22) = 1.4, p = .17. In the High stable (trajectory 5) group, externalizing was relatively lower than internalizing, t (200) = 2.1, p = .037. The High decreasing (trajectory 4) group had average internalizing and externalizing.

Figure 4, Panel C shows that all PAS trajectory groups had internalizing and externalizing scores that were close to the sample mean. The High PAS group had relatively lower internalizing and externalizing symptoms compares to the other trajectory groups.

Discussion

The aims of this study were to identify distinct developmental patterns in broad dimensions of temperament among typically developing infants, to determine whether these developmental patterns differ by sex, to evaluate how developmental patterns within each dimension of temperament relate to developmental patterns within the other dimensions of temperament, and to determine whether developmental patterns of infant temperament are associated with internalizing and externalizing behavior at 2 years of age. To accomplish these aims, we used a LCGM approach with a large community sample of infants. Our findings make several contributions, as described below.

Identifying distinct developmental patterns

Traditional multilevel modeling approaches, which have the ability to model only one overall developmental pattern, are an analytic mismatch with the majority of theorizing and conceptualizations of temperament and psychopathology, where the concept of risk and resilience is often evoked to postulate different developmental outcomes that may ensue from seemingly similar circumstances. The LCGM approach makes both empirical and conceptual/theoretical advances by identifying different developmental patterns of infant temperament. This approach allows one to model both the shape of the developmental pattern (the pattern of change/stability over time) and to estimate the approximate proportion of infants in the population who may follow one of these developmental patterns.

Consistent with our expectation that models with more than one latent trajectory would have better fit to the data than a model with a single latent trajectory, our findings revealed that the developmental patterns of temperament among a nonclinical sample of infants can be described by four or five distinct trajectories. Our findings suggest that the overarching dimensions of temperament develop along 4–5 different longitudinal trajectories, each with different patterns of initial level and stability/change over time. Furthermore, our analyses estimated the proportion of infants in the population following each of the developmental patterns, allowing for new insight into the range of developmental patterns of infant temperament in a community sample. Based on considerable evidence suggesting moderate longitudinal stability in parent-reported temperament (Bornstein et al., 2015), our expectation that the majority of infants would be classified to trajectories with little or modest longitudinal change was confirmed. Furthermore, we anticipated several trajectories with substantial change in temperament, that when combined with the modestly changing trajectories would recapitulate findings of overall moderate longitudinal stability

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

within the population. These expectations were largely confirmed across all three broad dimensions of temperament. The developmental patterns were consistent with previous reports of overall longitudinal stability/change, with mean increases in NE and PAS along with mean decreases in RCO (Gartstein et al., 2017; Gartstein & Rothbart, 2003; Holmboe et al., 2011; Mink et al., 2013; Olafsen et al., 2008).

Our analyses also identified trajectories that differed from the mostly parallel and slowly changing patterns of temperament that were displayed by most infants. These trajectories had sharply increasing or decreasing temperament scores, with much of this change occurring in early infancy, as shown by the generally steeper slopes between 3 and 6 months in Figures 1–3. Within the NE trajectories we identified a small predicted proportion of infants in the population who had among the highest NE scores at 3 months and the lowest NE scores at 12 months. Although the more typical pattern for infants with high NE at 3 months is to remain high in NE throughout infancy, the model predicted a small portion of these infants experience dramatic reductions in NE. We speculate that this latent NE class may overlap with the small group of infants who display persistent fussing and crying in the first few months and which resolves by 6 months of age (Wolke, Bilgin, & Samara, 2017). This may be a productive area for future research.

The model also predicted a second small group of infants with a significant increase in NE that was much larger than usual. We wonder to what extent this latent trajectory may represent what Chess and Thomas (Chess & Thomas, 2013) referred to as poorness of fit, which they used to describe a kind of vulnerability some children display when the characteristics and demands of the environment differ markedly from their capabilities or tendencies. In their view, poorness of fit has direct implications for overall developmental pattern, including the trajectory for temperament itself and ultimately for psychopathology. As we describe below, infants classified to this trajectory group later displayed relatively high internalizing and externalizing behavior at 2 years of age (see Figure 4), supporting our speculation that a large increase in negativity over the first year is an early indicator of developmental risk for later internalizing and externalizing problems. Together, our findings for NE suggest that developmental patterns of temperament (e.g., sharply increasing or decreasing trajectories of infant negativity) and not only intensity of temperament may contribute to risk and resilience. Furthermore, these findings point toward the potential for interventions that shift developmental trajectories to increase resilience among children.

