Published in final edited form as:

Child Dev. 2016 November ; 87(6): 1940–1955. doi:10.1111/cdev.12561.

A Twin Factor Mixture Modeling Approach to Childhood Temperament: Differential Heritability

Brandon G. Scott

Montana State University

Kathryn Lemery-Chalfant, Sierra Clifford, Jenn-Yun Tein, and Ryan Stoll

Arizona State University

H. Hill Goldsmith

University of Wisconsin-Madison

Abstract

Twin factor mixture modeling was used to identify temperament profiles, while simultaneously estimating a latent factor model for each profile with a sample of 787 twin pairs ($M_{age}=7.4$ years; $SD=.84$; 49% female; 88.3% Caucasian), using mother- and father-reported temperament. A 4-profile, 1-factor model fit the data well. Profiles included ‘Regulated, Typical Reactive’, ‘Well-regulated, Positive Reactive’, ‘Regulated, Surgent’, and ‘Dysregulated, Negative Reactive.’ All profiles were heritable, with heritability lower and shared environment also contributing to membership in the ‘Regulated, Typical Reactive’ and ‘Dysregulated, Negative Reactive’ profiles.

Keywords

temperament; twin; heritability; factor mixture modeling

Temperament forms the affective core of later personality and predicts a wide range of outcomes in childhood and beyond, which include self-regulation, social competence and psychopathology (Rothbart & Bates, 2006). Investigators conceptualize and measure temperament using either a dimensional or a categorical approach (Shiner & DeYoung, 2013) and both approaches offer distinct and valuable perspectives. However, different theoretical and statistical assumptions, combined with different operational definitions, measurement, and samples (community-based vs. extreme group; Shiner & DeYoung, 2013), make integrating findings from the two approaches challenging.

We used a twin factor mixture modeling (TFMM) approach (Muthén, Asparouhov, & Rebollo, 2006) to identify latent temperament profiles in twin children, while simultaneously modeling the underlying factor structure. Factor mixture modeling integrates the categorical and dimensional approaches by simultaneously modeling both heterogeneous, qualitatively distinct temperament categories and within-category individual

differences on multiple temperament dimensions. We further used the twin method to elucidate the genetic and environmental etiology of profiles and factors.

Overview of Temperament

Temperament is commonly defined as individual differences in reactivity and regulation, and theorized to be both relatively stable and at least partially biologically-based, although both biological and environmental factors contribute to temperament (Rothbart & Bates, 2006). Middle childhood involves the maturation of biological systems underlying self-regulation, behavioral inhibition, and approach, combined with increasing exposure to unfamiliar, structured and challenging environments and interactions with peers (Rueda et al., 2004; Shiner, 1998; Shiner & Caspi, 2003), making middle childhood an important time to study temperament. In particular, the middle years of elementary school (7-8 years of age) are a time of relative environmental stability, when children are established in school and have not yet reached the environmental changes characterizing the transition to adolescence (e.g., Larson, Richards, Moneta, Holmbeck, & Duckett, 1996), and most children have not yet experienced pubertal maturation and associated neurological development (Biro et al., 2010; Romeo, 2003). Individual differences in both reactive and regulatory aspects of temperament during these years may lay the groundwork for risk and resilience in multiple domains (e.g., social, academic) during the greater developmental upheaval of adolescence (Shiner & Caspi, 2003). Although some evidence shows stability of children's classification in certain temperamental profiles from infancy through early childhood (Beekman et al., 2015), few studies have taken a person-centered approach to temperament in the formative school years (e.g., Janson & Mathiesen, 2008) and none to our knowledge have considered indices of self-regulation as well as reactivity.

Dimensional and Categorical Views of Temperament

Both dimensional, variable-centered perspectives and categorical, person-centered perspectives impact modern temperament theory (Kagan, Snidman, Arcus, & Reznick, 1994; Rothbart & Bates, 2006). For example, Rothbart's effortful control factor is prominent in interdisciplinary research (Shiner et al., 2012), and predicts fewer behavior problems, particularly externalizing (Rothbart & Bates, 2006). On the categorical side, research suggests that highly behaviorally inhibited children are distinct from those in the moderate range; they are consistently highly inhibited across contexts and at risk for anxiety (Kagan et al., 1994). We believe that neither a dimensional nor a categorical approach is so powerful as to exclude the other approach; rather, both provide valuable information about the structure, etiology, and consequences of temperament.

Dimensional Models of Temperament

Dimensional approaches captures temperament with multiple distinct but related dimensions, measured continuously and typically derived using exploratory and confirmatory factor analysis. Rothbart's hierarchical model consists of three higher-order latent factors (surgency, negative affectivity, and effortful control), each composed of more specific, observable dimensions (e.g., attentional control is a subordinate aspect of effortful control; Rothbart Ahadi, Hershey, & Fisher, 2001). Factor analysis has consistently

supported this three-factor structure of childhood temperament, albeit with substantial age-dependent overlap of certain factor loadings, as well as cultural differences in factor structure (Ahadi, Rothbart, & Ye, 1993; Putnam & Rothbart, 2006).

The use of multiple continuous scales provides a multifactorial description of temperament, capturing even slight individual differences. However, investigators using a dimensional approach often focus on a single factor or dimension, or the additive or interactive effects of a few dimensions, rather than considering the child's temperament as a whole (Magnusson, 2001; Janson & Mathieson, 2008). This narrow focus is a major limitation (Kagan et al., 1994), as different dimensions co-occur and jointly influence behavior and outcomes (e.g., low fear combined with high activity in infancy predict depression and externalizing problems in early childhood; Colder, Mott, & Berman, 2002). Moreover, the dimensional approach yields a single estimate of genetic and environmental influences, intended to represent the whole population, whereas a categorical approach allows heritability to differ across subgroups within the population.

Categorical Models of Temperament

The categorical approach allows for heterogeneous profiles with differing distributions on observed and latent constructs (Magnusson, 2001), and addresses the possibility that certain combinations of traits are meaningful. For instance, in stressful situations, a child with high activity level combined with anger may exhibit conduct and social problems, whereas a highly active but less angry child may be sociable and involved in school activities.

However, most statistical methods used to identify subgroups (e.g., cluster analysis, latent profile analysis) assume that the underlying dimensional structure is homogeneous across all individuals in a profile, with no within-profile variation, or, stated another way, that the items are not allowed to correlate within each profile (i.e., conditional independence assumption; Muthén et al., 2006). This assumption of homogeneity does not seem to be reflected in temperament data, and treating individuals within a profile who show true differences as the same or reducing dimensional data to categorical latent variables results in a loss of information and a decrease in statistical power. In addition, some children may not fit into larger categories. For example, Thomas, Chess, and Birch (1970) found that only 65 percent of children were categorized as easy, difficult, or slow-to-warm up. Kagan and colleagues' (Kagan et al., 1994) have proposed that many temperament types may exist in the population, but the less frequent types may be difficult to identify with small sample sizes. Finally, most categorical research has defined groups using conceptual rather than statistical criteria.

Nevertheless, some researchers have used data-driven statistical methods, such as inverse factor analysis, cluster analysis, and latent profile analysis, to identify temperament typologies in childhood and adolescence (e.g., Caspi & Silva, 1995; Janson & Mathiesen, 2008; Van Den Akker, Dekovi, Prinzie, & Asscher, 2010). These studies have yielded between two and five temperament types. In one study, Aksan et al. (1999) uncovered two temperament types (controlled-nonexpressive and non-controlled-expressive) using configural frequency analysis on composite scores of parent-reported negative affectivity, approach, and effortful control with 42-54 month children. In a longitudinal study following

children from 18 months to 8-9 years, Jansen and Mathiesen (2008) uncovered five profiles (labeled undercontrolled, confident, unremarkable, inhibited, and uneasy) with parent-reported emotionality, shyness, sociability, and activity measures. These profiles are consistent with the five profiles found by Caspi and Silva (1995) in a sample of 3 year-old children using observed measures of lack of control, approach, and sluggishness. Taken together, there is a lack of convergence in the type and number of profiles being identified across multiple studies, potentially driven by differences in statistical methods, indicators included in analyses (e.g., most studies used measures of reactivity and not self-regulation), and the age ranges studied.

