

Temperament and its Association with Autism Symptoms in a High-risk Population

Nancy Garon¹ · Lonnie Zwaigenbaum^{2,3} · Susan Bryson⁴ · Isabel M. Smith⁴ · Jessica Brian^{5,6} · Caroline Roncadin⁷ · Tracy Vaillancourt⁸ · Vickie Armstrong⁴ · Lori-Ann R. Sacrey² · Wendy Roberts^{6,9}

Published online: 28 August 2015
© Springer Science+Business Media New York 2015

Abstract Temperament was investigated in a group of high-risk infants ($N=383$; 45 % girls) who had an older sibling with autism spectrum disorder (ASD), and in community control infants ($N=162$; 46 % girls) with no family history of ASD (low-risk). The infants were assessed at age 12 months using the Infant Behavior Questionnaire, and at 24 months using the Toddler Behavior Assessment Questionnaire. At 36 months, an independent blind diagnostic assessment for ASD was conducted using the Autism Diagnostic Interview-Revised (ADI-R) and the Autism Diagnostic Observation Schedule (ADOS). The results indicate not only differences in temperament traits between the high- and low-risk groups, but also differences in the structure of higher-order temperament factors. The results support the importance of early reactive temperament in the development of Effortful Control in the high-risk sample. Furthermore, Effortful Control at 24 months appears to play

a critical role in predicting later ASD symptoms (at 36 months). Taken together, these findings support the use of early temperament as an endophenotype for ASD.

Keywords Autism spectrum disorder · Temperament · Endophenotype · Regulation

Given our broad knowledge base of early temperament, measurement of temperament in populations at high risk of developmental disorders provides an informative framework for increasing our understanding of these disorders (Nigg 2006). Temperament and personality traits are closely linked to basic psychobiological systems, which in turn are governed by specific genes (Whittle et al. 2006). Thus, temperament may serve as an early endophenotype; that is, physiological or psychological / cognitive markers that can be measured in individuals with a given disorder and are also present more often in family members than in unrelated individuals (Gottesman and Gould 2003). Since endophenotypes are directly linked to genetic influences, they are particularly useful in the study of disorders with complex inheritance (Doyle et al. 2005), such as autism spectrum disorder (ASD).

Temperament research has a long history (see Rothbart 2011 for a review), with the most well-known framework originating from the work of Chess and Thomas (1990). The definition of temperament for the current study was based on that of Rothbart and her colleagues, built upon the pioneering work of Thomas and Chess. Rothbart and Bates (2006) defined temperament as individual differences in reactivity and regulation. Reactivity refers to emotional valence and intensity, which are dependent on interaction between the autonomic nervous system and the brain (Porges 2005). Factor analysis of childhood temperament traits using questionnaires developed by Rothbart and colleagues has consistently yielded

✉ Nancy Garon
ngaron@mta.ca

¹ Psychology Department, Mount Allison University, Sackville, New Brunswick E4L 1C7, Canada

² Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada

³ Autism Research Centre, Glenrose Rehabilitation Hospital, Edmonton, Alberta, Canada

⁴ Dalhousie University/IWK Health Centre, Halifax, Nova Scotia, Canada

⁵ Bloorview Research Institute, Toronto, Ontario, Canada

⁶ University of Toronto, Toronto, Ontario, Canada

⁷ Kinark Child and Family Services, Markham, Ontario, Canada

⁸ University of Ottawa, Ottawa, Ontario, Canada

⁹ The Hospital for Sick Children, Toronto, Ontario, Canada

three main factors – Extraversion, Negative Affectivity, and Effortful Control (Rothbart 2011). The first two factors are closely tied to emotional reactivity, while the third factor is regulatory. Extraversion is defined as an approach system; positive affect is an important component of this factor. Negative Affectivity refers to a child's propensity to experience distress, including sadness, anxiety and frustration. Finally, Effortful Control involves the ability to regulate reactivity through the use of attention (Rothbart and Posner 2006).

Temperament and ASD

ASD is a disorder involving problems with social communication and restricted/repetitive interests (American Psychiatric Association [APA], 2013). Accumulating evidence indicates that individuals with ASD show a specific temperament profile consisting of lower positive affect, higher negative affect, and lower regulation (e.g., Garon et al. 2009). Compared to typically developing children, children with ASD are perceived as exhibiting lower positive affect and neutral facial expression (i.e., little smiling and low approach; Adamek et al. 2011; Capps et al. 1993; Loveland et al. 1994; Konstantareas and Stewart 2006). In adulthood, individuals with ASD are reported to show lower novelty seeking and reward dependence (Anckarsäter et al. 2006; Soderstrom et al. 2002), higher shyness and introversion (Anckarsäter et al. 2006; Ozonoff, Garcia, Clark, and Lainhart 2005), and higher social withdrawal than typical controls. In contrast to positive affect, higher negative affect is common in children and adults with ASD, when compared to both typical and developmentally delayed controls (Adamek et al. 2011; Anckarsäter et al. 2006; Capps et al. 1993). Finally, lower levels of focusing attention, shifting attention, and inhibiting prepotent responses have been reported for children with ASD than for either typical or developmentally delayed controls (Adamek et al. 2011; Konstantareas and Stewart 2006). Similarly, children with ASD have been rated as less flexible, more difficult to distract and less goal-oriented (Brock et al. 2012). In adults with ASD, problems with inattention (Bradley and Isaacs 2006) and lower levels of self-directedness and responsibility (Anckarsäter et al. 2006; Soderstrom et al. 2002) also implicate low Effortful Control and Conscientiousness.