Among the RCO trajectories, we identified two latent trajectories that were unusual in the sense that only a small predicted proportion of infants were classification to them. One class of infants had relatively low RCO throughout infancy, and although the trajectory was stable, which is in keeping with the normative pattern, what sets this trajectory class apart was its unusually low level of RCO. It is important to note that although RCO scores within this class were low relative to the other classes, the actual scores were near the middle of the 7-point scale. Thus, it is not the case that these infants fail to demonstrate any regulatory or attentional capabilities, but they seem to do so much less than other infants. Analysis of fine-grained aspects of temperament may be useful for determining if infants classified to this trajectory are consistently low across various specific aspects of RCO or if they are particularly low on one or two aspects. The other unusual RCO trajectory had significantly increasing RCO that was initially (at 3 months) among the lowest in the sample and ended (at 12 months) among the highest. Interestingly, this trajectory had strong developmental linkage with the decreasing NE trajectory, suggesting that it is not only the intensity of temperament expression that is coordinated (see for example Putnam, Rothbart, & Gartstein, 2008), but also that developmental processes within reactive and regulatory aspects of temperament co-regulate each other over time. These findings point toward the need for more studies that go beyond the "balance" between regulatory and reactive aspects of temperament to examine how development within one aspect of temperament may contribute to the emergence of a new behavioral repertoire within other aspects of temperament.

In contrast to the NE and RCO trajectories, we did not identify any strongly diverging trajectories for PAS regarding either developmental pattern or overall level. The highest and lowest trajectory classes had relatively smaller predicted proportions of infants, 14.8% and 9.6% respectively, however, the developmental pattern was very similar across all trajectory groups. These findings are intriguing because they suggest that, in contrast to NE and RCO, individual differences in trajectories of PAS are more a matter of degree (i.e., the quantification of positive affect) than of different patterns of development. In other words, after accounting for initial level of PAS there may only be one overall trajectory for PAS during infancy, with individual differences in how much PAS infants display very early in development, and relatively stable increases in positivity regardless of initial levels.

Sex differences

Consistent with the majority of infant temperament studies (Bornstein et al., 2015; Else-Quest et al., 2006), we found little evidence of sex differences. Although we did observe some small mean sex differences in temperament, especially for NE, and for internalizing behavior at age 2 years, the manifest trajectory groups were not associated with infant sex and the relations between the trajectory groups and internalizing and externalizing behavior did not differ by sex. To the extent that the novel trajectories identified in our analysis capture the developmental patterns in typically developing infants, our findings indicate that male and female infants are equally likely to follow any of these developmental trajectories. These findings are consistent with previous work using different modeling methods (Gartstein et al., 2017; Janson & Mathiesen, 2008), which bolsters the conclusion that the development of temperament during infancy is similar for boys and girls. Nevertheless, there may be sex differences in developmental trajectories for more fine-grained aspects of temperament that should be examined in future work, as shown by Gartstein and colleagues who reported steeper increases in parent-reported fear among females (Gartstein et al., 2010).

Developmental coordination among temperament dimensions

Our interest in multiple dimensions of temperament is motivated by the possibility that development within one dimension of temperament constrains and/or enables co-development within other dimensions of temperament. Although previous work has examined trait-by-trait interactions (Gartstein et al., 2012), our analysis focused on the relations between the *developmental patterns* within different dimensions of temperament. Our modeling of co-development between broad temperament dimensions generated unique and valuable insights into developmental coordination between dimensions of temperament by showing how developmental patterns are interrelated across dimensions. If

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

there were no developmental coordination between trajectories, then one would expect all combinations of trajectories across dimensions to have equal probability. Higher probability estimates in these analyses indicate that classification to a latent trajectory within one dimension of temperament is strongly coordinated with classification to a specific latent trajectory on another dimension of temperament. Moderate probability estimates across multiple latent trajectories may suggest poor differentiation or specificity of trajectories in other dimensions of temperament given trajectories within one dimension of temperament. In contrast, moderate or strong probability estimates for only one or two trajectories may suggest developmental coordination across temperament dimensions. Low probabilities indicate dissociation of a specific latent trajectories in one dimension of temperament from latent trajectories in the other dimension of temperament.

Our analyses revealed different patterns of developmental coordination among and within the broad temperament dimensions. Developmental coordination between NE and RCO differed for trajectories with low versus high NE. Latent trajectories with relatively lower negativity were cross classified to latent trajectories exhibiting relatively higher regulatory ability, whereas trajectories with high negativity tended to be cross classified to trajectories displaying moderate or high regulatory ability. This suggests an interaction effect, as has been proposed previously (Nigg, 2006; Rothbart & Bates, 2006), such that the association between NE and RCO depends on the level of NE. In contrast, the developmental coordination between PAS and RCO was entirely positive, such that individuals with low positive affect were also more likely to follow a low regulation trajectory and vice versa.