Twin Factor Mixture Modeling: A Hybrid Method

TFMM, an advanced statistical technique that combines dimensional and person-centered approaches, provides an empirically-driven method that utilizes both categorical and dimensional modeling techniques (Muthén et al., 2006). More specifically, TFMM allows for not only the modeling of heterogeneous groups (i.e., profiles) within the population but also the specification of a latent dimensional model within each profile. This method has several advantages over traditional factor analysis and latent profile modeling. First, the specification of a dimensional model accounts for measurement error of parent-reported temperament not accounted for in traditional profile analysis (Lubke & Muthén, 2005), and as Muthén et al. (2006) suggests, “substantively more meaningful clusters might be found when allowing within-class correlations among items” (p. 317). Second, TFMM allows for the free estimation of factor loadings within each profile, which may show that the “factor [or factors] in the model might not be measured the same way across classes” (Clark et al., 2013; p. 689). Third, by utilizing the twin design, unique genetic and environmental (shared and non-shared) explanations of temperament profile membership can be estimated. In summary, TFMM allows for relaxing the assumption of conditional independence and homogeneity held in categorical and dimensional approaches, which may more accurately represent the data and provide more flexible estimations of the genetic and environmental influences.

Heritability of Temperament in Middle Childhood

Heritability is the proportion of variance in a trait that can be accounted for by genetic influences within a given sample at a given time (Plomin, DeFries, Knopik, & Neiderhiser, 2013). Twin studies find genetic influences on most dimensions of temperament in childhood, with estimates usually falling between .20 and .60 (Saudino, 2005), but genetic and environmental influences differ depending on the dimension of temperament, level of analysis (i.e., broad factors vs. subordinate dimensions), measurement (e.g., parent report vs. behavioral observation), and sample characteristics (e.g., increases in heritability with age; Saudino, 2005; Saudino & Micalizzi, 2015). We focus on parent-report of the higher-order factors negative emotionality, effortful control, and surgency, but acknowledge the importance of context and measurement, as well as the etiologically complex and hierarchical structure of temperament.

Most twin studies of temperament use infants and young children, with only a few examining the heritability of temperament in middle childhood. In both early and middle childhood, mother report of effortful control and surgency tends to be moderately to highly heritable, with little evidence of shared environmental influences. This finding is consistent at the subordinate level for approach, impulsivity and activity level (Deater-Deckard et al., 2010; Goldsmith et al., 1997; Lemery-Chalfant, Kao, Swann, & Goldsmith, 2013; Lemery-Chalfant, Doelger, & Goldsmith, 2008; Mullineaux, Deater-Deckard, Petrill, Thompson, & DeThorne, 2009; Saudino & Micalizzi, 2015). However, subordinate facets of temperament involving positive emotion often show shared environmental variance and low or nonsignificant heritability (e.g., Goldsmith et al., 1997; Mullineaux et al., 2009). In addition, facets of effortful control (inhibitory control) are consistently moderately to highly heritable, but show modest but significant shared environmental influences (Gagne & Saudino, 2010). Mother report of negative emotionality also shows mixed evidence of environmental influences at both the factor level and the subordinate level (Saudino & Micalizzi, 2015), with some but not all (e.g., Deater-Deckard et al., 2010) studies finding modest shared environmental influences on anger and social fear in early childhood (Emde, Robinson, Corley, Nikkari, & Zahn-Waxler, 2001; Goldsmith et al., 1997). Finally, one study found that a single set of common genetic influences accounted for the covariance among facets of negative emotionality, whereas shared environmental influences explained variance in mother-report of anger and sadness but not fear (Clifford, Lemery-Chalfant, & Goldsmith, 2015).

Only two genetically-informed studies considered heritability of temperament profiles across multiple dimensions. In the first, Matheny and Dolan (1980) found significantly higher monozygotic (MZ) than dizygotic (DZ) twin concordance for a descriptive temperament type representing six temperament dimensions (Compliant Morality, Applied Cognitive, Sociability, Tough-Mindedness, Emotionality, and Activity-Distractibility) in 7-10 year old twins. However, this study was limited by the small sample and descriptive analysis. In the second, Beekman et al. (2015) used latent profile analysis to derive temperament profiles in a larger longitudinal sample of adopted children at 9, 18, and 27 months. At each age, a four-profile solution provided the best fit, and two of the four profiles (negatively reactive and positively reactive) were evident at all ages and showed considerable stability in membership, with a stable fearful profile and a less stable active-reactive profile emerging at 18 months. Birth mother harm avoidance predicted higher likelihood of belonging to the fearful than the negatively reactive profile at 18 months, providing initial evidence of genetic influences on the fearful profile in toddlers. Adoptive mother and adoptive father harm avoidance, indexing shared environmental influences, predicted higher likelihood of belonging to the negatively reactive profile (relative to the positively reactive profile) at all three ages. These findings provide initial evidence for specificity in heritability and shared environmental influences on temperament profiles.

Present Study

The study had two aims. The first aim was to identify distinct latent temperament profiles in a large, representative twin sample of primarily seven to eight year old children. We used a TFFM approach (Muthén et al. 2006) that allowed us to specify both a categorical model of

temperament for the whole sample, as well as a dimensional model of temperament for each profile. Using a TFM model that maximizes the strengths of dimensional and categorical approaches allowed us to not only identify temperament profiles, but also account for measurement error and understand the profile specific relations between the observed temperament indicators and an unobserved latent factor (or factors). Moreover, past studies examining temperament profiles have relied on a few broad indicators of reactivity without considering regulation. The temperament measure used in this study (i.e., Children's Behavior Questionnaire; Rothbart et al., 2001) is a more comprehensive measure of temperament (including self-regulation), and assesses specific behaviors within a particular time frame, which may minimize rater biases such as twin contrast effects (Hwang & Rothbart, 2003). Thus, we address the statistical and measurement limitations of past studies, using an instrument that may provide a more fine-grained identification of temperament profiles.

Although the number and type of profiles has varied considerably across past studies, we predicted the identification of the following four profiles: 1) Dysregulated, Anger Reactive: moderate to high anger, low fear and shyness, high levels of surgency (e.g., impulsivity, positive approach, activity), low effortful control, 2) Dysregulated, Fear Reactive (Behaviorally Inhibited): high fear and shyness, low anger, low surgency, and low to moderate effortful control, 3) Well-Regulated, Positive Reactive: low levels of negative reactivity (anger, sadness, fear), moderate surgency, high positive reactivity, and high Effortful Control, and 4) Regulated, Typical Reactive: low to moderate negative reactivity, surgency, and effortful control. Moreover, we expected to find support for Rothbart's (see Rothbart & Bates, 2006) three-factor model of childhood temperament with the latent factors of negative reactivity, surgency, and effortful control.

The second aim was to estimate genetic and environmental influences on both temperament profiles and dimensional factors. We expected profiles characterized by high self-regulation and high negative emotionality (particularly fearfulness) to be highly heritable. Our predictions for profiles involving surgency or approach depend on profile composition, such that we expected a profile characterized primarily by high levels of the disinhibited aspects of surgency (e.g., impulsivity, activity level) to be highly heritable with no role for the shared environment, but a profile primarily characterized by positive emotion and approach to be largely influenced by the shared environment.