More recently, this pattern of low positive affect, high negative affect and low regulation has been confirmed in prospective studies of high-risk infants subsequently diagnosed with ASD (Bryson et al. 2007; Clifford et al. 2013; del Rosario et al. 2014; Filliter et al. 2015; Garon et al. 2009). Notably, two studies have indicated that infants who are subsequently diagnosed with ASD show distinctive temperament trajectories in the first year of life (Clifford et al. 2013; del Rosario et al. 2014). For example, Clifford et al. (2013) found

evidence of lower positive affect, higher negative affect, and lower levels of regulation in children from the high-risk (HR) group who were later diagnosed with ASD, compared to children in the low-risk (LR) group. Moreover, children who were in the HR group and who were not diagnosed with ASD also showed significant differences in temperament compared to the LR group. The overall pattern of findings from these two studies indicates that temperament distinguished children later diagnosed with ASD not only from LR controls, but also from their HR peers who did not develop ASD.

Development of Higher Order Temperament Traits

In a prospective study of infant siblings of children with ASD, Garon et al. (2009) found that the HR group was distinguished from a LR control group at 24 months by a combination of low scores on Effortful Control and high scores on negative affect scales from a parent-report inventory. HR infants who were diagnosed with ASD at 36 months were distinguished from their non-ASD counterparts by a combination of low scores on positive anticipation and attention shifting. These distinctions parallel the higher-order factors described by Evans and Rothbart (2009)— one factor was composed of scales related to positive affect and orienting sensitivity and the other, scales related to low negative affect and high effortful control. Two similar parallel meta-traits have been found in the personality literature (DeYoung 2006). DeYoung (2006) labeled the higher-order trait with positive affect as Plasticity (reflecting the tendency to explore) and the trait with negative affect, Stability (reflecting stable emotional, social and motivational regulation).

Patterns of correlations among temperament dimensions reported during childhood suggest a similar organization (Rothbart and Bates 2006). Furthermore, Shiner and DeYoung (2013) argued that antecedents of these meta-traits are present during infancy, with positive affect being a precursor to Plasticity and negative affect being a precursor to Stability. Evans and Rothbart (2009) suggested that these higher-order factors may reflect how the temperament dimensions of reactivity and regulation become organized over time. Of particular interest for ASD is that, in typical development, positive affect during infancy predicted later Effortful Control (Komsis et al. 2006), whereas no concurrent association between the two has been generally found.

The Present Study

The current study had five main goals. First, we were interested in specific temperament differences between children who had a sibling with ASD (HR) and those with no first-degree relatives with ASD (LR). We predicted that the HR group

would be rated higher by parents on scales assessing negative affect (e.g., fear) and lower on scales assessing regulation (e.g., inhibitory control). Second, we were interested in whether the reactive and regulatory aspects of temperament were organized similarly in HR and LR groups. For this latter purpose, we used structural equation modeling (SEM) to determine whether early reactive aspects of temperament (i.e., positive and negative affect) at 12 months predicted later reactive and regulatory aspects of temperament at 24 months. Based on previous findings in typically developing children (e.g., Komsis et al. 2006), we expected that positive affect would play a more important role than negative affect in predicting effortful control in both HR and LR groups.

The remaining goals were related to early temperamental heterogeneity within the HR group. First, we were interested in the associations between temperament factors (at 12 and 24 months) and ASD symptoms (at 36 months) in our SEM. Based on our and others' previous findings (e.g., Clifford et al. 2013), we expected that differences in positive affect and effortful control would be associated with ASD symptom severity. Second, we were interested in temperament differences among HR subgroups based on timing of diagnosis (Garon et al. 2009). Again, we expected to find differences between the early and later diagnosed groups on positive affect and regulation. Finally, given findings of a male bias in ASD prevalence and possible differences in phenotypic expression of symptoms in boys and girls (Zwaigenbaum et al. 2012), we included sex as a moderator in our analyses.

Method

Participants

The sample included 545 infants from our prospective infant sibling study who were followed from the first year until 36 months of age. Of these, 383 were HR infants with an older sibling with ASD, 98 of whom were diagnosed with ASD at 36 months (25.6 %). The ASD group comprised 50 children who were first diagnosed at 24 months ('Early Diagnosis') and 48 first diagnosed at 36 months ('Late Diagnosis'). The remaining 162 were LR control infants, with no known first- or second-degree relatives with ASD. Most infants in both the HR and LR groups were enrolled in the study at 6 months of age (84 % and 80 %, respectively), and the remainder by 12 months. All of the infants were born full-term (at least 37 weeks gestation). Participants were recruited through four major ASD diagnostic and treatment centers in Canada. Diagnoses of the older siblings were confirmed by expert clinical judgment based on DSM-IV-TR criteria (APA, 2000), and in most cases, included the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2000) and the Autism Diagnostic Interview-Revised (ADI-R; Lord et al.

1994). None of the older or younger siblings had a known genetic or chromosomal syndrome (e.g., fragile X syndrome) or neurological disorder (e.g., tuberous sclerosis) that could account for the ASD. The large majority of participants were Caucasian (85.3 %) and came from families classified as medium to high socioeconomic status (77.7 %). Participant characteristics are shown in Table 1.

Procedures

As part of the protocol for our larger prospective infant sibling study, the *Infant Behavior Questionnaire* (IBQ; Rothbart 1981) was completed by parents at 12 months and the *Toddler Behavior Assessment Questionnaire-Revised* (TBAQ-R; Goldsmith 1996; Rothbart et al. 2003) was completed by parents at 24 months. At 36 months, all children were assessed for ASD, blind to risk status and previous assessments, using a combination of the ADI-R (Lord et al. 1994), ADOS (Lord et al. 2000) and expert clinical judgment with reference to DSM-IV-TR. The children were also assessed on the *Mullen Scales of Early Learning* (MSEL; Mullen 1995) at 12, 24 and 36 months. Ethics approval was granted for each site involved in the study and written, informed consent was obtained from a parent or guardian of each infant participant.

Measures

Infant Behavior Questionnaire (IBQ) The IBQ (Rothbart 1981) is a reliable and well-validated measure of temperament designed for infants aged 3–12 months (for a review, see Rothbart and Bates 2006). The questionnaire is comprised of six subscales with 94 items. Items are scored on a Likert scale from 1 (*never*) to 7 (*always*) and are averaged for each scale. For the present sample, Cronbach alphas ranged from 0.73 to 0.86 for the scales (see Table 2).