These analyses reveal important differences in developmental coordination between the regulatory and reactive aspects of temperament. The concept of regulatory capacity, as applied to temperament, refers to processes that function to enhance or inhibit reactivity across both positive and negative dimensions (Rothbart, 1986). Our findings indicate that, at least during infancy, developmental coordination between RCO and NE may be primarily inhibitory such that infants classified to a relatively higher RCO trajectory tend to be cross classified to a lower NE trajectory. In contrast, the coordination between RCO and PAS may be primarily facilitative such that infants classified to a relatively higher RCO trajectory also tend to be classified to a higher PAS trajectory. These findings are consistent with the overall effects of regulatory capacity on behavior, which allows the infant to implement effective coping strategies by flexibly redirecting attention away from threatening or non-rewarding inputs and increasing attention to rewarding or relieving inputs (Derryberry & Rothbart, 1997).

Developmental coordination between NE and PAS was complex, with some trajectory groups showing more differentiation than others. Poor differentiation was observed for some NE trajectories with cross classifications distributed across many PAS trajectories, meaning that knowing which individuals were probabilistically classified to a specific NE trajectory provided little insight into their classification to one of the PAS trajectories. Nevertheless, there was also evidence of canalization in that the two moderate PAS trajectories were the most likely outcome of classification to any of the NE trajectories. In other words, infants had the highest probability of assignment to these moderate PAS trajectories regardless of their membership in the NE trajectories. To some extent, this result may arise from the fact that most infants (75.6%) were classified to the middle two trajectories. Finally, there was evidence of positive developmental coordination for NE and PAS for infants with high NE. Because both PAS and NE refer to reactive aspects of temperament, these findings suggest that common underlying developmental processes (i.e., morphological integration) could regulate reactivity of both positive and negative affect.

Trajectory groups and internalizing and externalizing behavior

Building upon a long tradition of developmental trajectory analysis that focuses on specific sub-populations to describe their trajectory over time (e.g., the work of Jerome Kagan), we show that manifest subgroups of children have unique developmental trajectories, and these trajectories have potential clinical meaning. For example, our analysis revealed that some trajectory groups had stronger implications for later internalizing and externalizing problems than other patterns. Findings consistent with those we report here have been produced using other methods (e.g., cluster analysis), providing converging evidence using rather different methods. Our findings make a unique and valuable contribution to identifying the main developmental patterns of infant temperament in the general population and strengthen the conclusions that subgroups of infants in the population follow different patterns of development, with meaningful implications of these patterns for developmental psychopathology.

Our findings highlight the utility of examining developmental patterns of temperament, as opposed to the only focusing on the intensity of temperament at a specific timepoint. As we discuss below, in some cases the intensity of temperament was not as relevant as the developmental pattern to later problem behavior. Overall, our findings suggest that developmental patterns in NE and RCO, but not PAS, predict later internalizing and externalizing behavior, which is largely consistent with findings from studies that focused more on temperament intensity than developmental patterns (Edwards & Hans, 2015; Gartstein et al., 2012).

NE

Consistent with our proposal that both developmental pattern and intensity of temperament make important contributions to later problem behavior, we found that trajectory groups with overall high NE and those with a pattern of increasing NE later displayed elevated internalizing and externalizing problems. Likewise, infants assigned to a low or decreasing NE trajectory groups demonstrated lower than average internalizing and externalizing symptoms.

In general, there was poor specificity between internalizing and externalizing problems within the NE trajectory groups, reinforcing the notion that NE tends to influence the occurrence of problem behavior rather than the type of problem behavior (Nigg, 2006). Still, it is instructive to note that infants assigned to the High NE trajectory had relatively greater elevations in externalizing compared to internalizing problems. This may reflect a tendency for highly negative infants to express their negativity behaviorally, but it contrasts to previous work suggesting an overall stronger association between infant NE (assessed cross-sectionally during infancy) and internalizing ($r = .24$) rather than externalizing ($r = .17$) symptoms at age 2 years (Gartstein et al., 2012). It may be instructive in future work to compare cross-sectional and longitudinal associations to determine the extent to which associations between temperament and later behavior problems differ as a function of measurement paradigm.

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

RCO

There appears to be a clear inverse gradient of association between RCO during infancy and later behavior symptoms. Infants assigned to the lowest two RCO trajectory groups had elevated internalizing and externalizing scores, with relatively higher scores among infants with the lowest RCO scores. In contrast, infants assigned to high or increasing RCO trajectory groups had the lowest internalizing and externalizing scores, with the lowest scores among infants classified to the highest RCO trajectory. These findings suggest a risk/resilience gradient across the RCO trajectory groups, consistent with previous reports using regression analysis (Gartstein et al., 2012). Of interest, the Increasing RCO trajectory group had lower than average internalizing and externalizing behaviors but the High decreasing RCO trajectory group did not. These two trajectory groups had approximately equivalent overall RCO scores, providing an example where the developmental pattern may be more important than the overall level.