Method

Participants

The sample consisted of 787 twin pairs (291 MZ, 237 same-sex DZ, 239 opposite-sex DZ, and 20 missing zygosity) participating in the longitudinal Wisconsin Twin Project focused on risk for child psychopathology (WTP; Lemery-Chalfant, Goldsmith, Schmidt, Arneson, & Van Hulle, 2006). Mothers ($n = 774$) and fathers ($n = 632$) completed measures of child temperament during middle childhood ($M = 7.4$ years; $SD = .84$; 49% female). Twin pairs were representative of Wisconsin, with 88.3% Caucasian, 5.1% African American, 4.3% Other/Mixed, 1.4% Native American, and .8% Asian. 3.4% of families earned \$20,000 or less, 6.5% earning \$20,001-50,000, 34.7% earning \$50,001-100,000, and 55.4% earned over

\$100,000. For fathers' education (mothers' in parentheses), .6% (.1%) had only a grade school education, 31.3% (17.2%) graduated from high school or had some high school, 31.2% (38.3%) had some college education, 23.2% (28.1%) were college graduates, and 13.7% (16.3%) held a graduate degree or had some graduate-level courses.

Measures

Zygosity Questionnaire for Young Twins (Goldsmith, 1991)—Mothers completed the *Zygosity Questionnaire for Young Twins* (Goldsmith, 1991), which measures physical similarities. The agreement of this questionnaire with genotyping is greater than 95% (Forget-Dubois et al., 2003; Price, Freeman, Craig, Petrill, Ebersole & Plomin, 2000). Observers also completed questions concerning zygosity after a home-visit. For 20 of the twin pairs, parents and observers did not agree and these pairs were omitted from genetic analyses.

Children's Behavior Questionnaire (CBQ; Rothbart et al., 2001)—The CBQ is a comprehensive parent-report assessment of temperament designed for three to eight year old children and measures temperament in both reactive and regulation domains. The CBQ uses focused questions that tap specific behaviors within a specific timeframe, thus minimizing rater bias and sibling contrast effects (Saudino, 2003). We obtained mother and father reports on a modified CBQ, consisting of 12 ten item scales: fear, shyness, sadness, anger, activity level, impulsivity, approach, smiling and laughter, low intensity pleasure, soothability, inhibitory control, and attentional focusing (mean α for mother report was .79 [$SD = .06$], for father report was .77 [$SD = .05$]). Items were rated on a 7-point scale ranging from "extremely untrue" to "extremely true" of your child over the past six months. Correlations ($p < .001$ for all bivariate correlations) between mother and father mean scores ranged from .42 (smiling and laughter) to .66 (impulsivity), and mother-father mean composites were formed. We next regressed mother-father mean composite scores on child age (linear), child age squared (quadratic), and sex (1 = girl, 2 = boy) and saved the unstandardized residuals to use as indicators in the analyses, as is standard practice to reduce potential biases when it is infeasible to incorporate additional covariates (McGue & Bouchard, 1984).

Procedure

Mothers completed the zygosity and demographic questionnaires over the telephone, and mothers and fathers independently completed the CBQ mailed to the home. In addition, families participated in a 4-hour home visit, which involved a number of assessments of the home environment and children's behavior that are not treated in these analyses.

Statistical Analyses

We used 'Type = Complex' in Mplus v7.11 to account for interdependence between cotwins for all analyses. We handled missing data using full information maximum likelihood (FIML) estimation (Enders & Bandalos, 2001).

Aim 1: Twin Factor Mixture Modeling—We followed the recommendations of Clark, Muthén, Kaprio, D'Onofrio, Viken, & Rose (2013) for fitting factor mixture models using

Mplus v7.11. The unstandardized residual scores of the 12 CBQ scales served as continuous indicators. The following is a step-by-step description of model testing:

1. We tested a series of latent profile models (i.e., 2-profile, 3-profile, 4-profile, etc.) and confirmatory factor models (i.e., 1-factor, 2-factor, 3-factor, etc.), independently. We set equality constraints across cotwins on measurement parameters (profile-specified means, variances/covariances) and structure parameters (profile proportion) for the latent profile models and on the factor structure and loadings for the confirmatory factor models. The best-fitting models served as upward bounds for the TFM models tested in step 2. That is, if the best-fitting latent profile model consisted of four profiles and confirmatory factor model was three factors, then the final TFM model tested was a four-profile, three-factor model. This method allowed us to compare more parsimonious models (i.e., latent profile or confirmatory factor model) and to determine whether a TFM model best fit the data.
2. We next tested a series of TFM models as illustrated in Figure 1 (see Muthén et al., 2006 for a more in-depth explanation). We set equality constraints across cotwins on the measurement parameters (profile-specific means, variances/covariances, factor loadings) and structure parameters (profile proportion). We also constrained factor variances equal to one and factor means to zero for each model to estimate standardized factor loadings and to make direct comparisons across profiles, as scale differences shift to factor loadings and covariances (Lubke & Neale, 2008). Note that we also tested the models using unconstrained factor variances. We allowed for the latent profile and factor variables to correlate across cotwins.

We compared all models using log-likelihood ratios, Bayesian Information Criteria (BIC; Schwarz, 1978) and the sample-size adjusted BIC (ABIC; Sclove, 1987). Mplus 7.11 does not provide Vuong-Lo-Mendell-Rubin likelihood ratio tests (LMR-LRT; Lo, Mendell, & Rubin, 2001) or parametric bootstrapped likelihood ratio tests (BLRT; McLachlan, 1987) for mixture models in wide data format. We specifically followed Muthén and colleagues' guidelines for model comparison (Clark et al., 2013; Lubke & Muthén, 2005; Muthén et al., 2006), evaluating the interpretability of the latent profiles and factors identified across empirically best-fitting models to select the model that best represented the data.

Aim 2: Twin Methodology—We report cotwin concordance for profile membership, as well as twin intraclass correlations for factor loadings, separately for MZ and DZ twins. We used the software OpenMx (Boker et al., 2011), which yields maximum likelihood estimates of scale means and parameters for covariances from the raw data matrix. We began with the full, saturated model (using a liability threshold model for profile concordance) and systematically equated estimates among zygosity groups to test mean and variance differences. The threshold model (Neale & Cardon, 1992) has been used widely in behavior genetics, and it assumes an underlying normal distribution of liability (representing a large number of independent genetic and environmental factors that create variation), with one or more threshold values that discriminate between the categories (or temperament profiles). Next, we tested univariate ACE or ADE models (C and D cannot be modeled

simultaneously) to estimate additive genetic influences (A), either nonadditive genetic (D) or shared environmental influences (C), and nonshared environmental influences (E). We dropped parameters to fit reduced models, using the likelihood-ratio chi-square test to compare the fit of these nested models to the full model. A nonsignificant difference in the chi square values between two models implies that the additional specification did not significantly reduce the fit; thus, the more restricted model is accepted as a more parsimonious model.

Results

Descriptive Statistics

We present phenotypic correlations, means, and standard deviations for the entire sample in Table 1. All correlations were in the expected directions.

Twin Mixture Modeling

Model Fit—Model fit indices (log-likelihood, BIC, and ABIC), entropy, and estimated number of parameters for the latent profile models, confirmatory factor models, and TFMMs are presented in Table 2. Initial examination of the model fit indices indicated that the 4-profile, 3-factor TFMM best fit the data ($-2LL = 16053.37$, $BIC = 33527.07$, and $ABIC = 32850.68$). However, we chose to exclude this model from subsequent analyses as it produced a non-positive definite matrix—due to the high intercorrelations ($r > .67$) among latent factors within profiles and across twin pairs—and required a greater number of parameters as compared to less complex models. The next best-fitting models were the 5-profile, 1-factor FMM ($-2LL = -16433.49$, $BIC = 34020.58$, $ABIC = 33471.22$), 4-profile, 2-factor FMM ($-2LL = -16427.54$, $BIC = 33982.01$, $ABIC = 33445.35$), and 4-profile, 1-factor FMM ($-2LL = -16610.75$, $BIC = 34148.38$, $ABIC = 33706.98$). Examination of profile plots revealed that the main difference between the 5-profile, 1-factor FMM and the 4-profile, 1-factor FMM was that a typical profile (i.e., mean levels across CBQ scales) in the latter was split into two profiles with minor severity differences (i.e., estimated CBQ mean scores slightly below and above the mean). The 4-profile, 2-factor FMM produced identical profiles as the 4-profile, 1-factor FMM, but at the cost of 30 more parameters. Moreover, examination of both the four profile, two factor and four profile, three factor models revealed that the factors were more highly correlated within most of the profiles (i.e., the two and three factors had correlations higher than .55 and sometimes higher than .90 within most of the profiles) than normally found using standard factor analysis procedures (Rothbart et al., 2001). In the end, we chose the 4-profile, 1-factor FMM for further analyses and interpretation. Note that factor variances ranged in magnitude from .21 to .37 when freely estimated in the 4-profile, 1-factor model, and did not appreciably impact the final model fit statistics ($-2LL = -16628.67$, $BIC = 34184.22$, $ABIC = 33742.82$, entropy = .78).