Toddler Behavior Assessment Questionnaire-Revised The TBAQ-R is a reliable and well-validated measure of temperament designed for infants aged 18–35 months (Goldsmith 1996; Rothbart et al. 2003). It is comprised of 13 scales, including the original four scales from the TBAQ (Activity Level, Anger/Frustration, Positive Anticipation, and Social Fear; Goldsmith 1996). Items are scored using the same Likert scale as the IBQ above. For the present sample, Cronbach alphas ranged from 0.68 to 0.89 for the scales (see Table 2).

Autism Diagnostic Observation Schedule The ADOS (Lord et al. 2000) is a semi-structured, standardized observational assessment designed to elicit behavior that is characteristic of ASD. It has been found to reliably distinguish children with ASD from typical and developmentally disabled non-autistic

Table 1 Participant characteristics (Mean Scores) as a function of group

Group	Mullen ELC			ADOS		ADI
	12 months	24 months	36 months	Social/ Communication 36 months	Severity Score 36 months	Total Score 36 months
ASD siblings <i>N</i> =95 (29 girls)	M=95.79 SD=16.62	M=86.74 SD=21.15	M=87.30 SD=23.71	M=11.54 SD=4.14	M=6.73 SD=2.07	M=23.83 SD=11.09
Non-ASD siblings <i>N</i> =278 (138 girls)	M=104.57 SD=14.66	M=106.78 SD=16.77	M=109.92 SD=17.45	M=4.02 SD=3.25	M=2.35 SD=1.72	M=5.91 SD=5.61
Controls <i>N</i> =161 (74 girls)	M=110.13 SD=11.79	M=118.34 SD=15.69	M=119.09 SD=15.82	M=3.16 SD=2.99	M=1.80 SD=1.39	M=3.85 SD=3.79

M Mean, *SD* Standard Deviation

controls (Lord et al. 2000). All ADOS administrations were conducted by examiners who had attained reliability according to the developers' criteria.

Autism Diagnostic Interview-Revised The ADI-R (Lord et al. 1994) is a structured parent interview designed to elicit information on developmental history, assessing various behaviors associated with ASD. This instrument has been found to discriminate children ASD from other developmental disabilities and has excellent inter-rater reliability (Lord et al. 1994). ADI-R interviews were conducted by research-reliable examiners.

Mullen Scales of Early Learning The MSEL (Mullen 1995) contains items assessing motor, cognitive, and language development in children from birth to 68 months. The test has been found to have good validity and test-retest reliability (Mullen 1995).

Statistical Approach

To assess overall differences on temperament scales between the LR and HR groups, a multivariate analysis of variance (MANOVA) was used with group (HR, LR) and sex as the independent variables. Dependent variables included all

Table 2 Means and standard deviations of temperament at 12 and 24 months as a function of group

Temperament scale		High-risk (<i>n</i> =373)	Control (<i>n</i> =161)	Group Comparisons			
	Scale internal reliability	M (SD)	M (SD)	<i>F</i>	η^2 [95 % CI]	<i>p</i>	Benjamini-Hochberg corrected α
IBQ (12 months)							
Distress to limitations	0.86	3.38 (0.89)	3.50 (0.78)	35.50*	0.071 [0.03; 0.12]	0.000	0.0080
Fear	0.76	3.21 (0.80)	2.88 (0.71)	19.65*	0.040 [0.01; 0.08]	0.000	0.0167
Smiling and laughter	0.82	4.92 (0.86)	5.09 (0.69)	3.68	0.008 [0.00; 0.03]	0.056	0.0333
Activity level	0.84	4.25 (0.88)	4.09 (0.89)	3.67	0.008 [0.00; 0.03]	0.056	0.0250
Soothability	0.82	5.23 (0.90)	5.30 (0.79)	0.60	0.001 [0.00; 0.02]	0.441	0.0417
Duration of orienting	0.73	2.70 (0.95)	2.66 (0.97)	0.32	0.001 [0.00; 0.01]	0.573	0.0500
TBQ (24 months)							
Anger	0.69	4.02 (0.92)	3.42 (0.85)	44.90*	0.091 [0.05; 0.14]	0.000	0.0045
Sadness	0.68	3.43 (0.76)	3.00 (0.66)	34.45*	0.072 [0.03; 0.12]	0.000	0.0090
Inhibitory Control	0.89	3.97 (0.93)	4.36 (0.80)	19.58*	0.042 [0.01; 0.08]	0.000	0.0136
Soothability	0.89	5.22 (0.86)	5.52 (0.59)	10.75*	0.024 [0.00; 0.06]	0.001	0.0227
Fear	0.78	3.83 (0.98)	3.52 (0.95)	10.51*	0.023 [0.00; 0.06]	0.001	0.0182
Attention Focus	0.78	3.96 (0.74)	4.17 (0.66)	8.49*	0.019 [0.00; 0.05]	0.004	0.0273
High Pleasure	0.77	4.71 (0.84)	4.91 (0.72)	5.32*	0.012 [0.00; 0.04]	0.022	0.0364
Low Pleasure	0.80	5.47 (0.76)	5.65 (0.63)	5.20*	0.012 [0.00; 0.04]	0.023	0.0318
Attention Shifting	0.82	4.81 (0.97)	5.01 (0.70)	3.04	0.007 [0.00; 0.03]	0.082	0.0409
Activity Level	0.71	4.25 (1.0)	4.30 (0.90)	0.23	0.001 [0.00; 0.01]	0.630	0.0500
Positive Anticipation	0.75	4.54 (0.84)	4.57 (0.82)	0.07	0.000 [0.00; 0.01]	0.799	0.0455

M Mean, *SD* Standard Deviation

temperament scales at 12 months and at 24 months. These were followed up by Benjamini-Hochberg corrected F -tests (Benjamini and Hochberg 1995). In this method, the p -values are ordered smallest to largest. The alpha level for each test is then set at $\frac{k*\alpha}{m}$ with k corresponding to the p -value's rank (e.g., lowest p -value=1) and m corresponding to the number of comparisons; in this case, six. The comparisons stop once one of the t -tests is rejected.