Similar to what was observed for the NE trajectory groups, there was relatively little differentiation between internalizing and externalizing symptoms based on the RCO trajectory groups. Among infants in the Low RCO trajectory group, externalizing was relatively more elevated than internalizing, although this comparison had small sample size and the results did not reach statistical significance. If this finding can be replicated, it may indicate that a low RCO trajectory is a specific risk factor for later externalizing behavior, which is consistent with the notion that under-control is a specific risk factor for externalizing behavior (Olson, Choe, & Sameroff, 2017; Olson, Sameroff, Kerr, Lopez, & Wellman, 2005). At the other end of the RCO scale, infants assigned to the High stable trajectory group had relatively more symptoms of internalizing than externalizing behavior, which is consistent with the notion that internalizing behavior is associated with overcontrol. Nevertheless, it is important to note that high regulatory ability appears to have an overall protective effect for later internalizing and externalizing behavior in our sample, which suggest that these infants may not be overcontrolled.

PAS

Although we did not observe marked elevations or reductions in behavior problems for any of the PAS trajectory groups, we note that some aspects of PAS may contribute to both internalizing (Dougherty, Klein, Durbin, Hayden, & Olino, 2010) and externalizing (Rothbart & Bates, 2006) problems in older children, and that other aspects of temperament in combination with PAS may either increase or decrease the association with problem behavior (Dougherty et al., 2010; Janson & Mathiesen, 2008). Furthermore, fine grained aspects of PAS (such as sociability and interest) may help to discriminate between children with anxiety and depression (Lonigan, Carey, & Finch, 1994), and other aspects such as high-intensity pleasure and shyness may regulate the conditional probability of internalizing and externalizing problems (Oldehinkel, Hartman, De Winter, Veenstra, & Ormel, 2004). We also note that our analysis of PAS trajectory groups and behavior problems may have been affected by differential dropout of children from the highest PAS trajectory group. Chi-square analysis of missing data found greater than expected missing data from the highest PAS trajectory group, χ^2 (4) = 24.30, $p < .001$, potentially reducing power for testing associations with behavior problems. Further study of the potential connections between PAS and behavior problems should conducted, especially among older children where associations may emerge.

Overall, the contribution of this developmental analysis of temperament to our understanding of developmental psychopathology is highlighted by the fact that these temperament trajectories are meaningfully related to later behavior problems above and beyond the intensity of temperament, at least for NE and RCO but less so for PAS. Developmental patterns in temperament provide a window into the processes by which behaviors become problematic before the appearance of overt behavior problems. Although our analysis cannot reveal why some infants may follow one particular trajectory and not another, we expect that the early life environment, especially parenting (Ryan & Ollendick, 2018), plays a significant role. We further speculate that infant temperament trajectories are the mediators that link the early life environment to later problem behavior. This speculation is strengthened by recent evidence from a relationship and co-parenting intervention study conducted during the perinatal period in which the effect of intervention on decreased internalizing behavior at 24 months of age was mediated by decreases in NE at 12 months of age (Tomfohr et al., 2020).

It is interesting to note that the level of temperament at the endpoint of the temperament trajectories (i.e., at 12 months of age) corresponds to the level of behavior problems at 2 years of age, suggesting that where infants "end up" on dimensions of temperament at 12 months of age may be sufficient to estimate later behavior problems.[1] Nevertheless, the trajectory analysis we conducted lends insight into the developmental pathways by which infants "arrive" at various levels of temperament at 12 months of age. Identifying these developmental pathways could be exploited to guide prevention efforts aimed at altering temperament at earlier stages of development to decrease the probability of later behavior problems.

Strengths and limitations

The novelty of our latent class trajectory analysis of infant temperament data, our linkage of temperament trajectories to internalizing and externalizing behavior, and our large sample size are strengths of this work. The group-based and person-centered analyses conducted here are developmental in their approach and therefore have several advantages over previous analyses. These advantages include the ability to identify qualitatively distinct subgroups within the population (which may not otherwise be identified based on theory or estimation of a single population growth curve), the ability to assess developmental linkages between dimensions of temperament, and the ability to formally test the relationship between the trajectories and later outcomes.

Nevertheless, some limitations of these analyses should be acknowledged when interpreting the findings. First, LCGM is a useful analytic tool for approximating the unknown population distribution of trajectories and for classifying individuals into different trajectory groups, but it should be recognized from the outset that these trajectory groups are a statistical abstraction, and one should avoid reifying them (Nagin, 2005). To make the problem of understanding developmental pathways more tractable, the LCGM approach identifies a limited set of latent trajectories that captures the main features of the underlying trajectories within the population, but in reality, the individual trajectories have a continuous distribution. Thus, the groups are heuristic, and their utility should be judged based on the insights that they yield.