4-Profile, 1 Factor FMM: Profile Membership—Estimated means of unstandardized CBQ scales are plotted in Figure 2. The largest profile ($n = 533$), labeled the ‘Regulated, Typical Reactive’ profile, consisted of children with mean levels of temperamental reactivity, surgency, and regulation. The next largest profile ($n = 483$) was a ‘Well-Regulated, Positive Reactive’ group of children with below or at mean levels of negative reactivity and surgency

(low activity, low impulsivity, mean approach), but above the mean on positive reactivity and self-regulation. The third largest profile ($n = 300$) was a 'Regulated, Surgent' group who had below mean levels of negative reactivity, above mean levels of both positive reactivity and surgency (high activity, high impulsivity, high approach), and mean levels of self-regulation. The smallest profile ($n = 258$) was labeled "Dysregulated, Negative Reactive," with above mean levels of negative reactivity, mean levels of surgency, and very low self-regulation. Note that allowing the factor variance to vary across profiles resulted in profiles that were remarkably similar to the four profiles just described. The only noticeable differences were that the profile that most closely resembled the 'Well-Regulated, Positive Reactive' type had closer to mean levels of attentional focusing and inhibitory control and that the profile that most closely resembled the 'Regulated, Surgent' profile had slightly higher than mean levels of attentional focusing and inhibitory control.

4 Profile, 1 Factor FMM: Profile-Specific Factor Loadings—We present standardized factor loadings for all profiles in Table 3. Results indicated that the 'temperament' latent factor explained significant variance for most observed indicators and that higher scores for the 'Well-Regulated, Positive Reactive', 'Regulated, Surgent', and 'Dysregulated, Negative Reactive' profiles reflected greater negative reactivity, higher surgency, and less effortful control (the factor score meaning for the 'Regulated, Typical Reactive' was reversed), but the pattern of factor loadings differed by profile. Specifically, the factor loadings for Low Intensity Pleasure, Smiling and Laughter, and Shyness were not significant for the 'Well-Regulated, Positive Reactive' or 'Regulated, Surgent' profiles, loadings for Fear and Sadness were not significant for the 'Regulated, Typical Reactive' or 'Dysregulated, Negative Reactive' profiles, and the Low Intensity Pleasure loading was not significant for the 'Dysregulated, Negative Reactive' profile.

Twin Concordance and Intraclass Correlations—Twin concordance for profile membership was high overall, and higher for MZ twins than DZ twins (Table 4), indicating heritability. The pattern of twin correlations for the temperament factor was consistent with an additive genetic model, with MZ correlations twice as high as DZ correlations (Table 4).

Twin ACE Liability Threshold Model Fitting with Profile Memberships—Saturated models indicated that thresholds could be equated across twin and zygosity groups for the 'Regulated, Typical', 'Well-Regulated, Positive Reactive', and 'Regulated, Surgent' profiles ($\chi^2(3) = 2.69, p = .44$; $\chi^2(3) = 6.85, p = .08$; $\chi^2(3) = 5.94, p = .11$, respectively), and could be equated across twins for the 'Dysregulated, Negative Reactive' profile ($\chi^2(2) = 1.88, p = .39$), but not zygosity ($\chi^2(3) = 13.39, p < .001$). Specifically, the threshold was estimated to be lower for DZ (.86) than MZ (1.21) twins, likely due to the small MZ (13% of MZ twins) compared to DZ (28% of DZ twins) sample size in the 'Dysregulated, Negative Reactive' profile. Next, ACE models were fit to estimate genetic and environmental effects on profile membership, beginning with the full model and dropping A and C in turn to test their significance. Parameter estimates and model fit are given in Table 5. Heritability was substantial for all profiles, ranging from .28 for the 'Regulated, Typical Reactive' profile to .91 for the 'Well-Regulated, Positive Reactive' profile. The ACE model fit best for the 'Regulated, Typical Reactive' and 'Well-Regulated,

Positive Reactive' profiles, with C estimated at .71 and .44, respectively. Estimates of nonshared environmental influences across profiles were small.

Twin ACE Model Fitting with the Factor Score—Univariate model fitting was applied to the temperament factor score raw data matrix. A saturated model indicated that means and variances could be equated across twin and zygosity, $\chi^2(6) = 5.43, p = .49$. Next, an ACE model was fit, and the significance of A and C was tested. As anticipated, the AE model fit best, with heritability estimated at .80; parameter estimates and model fit are given in Table 5.

Discussion

Our primary aim was to use a novel mixture modeling approach to identify temperament profiles in a large sample of school-aged twins. We add to the literature on temperament typologies (Aksan et al., 1999; Jansen & Mathiesen, 2008; Van Den Akker et al., 2010) by considering both reactivity and regulation, and by using a method that eliminates strong assumptions of other person-centered methods (e.g., no within-profile variance) and examines temperament typologies from a genetically-informed perspective. We uncovered four distinct profiles with one underlying latent factor, and this model provided a better fit than purely person-centered or variable-centered methods (i.e., latent profile analysis and factor analysis). Moreover, the pattern of significant factor loadings differed for each profile, such that variation on some CBQ scales was not explained by temperament type, supporting an integrated qualitative and quantitative approach. We also found differences in heritability across profiles, with both 'Surgent' and 'Well-Regulated' profiles highly heritable with no shared environmental effects, the 'Typical' profile largely influenced by the shared environment, and the 'Dysregulated' profile showing substantial genetic and shared environmental influences.

Temperament Profiles in Middle Childhood: Qualitative and Quantitative Features

In this initial application of TFMM to the study of childhood temperament, our findings indicated that modeling both the qualitative and quantitative features of temperament provided a more fine-grained classification than either alone. That is, the TFM model fit considerably better than a 5-profile LPM, with one less profile needed to explain variation in temperament when the underlying factor structure was also modeled, and also better than the 3-factor dimensional model. Two of our profiles closely map onto the hypothesized profile temperament typologies reported in past studies. For instance, similar to Jansen and Mathiesen's (2008) 'Unremarkable' profile, we found a 'Typical' profile of children (34%) with moderate positive and negative emotional reactivity, approach, shyness, and self-regulation. We also found a 'Surgent' profile (19%) similar to Jansen and Mathiesen's 'Confident' profile, defined in our sample by very low shyness, high activity, impulsivity, and positive emotionality, low-moderate negative emotionality, and moderate self-regulation. Finally, like Jansen and Mathiesen's 'Uneasy' profile, our 'Dysregulated' profile (16%) included children who were negatively reactive, high in shyness and anger, and poorly regulated relative to peers. Unlike past research (e.g., Jansen & Mathiesen, 2008; Van Den Akker et al., 2010), we did not find our hypothesized 'Fear Reactive, Dysregulated' (i.e.,

Inhibited) or 'Anger Reactive, Dysregulated' (i.e., Uninhibited) profiles based on discrete emotions of fear and anger; instead, our 'Dysregulated' profile was one of high negative emotionality and similar to Caspi and Silva's (1995) 'Undercontrolled' type, with moderate activity, impulsivity, and approach. Finally, many children (28%) belonged to a hypothesized 'Well-Regulated' profile that does not map onto any type found in past studies; this profile was characterized by consistently high self-regulation, high positive reactivity, and low negative reactivity, along with moderate approach and shyness. The emergence of this profile illustrates the importance of considering self-regulation when using person-centered or mixture method approaches. At the same time, the relative consistency of the other three profiles across different measures, methodological procedures (e.g., parent-report versus observations), and samples supports their generalizability even when measures of regulation are not considered.