To construct the temperament factors, two confirmatory factor analyses (CFAs) were conducted on the temperament scales at 12 and 24 months, using the method of maximum likelihood estimation from AMOS (Arbuckle 2006). Ninety-four items for the Infant Behavior Questionnaire and 144 items for the Toddler Behavior Questionnaire exceeded the ratio-to-parameter estimate of 10:1 recommended for CFA (Kline 2011). For each CFA, temperament scales were used as item parcels. Although the use of item parceling has been debated in the SEM literature (see Little et al. 2013), this strategy was advantageous for the current study. First, parcels compared to items typically show better psychometric characteristics such as higher reliability and have higher ratio of common-to-unique factor variance (Little 2013). Second, when a questionnaire has many items, the use of parcels can bring the ratio of parameters to sample size to acceptable levels (Little 2013).

Given SEM is sensitive to departures from normality, most variables were transformed due to either negative or positive skewness using Box Cox transformation (Osborne 2010). Each variable was also re-scaled to range from 0 to 1, as suggested by Little (2013). To provide a robust measure of symptoms (Wiggins et al. 2014), the CFA run for ASD symptoms at 36 months included scores from both ADOS and ADI-R. Five variables were included in this CFA: (1) ADOS Social Affect (SA) score, (2) ADOS Repetitive and Restricted Behavior (RRB) score, (3) ADI-R social score, (4) ADI-R communication score, and (5) ADI-R repetitive behavior score. To derive ADOS SA and RRB scores, scores from the original ADOS were transposed onto ADOS-2 algorithms (Lord et al. 2012).

Assumptions of multivariate normality were evaluated using Mahalanobis distances and Mardia's test for multivariate outliers (Mardia 1970). Separate analyses of Mahalanobis distances for the LR and HR group were conducted with all variables. These analyses revealed 9 multivariate outliers, all p 's < 0.001 (1 LR; 8 HR, including 3 with ASD). When these cases were removed, Mardia's test indicated conditions of multivariate normality were met for the LR group, $p=0.623$ and for the HR group, $p=0.187$. Table 1 shows the size of our final sample.

The three CFAs were tested for measurement invariance, to test for equivalence in latent constructs across the LR and HR groups. Four levels of measurement invariance are generally investigated when comparing groups (Widaman & Reise

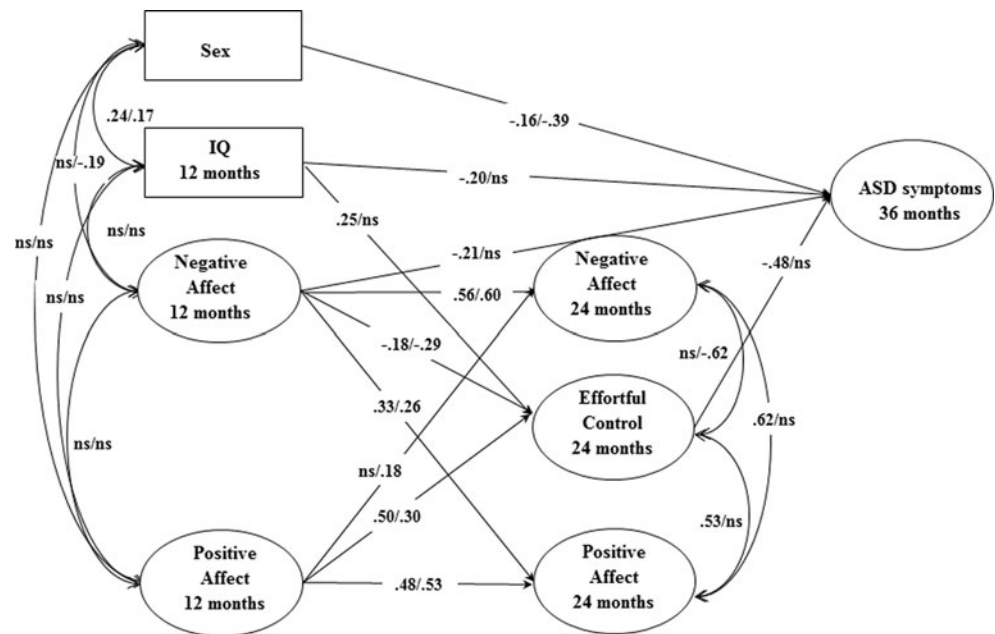
1997). The first level, *configural invariance*, tests whether the two groups have the same general factor structure. *Metric invariance* tests whether groups have the same factor loadings, whereas *scalar invariance* (Widaman and Reise 1997) tests the additional requirements of the same intercepts for the two groups. Finally, *residual invariance* (Widaman and Reise 1997) adds the constraint of equal error variances. Since each model was nested within the other, they were compared using chi-square differences. In addition, differences in the CFI $|\leq 0.01|$ (Cheung and Rensvold 2002) and RMSEA $|\leq 0.01|$ (Chen 2007) were used to evaluate equivalence since using the chi-square difference test may result in rejection equivalence for minor differences (Little 2013).

The three fitted CFAs were used as a foundation for the SEMs. The first fitted model consisted of the 12-month temperament factors predicting ASD symptoms at 36 months. MSEL Early Learning Composite scores at 12 months were used as a covariate. The model consisted of all possible calculated paths from the temperament factors and IQ to ASD symptoms at 36 months. Finally, to explore whether Effortful Control at 24 months mediated the predicted effect of early Positive Affect on ASD symptoms, the temperament factors at 24 months were added to the model (see Fig. 1 for final model). Preliminary analyses showed that sex reached metric invariance for all three CFAs ($p < 0.1$ for all), indicating that factor loadings were the same for boys and girls for temperament factors and ASD symptoms. As a result, boys and girls could be included in the same SEM and sex was only used as a covariate in the SEM models to control for this variable.