[1]We thank an anonymous reviewer for this observation.

https://doi.org/10.1017/S0954579420000991 Published online by Cambridge University Press

Second, the latent class analysis we conducted made several important simplifying assumptions (e.g., no within-class heterogeneity, zero variance–covariance structure), which were necessary to make the analysis tractable, but which overlook potentially important variability in the data. Supplementary Figures S1–S3 suggest there is within-trajectory variability. Our findings provide a useful starting point for replication studies that use more flexible approaches, such as growth mixture modeling, where the number and shape of each trajectory must be known a priori.

Third, our classification of infants to manifest groups, for the purpose of evaluating sex differences and implications for behavior problems, ignores uncertainty in group assignment. Although entropy in our model was adequate, there remains some uncertainty in the correct assignment of infants to manifest groups. As a result, our analysis likely underestimates the associations of trajectory groups with sex and behavior (van de Schoot et al., 2017).

Fourth, our community sample has relatively low sociodemographic risk, and because temperament is known to interact with environment (Bridgett et al., 2009; Chess & Thomas, 2013; Gartstein et al., 2010), it is not known to what extent similar results would be produced in a higher risk sample. Low sociodemographic risk samples offer insights into the developmental processes that are operating in large portions of the population, and they provide a comparison group for evaluating the similarities and differences in the pathways to developmental psychopathology for those children with greater exposure to sociodemographic risk factors. From a public health perspective, it is essential to understand the risks for poor outcome across all sociodemographic strata, including those deemed to be low risk, which represents many families in Canada, the United States, and many other first world countries.

Fifth, parent-report measures have limitations, such as recall bias, and may be influenced by characteristics of the parent. Such limitations may be partially offset by selecting a well-validated and reliable measure, such as the IBQ-R. A comparison of longitudinal measures of parent-report and laboratory-assessed infant fear produced remarkably similar findings for developmental patterns in fear and associations with later anxiety problems, decreasing concerns that perceptions of parents may be driving the findings (Gartstein et al., 2010). Nevertheless, the finding here would be strengthened by replication with observational data, although it would be challenging to repeatedly assess all three of the broad dimensions of temperament we have included here. For practical reasons, our analyses were limited to broad dimensions of temperament. The group-based modeling approach taken here can also be applied to more fine-grained aspects of temperament, and we believe such analyses would provide further insight into the processes by which infants differ from each other in the development of temperament.

Sixth, measures of temperament and measures of internalizing and externalizing behavior have at least some content and construct overlap, raising the possibility that measurement artifacts contribute to the observed associations (Egger & Angold, 2006; Nigg, 2006). Still, empirical analysis and expert consensus suggest that measures of temperament and psychopathology are clearly separable, and that removing items with potential content overlap does not affect the relation between temperament and psychopathology (Lemery et al., 2002). Our study also cannot rule out shared genetic contributions to both temperament and psychopathology. Furthermore, there are ongoing transactions between the infant acting on the environment and the environment acting on the infant such that the Person × Environment interactions over time may shape temperament behaviors into problem behaviors.

Finally, there is a need for advancing methods within the LCGM approach to allow for simultaneously modeling of multiple dimensions of temperament while at the same time estimating the associations between the groupings defined by these multiple trajectories with outcomes such as internalizing and externalizing behavior. This would allow for more nuanced conclusions about the ways that infant temperament as a multidimensional construct is associated with later behavioral outcomes.

Clinical implications

Early identification of temperament trajectories could inform early intervention efforts aimed as shifting children away from problematic trajectories. We found evidence suggesting that temperament trajectories could serve as either risk or resilience factors—increasing or decreasing later behavior problems. This was especially the case for those temperament trajectory groups at the extremes of the temperament dimension and for trajectory groups that displayed a great deal of change over infancy. These findings have at least two implications for early intervention. First, they suggest that it may be possible to incorporate information about infant temperament trajectories into the identification of those who may be at risk for later behavior problems. Identifying infants with increasing versus decreasing negativity, for example, could help to differentiate the risk profile for later behavior problems. Our data suggest that decreasing negativity is a significant resilience factor for later behavior problems, whereas increasing negativity is a significant risk factor. Second, early identification of temperament trajectories could inform later intervention efforts by helping to target the content of intervention toward the strengths and weaknesses of the child's temperament. For example, it may require relatively less effort to increase regulatory capabilities than decrease negative affectivity in a child low in the former and high in the latter. In contrast, a child who is high in both NE and regulatory capabilities might benefit from strategies that capitalize on existing regulatory ability to select alternate emotional responses or increase the latency and decrease the duration of negative emotions.