Our findings also demonstrated that profile membership alone could not capture all variability across temperament dimensions, and thus a latent factor was needed. Specifically, children in the 'Well-Regulated,' and 'Surgent' profiles showed consistently low shyness and high positive reactivity (smiling and laughter, low intensity pleasure), but varied on all dimensions of negative reactivity (fear, anger, sadness), surgency (activity level, approach, impulsivity) and effortful control (soothability, inhibitory control, attentional focusing). In contrast, fear and sadness did not vary within the 'Typical' or 'Dysregulated' profiles (e.g., 'Dysregulated' children had consistently high fear and sadness, and low low intensity pleasure). However, children in the 'Typical' and 'Dysregulated' did vary in surgency, effortful control, and (for the 'Typical' profile) smiling and laughter. The finding that children who are qualitatively similar in their overall temperament can vary considerably across specific dimensions may help explain why developmental trajectories for seemingly similar children diverge over time (Jansen & Mathiesen, 2008). When we estimated rather than fixed factor variances across profiles, our findings were not substantially altered, and entropy was lower (.77 vs. .80), which suggests that the profiles were closer together and less distinguishable.

Differential Heritability

Using the twin design, we observed that the genetic and environmental etiology of profile membership varied, which may have important implications for the study of temperament and children's risk for mood and behavior problems. Specifically, all profiles were heritable, but both the 'Well-Regulated' and 'Surgent' profiles were highly heritable with no shared environmental influences. The 'Dysregulated' profile showed moderate genetic and shared environmental effects, and membership in the 'Typical' profile was mostly influenced by shared environmental effects. Dimensional measures of temperament are often moderately to highly heritable, with no shared environmental effects (Saudino, 2005), but our findings suggest that shared environmental effects on patterns of traits may not be evident at the level of single dimensions. Findings point to the need for a holistic view, in that the whole of a child's temperament and behavior cannot be reduced to the additive effects of individual traits.

Moreover, our findings are consistent with studies of moderated heritability suggesting that the environment may facilitate or constrain the expression of genetic influences on emotion and behavior, such that, e.g., effortful control was more heritable in more chaotic homes, and negative emotionality was less heritable in safer, more structured homes (Lemery-Chalfant et al., 2013). Shanahan and Hofer (2005) describe multiple ways such gene-environment interplay can operate, but two of these ways are especially relevant to environmental influences on our 'Typical' group. First, an enriched environment may compensate for genetic risk, such that, e.g., environments characterized by warm and sensitive parenting, a physically and emotionally safe and secure climate, and effective teaching and modeling of self-regulation, may channel individuals away from dysregulated extremes toward a more moderate range (Sroufe, 1997). Secondly, social pressures may constrain the expression of genetically-influenced individual differences that may be evident in more permissive environments. For example, Kendler, Thornton, and Pedersen (2000) found cohort effects on the heritability of smoking in women, such that heritability increased across generations as social restrictions against women smoking lessened.

In the case of temperament, children are socialized to regulate their emotion and behavior in accordance with sociocultural norms and values. The earliest source of socialization is parenting, including responsiveness and mutuality, discipline, and the modeling and teaching of emotion-related knowledge and regulation (Eisenberg, Cumberland, & Spinrad, 1998; Feldman, Greenbaum, & Yirmiya, 1999). However, peers and teachers also play a role, and by middle childhood, children spend considerable time in school and peer settings where they are expected to focus their attention and inhibit both extreme positive affect or approach and extreme negative affect or withdrawal (Eisenberg, Valiente, & Eggum, 2010; Shiner et al., 1998). Importantly, such social constraints would not be expected to eliminate individual differences in temperament as normal-range variation, merely promote moderate rather than extreme levels. Our models were not informative about which aspects of the shared environment influence membership in our 'Typical' group, but one possibility is that profile membership reflects the combined effects of socialization, parenting, peer relations, and the home environment, while genetic influences on temperament exist at the within profile dimensional level. It is notable that the profile with the strongest genetic influences ('Well-Regulated') reflects traits that are adaptive and socially valued (self-regulation, positive emotion), which in a positive environment will be allowed or encouraged. The same may be true of some traits characteristic of the highly heritable 'Surgent' profile (e.g., positive emotion, approach). In contrast, impulsivity, approach, and activity are not always valued or adaptive, and put children at risk for externalizing problems (Eisenberg et al., 2010). However, children belonging to the 'Surgent' profile tended toward moderate regulation, and showed within-profile variation on impulsivity, activity, and approach, suggesting that this profile does not consist solely of disordered extremes.

In contrast to the 'Typical' profile, shared environmental influences on 'Dysregulated' profile membership likely reflect exposure to negative environments strong enough to contribute to individual differences over and above genetic variation. For instance, Jaffee, Caspi, Moffitt, Polo-Tomas, Price, and Taylor (2004) found that the effect of maltreatment on aggression was fully environmental. However, shared environmental risks other than maltreatment (e.g., life stress, family conflict, or poverty) may contribute to the negative

emotional reactivity and low self-regulation seen in our ‘Dysregulated’ profile. Intervention and prevention efforts would benefit from elucidating specific environmental risk factors for membership in the ‘Dysregulated’ profile.

Limitations and Future Directions

Several limitations and future directions are noteworthy. First, we must consider the possibility that our models may have identified profiles that truly do not exist in the population due to misspecification of the models and violations of certain assumptions (Bauer & Curran, 2004). For example, it is possible that the within-profile dimensional model was not specified correctly (i.e., latent factor structure actually differs for each profile) and resulted in the extraction of spurious latent profiles. Moreover, the model may have better captured a multivariate nonnormal distribution of the data that improved model fit and thus arbitrarily supported more latent profiles than exist. However, our results demonstrated that three of our four profiles closely align with our hypothesized profiles and profiles identified in past research (Jansen & Mathiesen, 2008; Caspi & Silva, 1995), thus providing confidence in our findings. Nevertheless, replication of our findings in other twin and non-twin samples would help bolster the robustness of our four profile solution. Second and a related point, a larger sample may have allowed a more sensitive test of the TFM models, and perhaps the identification of rarer temperament types, such as Kagan et al.’s (1994) behavioral inhibition. Third, a limitation of twin studies is that findings may not generalize to non-twin siblings and singletons. Most research supports the generalizability of twin data to non-twins (e.g., Angold, Erkanli, Silberg, Eaves, & Costello, 2002), but the use of TFFM to derive temperament profiles is novel and requires replication. Fourth, our sample was largely Caucasian, middle class, and from the upper Midwest, limiting generalizability. The cross-sectional design precludes an examination of stability and change in profile membership over time, and the narrow age range of our sample (primarily 7-8 years) limits generalizability of our findings in other developmental periods. Given the shared environmental influences on our high-risk ‘Dysregulated’ profile, future research also should examine which specific environmental factors predict membership in this profile, and whether these factors may be amenable to intervention.