Data were missing for 8.1 % of data points in the HR group and 4.7 % of data points in the LR group. Analyses were conducted to determine whether participants with missing data differed on sex, SES, or IQ; all differences were non-significant. Missing data were estimated using the Full Information Maximum Likelihood (FIML) in AMOS (Kline 2011). Maximum likelihood is relatively robust to bias even under conditions of moderate non-normality (Curran, West, and Finch 1996). Model fit was assessed via two criteria. The root mean square error of approximation (RMSEA) indicates how well the model parameters would fit the population covariance matrix, with 0.05 to 0.08 considered acceptable fit, and 0.01 to 0.05 considered close fit (Little 2013). The comparative fit index (CFI) corrects for sample size, with values of 0.90 to 0.95 indicative of acceptable fit, and values > 0.95 indicating good fit (Hu and Bentler 1999).

To investigate differences in temperament factors within the HR group, a multivariate profile analysis was conducted with group (early-, late-, and non-ASD sibs) as the independent variable, and temperament factors at 12 and 24 months as the dependent variables. The five temperament factors calculated in the CFA were used as the dependent measures.

Fig. 1 The structural equation models for the high-risk and control group. High-risk/low-risk groups. Values represent standardized coefficient that are statistically significant at $p < 0.05$. Non-significant parameters remain in the model but are not displayed in the figure



Results

Temperament Differences between the HR and LR Groups

The means and standard deviations for all temperament scales by group and age are shown in Table 2. For the 12-month temperament scale, the group effect was significant, $F(6, 462) = 7.14$, $p < 0.001$, $\eta^2 = 0.085$, 95 % CI [0.03; 0.13]. The Sex and Sex X Group interaction were both non-significant, both F 's < 1 . The group effect was followed up by Benjamini-Hochberg corrected t-tests for each scale. Using this method, Distress to Limitations and Fear were found to be significant. As can be seen in Table 2, at 12 months the HR group was rated by parents as higher on Distress to Limitations and Fear compared with LR controls.

For the 24-month temperament scales, the group effect was also significant, $F(11, 437) = 7.42$, $p < 0.001$, $\eta^2 = 0.157$, 95 % CI [0.08; 0.20]. The main effect of sex was significant, $F(11, 437) = 4.58$, $p < 0.001$, $\eta^2 = 0.010$, 95 % CI [0.04; 0.14] but the Sex X Group interaction was non-significant, $F(11, 437) < 1$. There were significant sex differences for Inhibitory Control, $F(1, 447) = 15.62$, $p < 0.001$, $\eta^2 = 0.034$, 95 % CI [0.01; 0.07], Activity Level, $F(1, 447) = 11.51$, $p < 0.001$, $\eta^2 = 0.025$, 95 % CI [0.01; 0.06], Anger, $F(1, 464) = 10.75$, $p = 0.001$, $\eta^2 = 0.024$, 95 % CI [0.01; 0.06] and Attention Shifting, $F(1, 447) = 9.50$, $p = 0.001$, $\eta^2 = 0.021$, 95 % CI [0.00; 0.05]. Boys were rated by parents as higher on Activity Level ($M = 4.42$, $SD = 0.92$) and Anger ($M = 3.95$, $SD = 0.92$) compared with girls for Activity Level ($M = 4.09$, $SD = 0.99$) and Anger ($M = 3.69$, $SD = 0.95$). In contrast, boys were lower on Inhibitory Control ($M = 3.95$, $SD = 0.88$) and Attention Shifting ($M = 4.77$,

$SD = 0.90$), compared with girls for Inhibitory Control ($M = 4.27$, $SD = 0.91$) and Attention Shifting ($M = 4.99$, $SD = 0.89$).

There were significant group (HR vs. LR) differences on Anger, Sadness, Inhibitory Control, Soothability, Fear, Attention Focus, High Pleasure, and Low Pleasure. Temperament means for both groups, shown on Table 2, reveal that the HR group was rated by parents as higher than controls on Fear, Sadness, and Anger, but lower on Inhibitory Control, Soothability, Attention Focus, High Pleasure, and Low Pleasure.

Confirmatory Factor Analysis for IBQ at 12 months

A preliminary CFA indicated that Orienting did not load significantly onto any factor and so was removed from further analysis. Table 3 provides a summary of the results of the group measurement invariance testing and standardized factor loadings. Smiling and Soothability at 12 months loaded significantly onto the Positive Affect factor, whereas Activity Level and Distress to Limitations loaded significantly onto the Negative Affect factor. Fear cross-loaded negatively onto the Positive Affect factor. The unconstrained model estimated for the two groups separately showed adequate fit, providing support for configural invariance. Constraining the factor loadings to be equivalent for the groups to test for metric invariance resulted in an adequately fitted model; the difference between the models was not significant (Table 3), suggesting that factor loadings were equivalent for HR and LR groups. Constraining the intercepts to be equivalent across group for scalar invariance produced a poorer fitting model (see Table 3), suggesting group differences on subscale intercepts. Examining critical ratios indicated group differences on

Table 3 Summary of results from confirmatory factor analysis and structural equation modelling