Conclusion

These findings suggest that developmental patterns of infant temperament can be summarized by 4–5 different latent trajectories and that both intensity and pattern of stability/change in temperament operate as risk and resilience factors for later problem behavior.

Supplementary Material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954579420000991.

Acknowledgments. The authors gratefully acknowledge the participants of the Alberta Pregnancy Outcomes and Nutrition (APrON) study and the APrON Study Team, whose members include B.J. Kaplan, C.J. Field, D. Dewey, R.C. Bell, F.P. Bernier, M. Cantell, L.M. Casey, M. Eliasziw, A. Farmer, L. Gagnon, L. Goonewardene, D.W. Johnston, L. Kooistra, D.P. Manca, J.W. Martin, L.J. McCargar, M. O'Beirne, V.J. Pop, and N. Singhal.

Financial Support. This work was supported by the Alberta Children's Hospital Foundation, Alberta Innovates Health Solutions, and a Career Development Award from the Canadian Child Health Clinician Scientist Program (GFG).

Conflicts of Interest. The authors declare they have no conflicts of interest.

References

Achenbach, T. M., Ivanova, M. Y., Rescorla, L. A., Turner, L. V., & Althoff, R. R. (2016). Internalizing/externalizing problems: Review and recommendations for clinical and research applications. *Journal of the American Academy of Child Adolescent Psychiatry*, *55*, 647–656. doi:10.1016/j.jaac.2016.05.012

Achenbach, T. M., & Rescorla, L. A. (2010). *Manual for the ASEBA preschool forms and profiles*. Burlington, VT: University of Vermont.

Akhnikh, S., Engelberts, A. C., van Sleuwen, B. E., L'Hoir, M. P., & Benninga, M. A. (2014). The excessively crying infant: Etiology and treatment. *Pediatric Annals*, *43*, e69–75. doi:10.3928/00904481-20140325-07

Bornstein, M. H., Putnick, D. L., Gartstein, M. A., Hahn, C. S., Auestad, N., & O'Connor, D. L. (2015). Infant temperament: Stability by age, gender, birth order, term status, and socioeconomic status. *Child Development*, *86*, 844–863. doi:10.1111/cdev.12367

Braungart-Rieker, J. M., Hill-Soderlund, A. L., & Karrass, J. (2010). Fear and anger reactivity trajectories from 4 to 16 months: The roles of temperament, regulation, and maternal sensitivity. *Developmental Psychology*, *46*, 791–804. doi:10.1037/a0019673

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. doi:10.1016/j.infbeh.2008.10.007

Brooker, R. J., Buss, K. A., Lemery-Chalfant, K., Aksan, N., Davidson, R. J., & Goldsmith, H. H. (2013). The development of stranger fear in infancy and toddlerhood: Normative development, individual differences, antecedents, and outcomes. *Developmental Science*, *16*, 864–878. doi:10.1111/desc.12058

Chess, S., & Thomas, A. (2013). *Goodness of fit: Clinical applications, from infancy through adult life*. Taylor & Francis.

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, *8*, 597–600.

Costa, R., & Figueiredo, B. (2011). Infant's psychophysiological profile and temperament at 3 and 12 months. *Infant Behavior and Development*, *34*, 270–279. doi:10.1016/j.infbeh.2011.01.002

Degnan, K. A., & Fox, N. A. (2007). Behavioral inhibition and anxiety disorders: Multiple levels of a resilience process. *Development and Psychopathology*, *19*, 729–746. doi:10.1017/S0954579407000363

De Pauw, S. S., & Mervielde, I. (2010). Temperament, personality and developmental psychopathology: A review based on the conceptual dimensions underlying childhood traits. *Child Psychiatry & Human Development*, *41*, 313–329. doi:10.1007/s10578-009-0171-8

Derryberry, D., & Rothbart, M. K. (1997). Reactive and effortful processes in the organization of temperament. *Development and Psychopathology*, *9*, 633–652. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9448999

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*, 462–488. doi:10.1521/jscp.2010.29.4.462

Edwards, R. C., & Hans, S. L. (2015). Infant risk factors associated with internalizing, externalizing, and co-occurring behavior problems in young children. *Developmental Psychology*, *54*(4), 489–499. doi:10.1037/a0038800

Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry*, *47*, 313–337. doi:10.1111/j.1469-7610.2006.01618.x

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin*, *132*, 33–72. doi:10.1037/0033-2909.132.1.33

Erickson, N. L., Gartstein, M. A., & Beauchaine, T. P. (2017). Infant predictors of toddler effortful control: A multi-method developmentally sensitive approach. *Infant and Child Development*, *26*, e1971. doi:10.1002/icd.1971

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. doi:10.1037/a0018898

Gartstein, M. A., Gonzalez, C., Carranza, J. A., Ahadi, S. A., Ye, R., Rothbart, M. K., & Yang, S. W. (2006). Studying cross-cultural differences in the development of infant temperament: People's Republic of China, the United States of America, and Spain. *Child Psychiatry & Human Development*, *37*, 145–161.