Acknowledgments

Research reported in this publication was supported by the National Institute of Mental Health under Award Numbers R01MH59785, R01MH101504, R01MH10150, R01HD079520 P30HD003352, P50MH100031, and T32MH018387. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Ahadi SA, Rothbart MK, Ye RM. Children’s temperament in the US and China: Similarities and differences. *European Journal of Personality*. 1993; 7:359–378. doi:10.1002/per.2410070506.
- Aksan N, Goldsmith HH, Smider NA, Essex MJ, Clark R, Hyde JS, Klein MH, Vandell DL. Derivation and prediction of temperamental types among preschoolers. *Developmental Psychology*. 1999; 35:958–971. doi:10.1037/0012-1649.35.4.958. [PubMed: 10442865]
- Angold A, Erkanli A, Silberg J, Eaves L, Costello EJ. Depression scale scores in 8-17-year-olds: Effects of age and gender. *Journal of Child Psychology and Psychiatry*. 2002; 43:1052–1063. doi: 10.1111/1469-7610.00232. [PubMed: 12455926]

- Bauer DJ, Curran PJ. The integration of continuous and discrete latent variable models: potential problems and promising opportunities. *Psychological Methods*. 2004; 9:3–29. [PubMed: 15053717]
- Beekman C, Neiderhiser JM, Buss KA, Loken E, Moore GA, Leve LD, Ganiban JM, Shaw DS, Reiss D. The development of early profiles of child temperament: Characterization, continuity, and etiology. *Child Development*. 2015 doi: 10.1111/cdev.12417.
- Biro FM, Galvez MP, Greenspan LC, Succop PA, Vangeepuram N, Pinney SM, Wolff MS. Pubertal assessment method and baseline characteristics in a mixed longitudinal study of girls. *Pediatrics*. 2010; 126(3):e583–e590. doi: 10.1542/peds.2009-3079. [PubMed: 20696727]
- Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T, Fox J. OpenMx: An open source extended structural equation modeling framework. *Psychometrika*. 2011; 76:306–317. doi:10.1007/s11336-010-9200-6. [PubMed: 23258944]
- Caspi A, Silva PA. Temperamental qualities at age three predict personality traits in young adulthood: Longitudinal evidence from a birth cohort. *Child Development*. 1995; 66:486–498. doi: 10.1111/j.1467-8624.1995.tb00885.x. [PubMed: 7750379]
- Clark SL, Muthén B, Kaprio J, D’Onofrio BM, Viken R, Rose RJ. Models and strategies for factor mixture analysis: An example concerning the structure underlying psychological disorders. *Structural Equation Modeling: A Multidisciplinary Journal*. 2013; 20:681–703. doi: 10.1080/10705511.2013.824786.
- Clifford S, Lemery-Chalfant K, Goldsmith HH. The unique and shared genetic and environmental contributions to fear, anger, and sadness in childhood. *Child Development*. 2015; 86(5):1538–1556. doi: 10.1111/cdev.12394. [PubMed: 26182850]
- Colder CR, Mott JA, Berman AS. The interactive effects of infant activity level and fear on growth trajectories of early childhood behavior problems. *Development and Psychopathology*. 2002; 14:1–23. doi:10.1017/S0954579402001013. [PubMed: 11893087]
- Deater-Deckard K, Beekman C, Wang Z, Kim J, Petrill SA, Thompson LA, DeThorne LS. Approach/positive anticipation, frustration/anger, and overt aggression in childhood. *Journal of Personality*. 2010; 78:991–1010. doi: 10.1111/j.1467-6494.2010.00640.x. [PubMed: 20573133]
- Eisenberg N, Cumberland A, Spinrad T. Parental socialization of emotion. *Psychological Inquiry*. 1998; 9(4):241–273. doi: 10.1207/s15327965pli0904_1. [PubMed: 16865170]
- Eisenberg N, Valiente C, Eggum ND. Self-regulation and school readiness. *Early Education and Development*. 2010; 25(5):681–698. doi: 10.1080/10409289.2010.497451.
- Emde, RN., Robinson, JL., Corley, RP., Nikkari, D., Zahn-Waxler, C. Reactions to restraint and anger-related expressions during the second year. In: Emde, RN., Hewitt, JK., editors. *Infancy to early childhood: Genetic and environmental influences on developmental change*. Oxford University Press; New York, NY: 2001. p. 127-140.
- Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*. 2001; 8:430–457. doi:10.1207/S15328007SEM0803_5.
- Feldman R, Greenbaum CW, Yirmiya N. Mother-infant affect synchrony as an antecedent of the emergence of self-control. *Developmental Psychology*. 1999; 35:223–231. doi: 10.1037/0012-1649.35.1.223. [PubMed: 9923477]
- Forget-Dubois N, Pérusse D, Turecki G, Girard A, Billette JM, Rouleau G, Boivin M, Malo J, Tremblay RE. Diagnosing zygosity in infant twins: Physical similarity, genotyping, and chorionicity. *Twin Research*. 2003; 6:479–485. doi:10.1375/136905203322686464. [PubMed: 14965457]
- Gagne JR, Saudino KJ. Wait for it! A twin study of inhibitory control in early childhood. *Behavior Genetics*. 2010; 40:327–337. doi: 10.1007/s10519-009-9316-6. [PubMed: 19936910]
- Goldsmith HH. A zygosity questionnaire for young twins: A research note. *Behavior Genetics*. 1991; 21:257–269. doi:10.1007/BF01065819. [PubMed: 1863259]
- Goldsmith HH, Buss KA, Lemery KS. Toddler and childhood temperament: Expanded content, stronger genetic evidence, new evidence for the importance of environment. *Developmental Psychology*. 1997; 33:891–905. doi:10.1037/0012-1649.33.6.891. [PubMed: 9383612]