Model and number of factors	χ^2	<i>df</i>	<i>p</i>	RMSEA	CFI	Model Comparison	$\Delta\chi^2$	<i>p</i>	Δ RMSEA	Δ CFI	
12-month Temperament											
1. Configural Invariance (a)	7.11	5	0.212	0.028	0.993						
2. Metric Invariance (b)	11.54	8	0.173	0.029	0.988	A vs B	4.43	0.351	0.001	0.005	
3. Scalar Invariance (c)	53.71	13	0.000	0.077	0.857	B vs C	42.17	<0.001	0.048	0.131	
4. Residual Invariance (d)	56.96	18	0.000	0.064	0.863	C vs D	3.25	0.666	-0.013	-0.006	
Factor loadings (standardized)											
IBQ Scales											
	Positive Affect (HR/LR)				Negative Affect (HR/LR)						
Smiling/Laughter	0.96/0.80										
Soothability	0.42/0.35										
Activity					0.51/0.43						
Distress to Limitation					0.93/0.89						
Fear	-0.30/-0.23				0.37/0.34						
24-month Temperament											
1. Configural Invariance (a)	178.16	69	0.000	0.055	0.931						
2. Metric Invariance (b)	199.20	81	0.000	0.052	0.926	A vs B	21.045	0.049	-0.003	0.005	
3. Scalar Invariance (c)	282.05	92	0.000	0.062	0.880	B vs C	82.86	<0.001	0.010	0.046	
4. Residual Invariance (d)	295.62	103	0.000	0.059	0.887	C vs D	12.43	0.258	-0.003	0.001	
Factor loadings (standardized)											
TBQ Scales											
	Positive Affect (HR/LR)				Negative Affect (HR/LR)			Effortful Control (HR/LR)			
High Pleasure	0.36/0.37										
Activity Level	0.49/0.52							-0.69/-0.61			
Positive Anticipation	0.74/0.71										
Fear	-0.56/-0.52				0.56/0.47						
Sadness					0.76/0.67						
Anger					0.75/0.67			-0.34/-0.28			
Soothability								0.58/0.55			
Low Pleasure	0.19/0.19							0.54/0.45			
Inhibitory Control	-0.21/-23							0.95/0.84			
Attention Shifting								0.84/0.81			
Attention Focus								0.41/0.35			

Bold indicate level of variance passed for each CFA

Fear ($z=4.48, p<0.001$) and Anger ($z=6.11, p<0.001$), with the HR group having a higher intercept for Fear (HR=0.47, LR=0.40) and Anger (HR=0.51, LR=0.41).

Confirmatory Factor Analysis for TBQ at 24 months

At 24 months, the CFA of the unconstrained model with the two groups showed adequate fit, again providing support for configural invariance for the two groups. Table 3 shows the loadings from the CFA and test of group factorial invariance. High Pleasure, Activity Level and Positive Anticipation loaded positively onto a Positive Affect factor, whereas Fear and Inhibitory Control loaded negatively onto this factor. Fear, Sadness, and Anger loaded positively onto a Negative Affect factor. Finally, Soothability, Low Pleasure, Inhibitory Control, Attention Shifting and Attention Focus loaded

positively onto an Effortful Control factor, whereas Activity Level and Anger loaded negatively onto this factor. Constraining the factor loadings to be equivalent for the groups to test for metric invariance resulted in an adequately fitted model. While $\Delta\chi^2(12)=21.04, p=0.050$, indicated a marginal difference between the configural and metric models, using Δ CFI criteria of $|0.01|$ (Cheung and Rensvold 2002), the Δ CFI=0.005 and slight improvement in RMSEA (Chen 2007) provided evidence for the groups having passed metric invariance. Constraining intercept loadings for scalar invariance, however, resulted in a significantly poorer fitting model (see Table 3), suggesting that groups were different on subscale intercepts. Examination of critical ratios indicated a significant group difference on the majority of subscale intercept, including Fear, Anger, Sadness in which the HR had higher intercepts and Soothability, Low pleasure, Inhibition,

High Pleasure, and Attention Focus in which the HR had lower intercepts.

Examining the associations among the temperament factors indicated some interesting differences between the groups. While the association between Positive Affect and Negative Affect was not significantly different for the HR ($\beta=0.60$) and LR ($\beta=0.39$) groups ($z=1.73$, $p=0.08$), the association between Negative Affect and Effortful Control was significantly different for the two groups ($z=3.17$, $p=0.002$). Specifically, the association was not significant for the HR group ($\beta=-0.10$, $p=0.190$), but was significant for the LR group ($\beta=-0.58$, $p<0.001$). Finally, the association between Positive Affect and Effortful Control was significantly different for the two groups ($z=2.45$, $p=0.014$), with the association being stronger for the HR group ($\beta=0.55$, $p<0.001$) in comparison to the LR group ($\beta=0.26$, $p=0.018$).

Confirmatory Factor Analysis for ASD symptoms

For ASD symptoms at 36 months, a one-factor CFA of the unconstrained model with the two groups showed good fit, $\chi^2(4)=4.25$, $p=0.373$, CFI=1.0, RMSEA=0.011, again providing support for configural invariance for the two groups. For the HR group, the loadings for the ASD symptom factor were ADOS SA (0.69), ADOS RRB (0.57), ADI-R Social (0.87), ADI-R Communication (0.89), and ADI-R repetitive behavior (0.68), with all loadings being highly significant, $p<0.001$. For the LR group, loadings for the ASD symptom factor were ADOS SA (0.43), ADOS RRB (0.33), ADI-R social (0.41), ADI-R communication (0.64), and ADI-R repetitive behavior (0.41), with all loadings being highly significant, $p<0.001$. Constraining the factor loadings to be equal resulted in a good fitting model, $\Delta\chi^2(2)=0.788$, $p=0.753$, $\Delta\text{CFI}<0.001$. Constraining the intercepts to be the same for both groups resulted in a model with a significantly poorer fit, $\Delta\chi^2(5)=124.752$, $p<0.001$, $\Delta\text{CFI}=0.12$, with all intercepts being significantly higher for the HR group as compared to the LR group (all p 's<0.001).

Predicting Effortful Control and ASD symptoms

The first model consisted of only the 12-month temperament factors, sex, IQ at 12 months, and ASD symptoms at 36 months. This model showed an adequate fit, $\chi^2(99)=154.39$, $p<0.001$, CFI=0.961, RMSEA=0.032. For the HR group, the majority of paths were significant. Positive Affect ($\beta=-0.30$, $p<0.001$), sex ($\beta=-0.21$, $p<0.001$) and 12-month IQ ($\beta=-0.33$, $p<0.001$) significantly predicted ASD symptoms at 36 months. Negative Affect at 12 months was not significantly associated with later ASD symptoms ($\beta=-0.11$, $p=0.113$). For the LR group, only sex was significantly associated with later ASD symptoms ($\beta=-0.38$, $p<0.001$).