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. https://doi.org/10.1016/j.infbeh.2004.09.003

Gartstein, M. A., Prokasky, A., Bell, M. A., Calkins, S., Bridgett, D. J., Braungart-Rieker, J., … Seamon, E. (2017). Latent profile and cluster analysis of infant temperament: Comparisons across person-centered approaches. *Developmental Psychology*, *53*, 1811–1825. doi:10.1037/dev0000382

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. doi:10.1002/imhj.21312

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

Hemmi, M. H., Wolke, D., & Schneider, S. (2011). Associations between problems with crying, sleeping and/or feeding in infancy and long-term behavioural outcomes in childhood: A meta-analysis. *Archives of Disease in Childhood*, *96*, 622–629. doi:10.1136/adc.2010.191312

Hesketh, T., & Xing, Z. W. (2006). Abnormal sex ratios in human populations: Causes and consequences. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 13271–13275. doi:10.1073/pnas.0602203103

Holmboe, K., Nemoda, Z., Fearon, R. M., Sasvari-Szekely, M., & Johnson, M. H. (2011). Dopamine D4 receptor and serotonin transporter gene effects on the longitudinal development of infant temperament. *Genes, Brain and Behavior*, *10*, 513–522. doi:10.1111/j.1601-183X.2010.00669.x

Janson, H., & Mathiesen, K. S. (2008). Temperament profiles from infancy to middle childhood: Development and associations with behavior problems. *Developmental Psychology*, *44*, 1314–1328. doi:10.1037/a0012713

Jones, B. L., Nagin, D. S., & Roeder, K. (2001). A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods & Research*, *29*, 374–393. doi:10.1177/0049124101029003005

Kaplan, B. J., Giesbrecht, G. F., Leung, B. M., Field, C. J., Dewey, D., Bell, R. C., … Team, A. P. S. (2014). The Alberta Pregnancy Outcomes and Nutrition (APrON) cohort study: Rationale and methods. *Maternal and Child Nutrition*, *10*, 44–60. doi:10.1111/j.1740-8709.2012.00433.x

Kramer, M. S., Platt, R. W., Wen, S. W., Joseph, K. S., Allen, A., Abrahamowicz, M., … Fetal/Infant Health Study Group of the Canadian Perinatal Surveillance, S. (2001). A new and improved population-based Canadian reference for birth weight for gestational age. *Pediatrics*, *108*, E35. doi:10.1542/peds.108.2.e35

Kristensen, S., Henriksen, T. B., & Bilenberg, N. (2010). The Child Behavior Checklist for Ages 1.5–5 (CBCL/1½–5): Assessment and analysis of parent-and caregiver-reported problems in a population-based sample of Danish preschool children. *Nordic Journal of Psychiatry*, *64*, 203–209. doi:10.3109/08039480903456595

Lemery, K. S., Essex, M. J., & Smider, N. A. (2002). Revealing the relation between temperament and behavior problem symptoms by eliminating measurement confounding: Expert ratings and factor analyses. *Child Development*, *73*, 867–882. doi:10.1111/1467-8624.00444

Lonigan, C. J., Carey, M. P., & Finch, A. J. (1994). Anxiety and depression in children and adolescents - negative affectivity and the utility of self-reports. *Journal of Consulting and Clinical Psychology*, *62*, 1000–1008. doi:10.1037/0022-006x.62.5.1000

Mink, D., Henning, A., & Aschersleben, G. (2013). Visual habituation tasks: The role of infant temperament. *Infant Behavior and Development*, *36*, 377–390. doi:10.1016/j.infbeh.2013.03.005

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*, *4*, 139–157. doi:10.1037/1082-989X.4.2.139

Nagin, D. S. (2005). *Group-based modeling of development*. Cambridge, MA: Harvard University Press.

Nagin, D. S., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods*, *6*, 18–34. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11285809 doi:10.1037/1082-989X.6.1.18

Nagin, D. S., & Tremblay, R. E. (2005). What has been learned from group-based trajectory modeling? Examples from physical aggression and other problem behaviors. *Annals of the American Academy of Political and Social Science*, *602*, 82–117. doi:10.1177/0002716205280565

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

Olafsen, K., Kaaresen, P., Handegard, B., Ulvund, S., Dahl, L., & Rønning, J. (2008). Maternal ratings of infant regulatory competence from 6 to 12 months: Influence of perceived stress, birth-weight, and intervention: A randomized controlled trial. *Infant Behavior & Development*, *31*, 408–421.