- Hwang J, Rothbart MK. Behavior genetics studies of infant temperament: Findings vary across parent-report instruments. *Infant Behavior and Development*. 2003; 26(1):112–114. doi: 10.1016/S0163-6383(02)00173-X.
- Jaffee SR, Caspi A, Moffitt TE, Polo-Tomas M, Price T, Taylor A. The limits of child effects: Evidence for genetically mediated child effects on corporal punishment, but not on maltreatment. *Developmental Psychology*. 2004; 40:1047–1058. doi:10.1037/0012-1649.40.6.1047. [PubMed: 15535755]
- Janson H, Mathiesen KS. Temperament profiles from infancy to middle childhood: Development and associations with behavior problems. *Developmental Psychology*. 2008; 44:1314–1328. doi: 10.1037/a0012713. [PubMed: 18793065]
- Kagan, J., Snidman, N., Arcus, D., Reznick, JS. Galen's prophecy: Temperament in human nature. Basic Books; New York, NY: 1994. doi:10.5860/choice.32-2419
- Kendler KS, Thornton LM, Pedersen NL. Tobacco consumption in Swedish twins reared apart and reared together. *Archives of General Psychiatry*. 2000; 57(9):886–892. doi: 10.1001/archpsyc.57.9.886. [PubMed: 10986552]
- Larson RW, Richards MH, Moneta G, Holmbeck G, Duckett E. Changes in adolescents' daily interactions with their families from ages 10 to 18: Disengagement and transformation. *Developmental Psychology*. 1996; 32(4):744–754. doi: 10.1037/0012-1649.32.4.744.
- Lemery-Chalfant K, Doelger L, Goldsmith HH. Genetic relations between effortful and attentional control and symptoms of psychopathology in middle childhood. *Infant and Child Development*. 2008; 17:365–385. doi:10.1002/icd.581. [PubMed: 27076792]
- Lemery-Chalfant K, Goldsmith HH, Schmidt NL, Arneson CL, Van Hulle CA. Wisconsin Twin Project: Current directions and findings. *Twin Research and Human Genetics*. 2006; 9:1030–1037. doi:10.1375/183242706779462363. [PubMed: 17254447]
- Lemery-Chalfant K, Kao K, Swann G, Goldsmith HH. Childhood temperament: passive gene-environment correlation, gene-environment interaction, and the hidden importance of the family environment. *Development and Psychopathology*. 2013; 25:51–63. doi:10.1017/S0954579412000892. [PubMed: 23398752]
- Lo Y, Mendell N, Rubin D. Testing the number of components in a normal mixture. *Biometrika*. 2001; 88:767–778. doi:10.1093/biomet/88.3.767.
- Lubke GH, Muthén B. Investigating population heterogeneity with factor mixture models. *Psychological Methods*. 2005; 10:21–39. doi:10.1037/1082-989X.10.1.21. [PubMed: 15810867]
- Lubke G, Neale M. Distinguishing between latent classes and continuous factors with categorical outcomes: Class invariance of parameters of factor mixture models. *Multivariate Behavioral Research*. 2008; 43:592–620. doi:10.1080/00273170802490673. [PubMed: 20165736]
- Magnusson D. The holistic-interactionistic paradigm: Some directions for empirical developmental research. *European Psychologist*. 2001; 6:153–162. doi:10.1027/1016-9040.6.3.153.
- Matheny AP, Dolan AB. A twin study of personality and temperament during middle childhood. *Journal of Research in Personality*. 1980; 14:224–234. doi:10.1016/0092-6566(80)90030-6.
- McGue M, Bouchard TJ. Adjustment of twin data for the effects of age and sex. *Behavior Genetics*. 1984; 14:325–343. doi:10.1007/BF01080045. [PubMed: 6542356]
- McLachlan GJ. On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture model. *Applied Statistics*. 1987; 36:318–324. <http://www.jstor.org/stable/2347790>.
- Mullineaux PY, Deater-Deckard K, Petrill SA, Thompson LA, DeThorne LS. Temperament in middle childhood: A behavioral genetic analysis of fathers' and mothers' reports. *Journal of Research in Personality*. 2009; 43:737–746. doi:10.1016/j.jrp.2009.04.008. [PubMed: 20216930]
- Muthén B, Asparouhov T, Rebollo I. Advances in behavioral genetics modeling using Mplus: Applications of factor mixture modeling to twin data. *Twin Research and Human Genetics*. 2006; 9:313–324. doi:10.1375/twin.9.3.313. [PubMed: 16790142]
- Neale, M., Cardon, L. Methodology for genetic studies of twins and families. Springer Science & Business Media; 1992.
- Price TS, Freeman B, Craig IW, Petrill SA, Ebersole L, Plomin R. Infant zygosity can be assigned by parental report questionnaire data. *Twin Research*. 2000; 3:129–133. [PubMed: 11035484]

- Plomin, R., DeFries, J.C., Knopik, V.S., Neiderhiser, J.M. Behavioral genetics. 6th. Worth; New York, NY: 2013.
- Putnam SP, Rothbart MK. Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment*. 2006; 87:102–112. doi:10.1207/s15327752jpa8701_09. [PubMed: 16856791]
- Romeo RD. Puberty: A period of both organizational and activational effects of steroid hormones on neurobehavioral development. *Journal of Neuroendocrinology*. 2003; 15(12):1185–1192. doi: 10.1111/j.1365-2826.2003.01106.x. [PubMed: 14636181]
- Rothbart MK, Ahadi SA, Hershey KL, Fisher P. Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development*. 2001; 72:1394–1408. doi: 10.1111/1467-8624.00355. [PubMed: 11699677]
- Rothbart, MK., Bates, J.E. Temperament in children's development. In: Eisenberg, N., Damon, W., Lerner, R.M., editors. *Social, emotional, and personality development*. 6th. Vol. 3. Wiley; Hoboken, NJ: 2006. p. 99-166. *Handbook of child psychology*
- Rueda MR, Fan J, McCandliss BD, Halparin JD, Gruber DB, Pappert Lercari L, Posner MI. Development of attentional networks in childhood. *Neuropsychologia*. 2004; 42:1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012.
- Saudino KJ. Behavioral genetics and child temperament. *Journal of Developmental and Behavioral Pediatrics*. 2005; 26:214–223. doi:10.1097/00004703-200506000-00010. [PubMed: 15956873]
- Saudino KJ, Micalizzi L. Emerging trends in behavioral genetic studies of child temperament. *Child Development Perspectives*. 2015; 9(3):144–148. doi: 10.1111/cdep.12123. [PubMed: 26417384]
- Schwarz G. Estimating the dimension of a model. *The Annals of Statistics*. 1978; 6:461–464. doi: 10.2307/2958889.
- Shanahan M, Hofer S. Social context in gene-environment interactions: Retrospect and prospect. *Journal of Gerontology*. 2005; 60B(Special Issue 1):65–76. doi: 10.1093/geronb/60.Special_Issue_1.
- Sclove L. Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*. 1987; 52:333–343. doi:10.1007/BF02294360.
- Shiner RL. How shall we speak of children's personalities in middle childhood? A preliminary taxonomy. *Psychological Bulletin*. 1998; 124:308–332. doi:10.1037/0033-2909.124.3.308. [PubMed: 9849111]
- Shiner RL, Buss KA, McClowry SG, Putnam SP, Saudino KJ, Zentner M. What is temperament now? Assessing progress in temperamental research on the twenty-fifth anniversary of Goldsmith et al. (1987). *Child Development Perspectives*. 2012; 6:436–444. doi:10.1111/j.1750-8606.2012.00254.x.
- Shiner R, Caspi A. Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of Child Psychology and Psychiatry*. 2003; 44(1):2–32. doi: 10.1111/1469-7610.00101. [PubMed: 12553411]
- Shiner, R.L., DeYoung, C.G. *The Oxford handbook of developmental psychology*, Vol. 2: Self and other. Oxford University Press; New York, NY: 2013. The structure of temperament and personality traits: A developmental perspective; p. 113-141.
- Sroufe LA. Psychopathology as an outcome of development. *Development and Psychopathology*. 1997; 9:251–268. doi: 10.1017/S0954579497002046. [PubMed: 9201444]
- Thomas A, Chess S, Birch HG. The origin of personality. *Scientific American*. 1970; 223(2):102–109. Retrieved from <http://search.proquest.com/docview/615738572?accountid=28148>. [PubMed: 5447967]
- Van Den Akker AL, Dekovi M, Prinzie P, Asscher JJ. Toddlers' temperament profiles: Stability and relations to negative and positive parenting. *Journal of Abnormal Child Psychology*. 2010; 38:485–495. doi:10.1007/s10802-009-9379-0. [PubMed: 20058064]

Figure 1.

Twin Factor Mixture Model for CBQ Temperament Scales. *Note:* Broken arrow lines indicate that the factor loadings and factor covariance was allowed to vary across profiles. Factor variance was set to equal one across profiles. Adapted from Muthén et al. (2006).

Figure 2.
Estimated Means for CBQ Scales across the Four Profiles

Table 1

Zero-order Correlations and Descriptive Statistics for Study Variables (n = 1574)

	1	2	3	4	5	6	7	8	9	10	11	12	13	Mean	SD
1. Age														7.40	.84
2. Sex	.00	-													
CBQ Scales															
3. Fear	-.09**	-.13**												3.86	.88
4. Shyness	.00	-.15**	.23**											3.59	1.00
5. Sadness	-.06*	-.11**	.41**	.13**										3.88	.68
6. Anger	-.06*	.14**	.24**	.07*	.59**									4.49	.86
7. Activity Level	-.10**	.26**	.02	-.23**	.07*	.35**								4.51	.78
8. Impulsivity	-.06*	.26**	-.09**	-.56**	.06*	.32**	.66**							4.39	.70
9. Approach	-.12**	.12**	.20**	-.16**	.36**	.47**	.50**	.52**						4.99	.59
10. Smiling/Laughter	-.08**	-.02	-.06*	-.39**	-.05	-.11**	.15**	.22**	.25**					5.65	.52
11. Low Int. Pleasure	-.08**	-.29**	.05*	-.01	.06*	-.22**	-.32**	-.20**	.00	.48**				5.28	.57
12. Soothability	.03	.00	-.29**	-.14**	-.43**	-.59**	-.27**	-.19**	-.33**	.27**	.28**			4.71	.70
13. Inhibitory Control	.08**	-.24**	-.14**	.09**	-.28**	-.57**	-.58**	-.58**	-.48**	.15**	.39**	.56**		4.73	.90
14. Att. Focusing	.03	-.17**	-.15**	.05*	-.21**	-.38**	-.46**	-.42**	-.29**	.15**	.42**	.40**	.72**	4.53	.75

Note.