Next, the full model was fitted by incorporating the temperament factors at 24 months. The model showed adequate

fit, $\chi^2(404)=667.11$, $p<0.000$ (CFI=0.924, RMSEA=0.035). Non-significant paths for both groups were trimmed from the model. The new trimmed model was not significantly different, $\chi^2(6)=8.413$, $p=0.422$, from the full model. The final model had adequate fit, $\chi^2(410)=673.12$, $p<0.000$, CFI=0.924, RMSEA=0.035. Figure 1 shows the final trimmed model for both groups.

For the HR group, there was a strong positive association between respective 12-month and 24-month reactive factors, Negative Affect ($\beta=0.56$, $p=0.001$) and Positive Affect ($\beta=0.48$, $p=0.001$), indicating continuity in these constructs over that interval. As predicted, there was a strong association between Positive Affect at 12 months and Effortful Control at 24 months ($\beta=0.50$, $p=0.001$), whereas the association between Negative Affect at 12 months and Effortful Control at 24 months was significant, but small ($\beta=-0.18$, $p=0.01$).

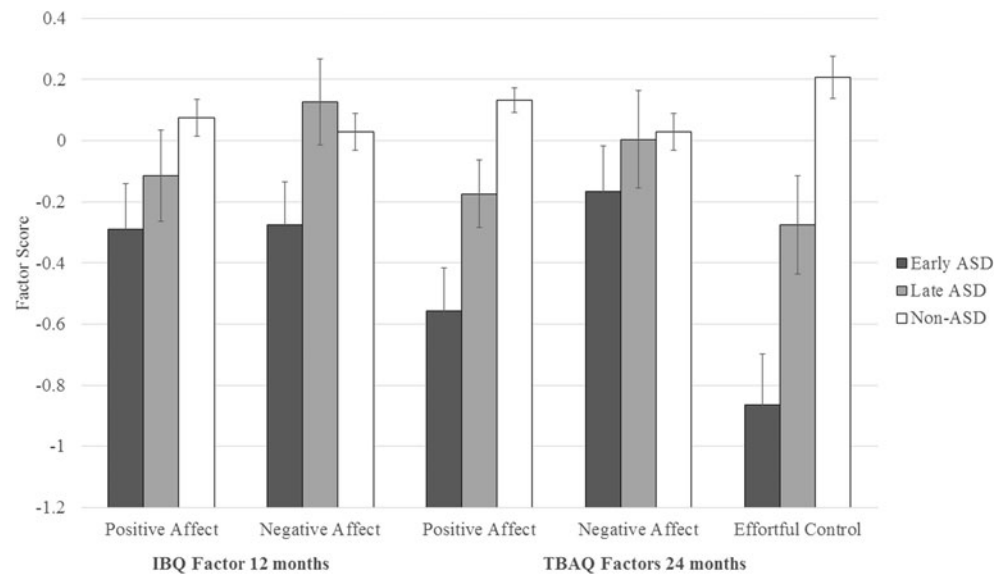
Four variables significantly predicted ASD symptoms at 36 months. Being female predicted lower levels of ASD symptoms ($\beta=-0.16$, $p<0.001$). Lower IQ at 12 months ($\beta=-0.20$, $p<0.001$) combined with lower Effortful Control score at 24 months ($\beta=-0.48$, $p<0.001$) predicted more ASD symptoms. Interestingly, Negative Affect at 12 months ($\beta=-0.21$, $p<0.001$) with 24-month temperament incorporated into the model, significantly predicted ASD symptoms (see Discussion). In contrast, the association between Positive Affect at 12 months and ASD symptoms was no longer significant ($\beta=-0.05$, $p=0.569$) (path removed in the trimmed model in Fig. 1). Bootstrapping was used to estimate the indirect effect of Positive Affect at 12 months on ASD symptoms at 36 months. The indirect effect was significant (95 % C.I., -0.038 to -0.019 , $p<0.001$), suggesting that the association between Positive Affect at 12 months and ASD symptoms at 36 months was fully mediated by its association with 24-month Effortful Control.

The LR group showed similar continuity between the 12-month and 24-month reactive temperament factors, Negative Affect ($\beta=0.60$, $p=0.001$) and Positive Affect ($\beta=0.53$, $p<0.001$). As with the HR group, higher Positive Affect at 12 months ($\beta=0.30$, $p=0.003$), and lower Negative Affect at 12 months ($\beta=-0.29$, $p=0.004$) predicted higher Effortful Control at 24 months at 24 months. In contrast to the HR findings, sex alone significantly predicted ASD symptoms at 36 months ($\beta=-0.39$, $p<0.001$) while IQ was marginally significant ($\beta=-0.17$, $p=0.099$).

Temperament within the HR group, Stratified by Diagnostic Outcome

Figure 2 shows the temperament factor scores as a function of diagnostic status in the HR group. Results of the MANOVA indicated a significant group effect, $F(2, 367)=22.54$, $\eta^2=0.109$, $p<0.001$, 95 % CI [0.05; 0.17] significant main effect of temperament scale, $F(4, 364)=2.75$, $\eta^2=0.029$, $p=0.028$,

Fig. 2 Scores on the temperament factors as a function of diagnosis for the high-risk sample



95 % CI [0.00; 0.06], and a significant Group X Temperament Scale interaction, $F(4, 364)=3.21$, $\eta^2=0.034$, $p=0.001$, 95 % CI [0.01; 0.05] suggesting different temperament profiles for the three groups. The main effect and all interactions with sex were non-significant, all p 's > 0.05. The significant interaction of Group X Temperament was followed up with two profile analyses comparing the non-diagnosed HR group with the diagnosed HR group as a whole, and a second comparing early- versus late-diagnosed HR children.