Oldehinkel, A. J., Hartman, C. A., De Winter, A. F., Veenstra, R., & Ormel, J. (2004). Temperament profiles associated with internalizing and externalizing problems in preadolescence. *Development and Psychopathology*, *16*, 421–440. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15487604 doi:10.1017/S0954579404044591

Olson, S. L., Choe, D. E., & Sameroff, A. J. (2017). Trajectories of child externalizing problems between ages 3 and 10 years: Contributions of children's early effortful control, theory of mind, and parenting experiences. *Development and Psychopathology*, *29*, 1333–1351. doi:10.1017/S095457941700030X

Olson, S. L., Sameroff, A. J., Kerr, D. C., Lopez, N. L., & Wellman, H. M. (2005). Developmental foundations of externalizing problems in young children: The role of effortful control. *Development and Psychopathology*, *17*, 25–45. doi:10.1017/S0954579405050029

Planalp, E. M., & Braungart-Rieker, J. M. (2015). Trajectories of regulatory behaviors in early infancy: Determinants of infant self-distraction and self-comforting. *Infancy*, *20*, 129–159. doi:10.1111/infa.12068

Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2013). Development and assessment of short and very short forms of the Infant Behavior Questionnaire-Revised. *Journal of Personality Assessment*, *96*(4), 445–458. doi:10.1080/00223891.2013.841171

Putnam, S. P., Rothbart, M. K., & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development: An International Journal of Research and Practice*, *17*(4), 387–405.

Rettew, D. C., & McKee, L. (2005). Temperament and its role in developmental psychopathology. *Harvard Review of Psychiatry*, *13*, 14–27. doi:10.1080/10673220590923146

Rothbart, M. K. (1986). Longitudinal observation of infant temperament. *Developmental Psychology*, *22*, 356–365. doi:10.1037/0012-1649.22.3.356

Rothbart, M. K. (1989). Temperament and development. In G. A. Kohnstamm, J. E. Bates & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 187–247). New York: John Wiley and Sons, Inc.

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

Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. Brown (Eds.), *Advances in developmental psychology* (1st ed., Vol. *1*, pp. 33–86). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rothbart, M. K., & Posner, M. I. (2006). Temperament, attention, and developmental psychopathology. In D. C. a. D. Cohen (Ed.), *Developmental psychopathology* (2nd ed., Vol. *2*, pp. 465–501). Hoboken, NJ: John Wiley & Sons Inc.

Rutter, M. (1984). Continuities and Discontinuities in Socioemotional Development. In R. N. Emde & R. J. Harmon (Eds.), *Continuities and discontinuities in development* (pp. 41–68). Boston, MA: Springer US.

Ryan, S. M., & Ollendick, T. H. (2018). The interaction between child behavioral inhibition and parenting behaviors: Effects on internalizing and externalizing symptomology. *Clinical Child and Family Psychology Review*, *21*, 320–339. doi:10.1007/s10567-018-0254-9

Sroufe, L. A., & Rutter, M. (1984). The domain of developmental psychopathology. *Child Development*, *55*, 17–29. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6705619 doi:10.2307/1129832

Thomas, A., & Chess, S. (1977). *Temperament and development*. Temperament and development. xv, 270 pp. Oxford, England: Brunner/Mazel; England.

Tomfohr, L., Giesbrecht, G. F., Madsen, J. W., MacKinnon, A., Le, Y., & Doss, B. (2020). Improved child functioning following brief relationship enhancement and co-parenting interventions during the transition to parenthood. *International Journal of Environmental Research and Public Health*, *17*, 1–16.

van de Schoot, R., Sijbrandij, M., Winter, S. D., Depaoli, S., & Vermunt, J. K. (2017). The GRoLTS-checklist: Guidelines for reporting on latent trajectory studies. *Structural Equation Modeling*, *24*(3), 451–467. doi:10.1080/10705511.2016.1247646

Wolke, D., Bilgin, A., & Samara, M. (2017). Systematic review and meta-analysis: Fussing and crying durations and prevalence of colic in infants. *The Journal of Pediatrics*, *185*, 55–61e54. doi:10.1016/j.jpeds.2017.02.020

Zahn-Waxler, C., Crick, N. R., Shirtcliff, E. A., & Woods, K. E. (2015). The origins and development of psychopathology in females and males. *Developmental Psychopathology: Volume One: Theory and Method*, 76–138. doi:10.1002/9780470939383.ch4