Means and standard deviations on diagonal. Int. = Intensity; Att. = Attentional. Sex was coded (1 = girl, 2 = boy). Correlations were derived in Mplus v7.11 using 'Type=Complex' to account for twin interdependence.

** p .001,

* p .05;

Table 2

Model Fit Indices for Latent Profile, Confirmatory Factor, and TFM Models

Model	Profiles	Factors	# of Parameters	Entropy	Log-likelihood	BIC	ABIC
LPA							
	2	-	50	.84	-18906.62	38146.66	37987.88
	3	-	65	.83	-18364.14	37161.71	36955.30
	4	-	81	.82	-17976.03	36492.19	36234.97
	5	-	98	.83	-17619.16	35891.80	35580.60
CFA							
	-	1	49	-	-18155.50	36637.75	36482.45
	-	2	52	-	-17914.85	36176.45	36011.33
	-	3	61	-	-17218.79	34844.34	34650.63
TFMA-1							
	2	1	77	.77	-17326.80	35167.05	34922.54
	3	1	107	.74	-16878.86	34471.22	34131.44
	4	1	139	.80	-16610.75	34148.38	33706.98
	5	1	173	.80	-16433.49	34020.58	33471.22
TFMA-2							
	2	2	86	.72	-17108.67	34790.81	34517.71
	3	2	125	.78	-16670.34	34174.20	33777.26
	4	2	169	.81	-16427.54	33982.01	33445.35
	<i>5</i>	<i>2</i>	<i>No Convergence</i>				
TFMA-3							
	2	3	105	.77	-16609.42	33918.99	33585.67
	3	3	159	.82	-16292.39	33645.03	33140.13
	4	3	213	.79	-16053.37	33527.07	32850.68
	<i>5</i>	<i>3</i>	<i>No Convergence</i>	-	-	-	-

Note: LPA = Latent Profile Analysis; CFA = Confirmatory Factor Analysis; TFMA = Twin Factor Mixture Analysis; BIC = Bayesian Information Criterion; ABIC = Adjusted Bayesian Information Criterion. **Bold** = "Best-fitting" model for that category based on entropy, log-likelihood, BIC, and ABIC.

Table 3

Standardized Factor Loadings for 4-Profile, 1-Factor TFM Model

CBQ Scales	Regulated, Typical Reactive	Well-Regulated, Positive Reactive	Regulated, Surgent	Dysregulated, Negative Reactive
Fear	0.13	0.51 ***	0.56 ***	-0.10
Shyness	0.52 ***	0.07	0.13	-0.61 ***
Sadness	-0.05	0.75 ***	0.80 ***	0.05
Anger	-0.35 ***	0.80 ***	0.82 ***	0.56 ***
Activity Level	-0.73 ***	0.40 ***	0.48 ***	0.77 ***
Impulsivity	-0.82 ***	0.40 ***	0.55 ***	0.84 ***
Approach	-0.62 ***	0.70 ***	0.76 ***	0.62 ***
Smiling and Laughter	-0.31 *	0.06	0.14	0.53 ***
Low Intensity Pleasure	0.26 **	-0.06	0.21	-0.23
Soothability	0.22 *	-0.58 ***	-0.75 ***	-0.57 ***
Inhibitory Control	0.84 ***	-0.61 ***	-0.82 ***	-0.84 ***
Attentional Focusing	0.71 ***	-0.37 ***	-0.47 ***	-0.63 ***

Note:

Factor loadings were averaged across cotwins.

 $p < .001$,

**
 $p < .01$

*
 $p < .05$.

Table 4

Twin Similarity in Temperament Profile Membership and Factor Score

Twin Profile Concordance		MZ		DZ	
Regulated, Typical Reactive		Twin 2		Twin 2	
		0	1	0	1
Twin 1	0	195 (67%)	6 (2%)	0	265 (56%)
	1	7 (2%)	83 (29%)	1	41 (9%)
					132 (28%)
Well-Regulated, Positive Reactive		Twin 2		Twin 2	
		0	1	0	1
Twin 1	0	169 (58%)	21 (7%)	0	282 (59%)
	1	18 (6%)	83 (29%)	1	66 (14%)
					69 (15%)
Regulated, Surgent		Twin 2		Twin 2	
		0	1	0	1
Twin 1	0	206 (71%)	18 (6%)	0	338 (71%)
	1	19 (7%)	48 (16%)	1	59 (12%)
					22 (5%)
Dysregulated, Negative Reactive		Twin 2		Twin 2	
		0	1	0	1
Twin 1	0	252 (87%)	6 (2%)	0	343 (72%)
	1	7 (2%)	26 (9%)	1	34 (7%)
					53 (11%)
Twin Intraclass Correlations for Factor Score		.81		.40	

Note: MZ = monozygotic, DZ=dizygotic.

Table 5
ACE Parameter Estimates and Threshold Model Fit for Temperament Profiles and Factor Score

Profile	Model	-2LL	df	-2LL	df	p	AIC	A	C or D	E
Regulated, Typical Reactive	ACE	1508.02	1531	–	–	–	–1553.98	.28 (.17-.42)	.71 (.57-.81)	.01 (.003-.031)
	AE	1563.44	1532	55.42	1	< .001	–1500.56			
	CE	1535.00	1532	26.98	1	< .001	–1529.00			
	E	1967.28	1533	459.26	2	< .001	–1098.72			
Well-regulated, Positive Reactive	ACE	1678.32	1531	–	–	–	–1383.68	.70	.20	.10
	AE	1680.50	1532	2.18	1	.14	–1383.50	.91 (.84-.95)	–	.09 (.05-.16)
	CE	1704.37	1532	26.05	1	< .001	–1359.63			
	E	1887.16	1533	208.84	2	< .001	–1178.84			
Regulated, Surgent	ADE*	1380.18	1531	–	–	–	–1681.82	.20	.65	.14
	AE	1383.30	1532	3.11	1	.08	–1680.70	.84 (.73-.91)	–	.16 (.09-.27)
	E	1496.19	1533	116.01	2	< .001	–1569.81			
	ACE	1162.18	1531	–	–	–	–1899.82	.52 (.31-.77)	.44 (.21-.63)	.03 (.01-.09)
Dysregulated, Negative Reactive	AE	1174.02	1532	11.84	1	< .001	–1889.98			
	CE	1182.68	1532	20.49	1	< .001	–1881.32			
	E	1367.20	1533	205.01	2	< .001	–1698.80			
Temperament Factor Score	ACE	3512.38	1530	–	–	–	452.38	.78 (.62-.88)	.02 (.00-.16)	.20 (.18-.24)
	AE	3512.43	1531	.05	1	.82	450.43	.80 (.71-.88)	–	.20 (.18-.24)
	CE	3609.51	1531	97.13	1	< .001	547.51			
	E	3900.48	1532	388.10	2	< .001	838.48			

Note. –2LL = –2 log likelihood; df = degrees of freedom; = change; p = probability; AIC = Akaike's Information Criterion; A, C, D and E are standardized squared parameter estimates for additive genetic, common environment, dominant genetic, and nonshared environment factors, respectively. The most parsimonious model is indicated in bold, and standardized confidence intervals are reported in parentheses. The most parsimonious final model is indicated in bold.

*We began with the ADE model because DZs where less than half as concordant than MZs.