Diagnosed versus Non-Diagnosed HR Groups The group effect was significant, $F(1, 369)=34.78$, $\eta^2=0.086$, $p<0.001$, 95 % CI [0.04; 0.14], while the main effect of temperament scale was not, $F(4, 366)=1.54$, $\eta^2=0.017$, $p=0.189$, 95 % CI [0.00; 0.04]. The interaction of Group X Temperament Scale was significant, $F(4, 366)=6.03$, $p<0.001$, $\eta^2=0.062$, 95 % CI [0.02; 0.11]. The significant interaction was followed up with group comparisons on each temperament factor using Benjamini-Hochberg corrected t-tests. Using this method, the diagnosed group differed from the non-diagnosed HR group on the Effortful Control factor, $t(371)=5.99$, $p<0.001$, $d=0.66$, 95 % CI [0.47; 0.94], Positive Affect at 24 months, $t(371)=5.44$, $p<0.001$, $d=0.56$, 95 % CI [0.40; 0.87], and Positive Affect at 12 months, $t(371)=2.40$, $p=0.017$, $d=0.27$, 95 % CI [0.05; 0.51].

Early- versus Late-Diagnosed HR Group The group effect was significant, $F(1, 95)=7.30$, $p=0.008$, $\eta^2=0.071$, 95 % CI [0.01; 0.18]. The early-diagnosed group had a lower overall temperament score ($M=-0.43$, $SE=0.09$), compared with the late-diagnosed group ($M=-0.09$, $SE=0.09$). Although the main effect of temperament scale, $F(4, 92)=3.81$, $p=0.007$, $\eta^2=0.142$, 95 % CI [0.01; 0.25] was significant, the interaction of Group X Temperament Scale effect, $F<1$ was not

significant, suggesting the key difference between the early- and late-diagnosed groups was in parents' overall temperament intensity ratings. Figure 2 shows that the two diagnosed groups had similar temperament profiles, and that the early-diagnosed group had lower scores on all temperament factors, compared with the late-diagnosed group.

Discussion

The main goals of the current study were to examine early differences in a sample of infants at high risk for developing ASD compared to a LR sample, and to examine temperament differences within the HR sample. In relation to this, we were interested in potential differences in how early reactive components of temperament might predict later regulatory traits in the HR versus the LR samples. We were also interested in the heterogeneity of temperament within this infant sibling sample and how these differences in temperament might be associated with later ASD symptoms. Our results extend previous findings by indicating that early reactive temperament components jointly predict later regulatory aspects of temperament at 24 months in both LR and HR, and that this regulatory aspect of temperament in turn predicts later ASD symptoms at 36 months in the HR sample.

Temperament Differences for High- versus Low-risk Groups

The HR group as a whole was characterized by high levels of negative affect. Multivariate analyses indicated group differences on all temperament scales that loaded on the Negative Affect factor at both 12 and 24 months. For instance, the HR group was rated higher on scales assessing fear and anger at 12

and 24 months. Furthermore, parents rated their HR children as lower on 5 of the 7 of the scales that loaded on the Effortful Control factor, indicating not only higher negative affect but also lower levels of regulation in this group. This is consistent with evidence that individuals with the broader autism phenotype express higher levels of negative emotion than the general population (Wainer et al. 2011).

Findings from SEM provided evidence of differences in association of temperament factors among HR and LR infants at 24 months. Although a strong positive relation between the Positive and Negative Affect factors was found for the HR group, the control group showed no such association. One possible explanation for this is a disturbance in the autonomic nervous system, such as higher sympathetic arousal (Goodwin et al. 2006; Kushki, Brian, Dupuis, and Anagnostou 2014) or lower parasympathetic arousal (Porges 2005), either of which may influence both positive and negative affect. Emotional states such as happiness, fear, and anger are associated with increased sympathetic arousal (Rainville, Bechara, Naqvi, and Damasio 2006) with this activation being subject to modulation by the parasympathetic nervous system and frontal brain systems (Thayer et al. 2012). Lower levels of regulatory influences by the parasympathetic nervous system and by the frontal brain system, as reported in individuals with ASD (Eilam-Stock et al. 2014; Porges 2005), may lead to a stronger association between positively and negatively valenced emotion in this population.

The SEM finding relates to the two meta-traits discussed in the introduction, Plasticity and Stability, both of which include reactive and regulatory components and have been hypothesized to reflect integration of temperament traits through development (Evans and Rothbart 2009). Recall that Plasticity includes high positive loadings from scales measuring positive affect and regulation, whereas Stability includes negative loadings from scales that assess negative affect and positive loadings from scales that assess regulation (DeYoung 2014). In the LR sample, the strong negative association between Negative Affect and Effortful Control factors is consistent with past findings of formation of a Stability meta-trait in typically developing children. Previous reports of this negative association have been taken to indicate that Effortful Control is important for regulating negative emotions (Rothbart and Bates 2006). The failure to find a significant negative correlation in the present study between the Effortful Control and Negative Affect factors in the HR group could indicate a problem with integration of the brain networks underlying these temperament traits.

Our LR control group findings replicated work suggesting an important role for positive affect in predicting later Effortful Control in typical infant development (e.g., Komsis et al. 2006). Again, these findings are relevant for the early development of the two meta-traits. The longitudinal association of positive affect at 12 months with later 24-month

regulation may reflect integration of the networks underlying temperament traits to form the meta-trait of Plasticity. In fact, the presence of high Positive Affect at 12 months and its association with Effortful Control at 24 months seemed to distinguish those HR children who were not diagnosed with ASD from those who were. Perhaps a certain level of positive affect is necessary for the development of the regulatory aspect of temperament.

Heterogeneity Within the High-Risk Group and Association with ASD Symptoms

A critical finding in the SEM was the negative association between 24-month Effortful Control and subsequent ASD symptoms and diagnosis. A temperamental profile characterized by lower regulation has been one of the most consistent findings in studies of children with ASD (Adamek et al. 2011; Clifford et al. 2013). The current profile analysis further supported the importance of this construct in the prediction of ASD in HR infants. Also notable was the association between Positive Affect at 12 months and ASD symptoms at 36 months, and the finding that the inclusion of Effortful Control at 24 months eliminated this association. The implication is that 12-month Positive Affect influences later ASD symptoms through its effect on Effortful Control. This finding was further supported by the profile analysis, indicating that the critical difference between HR children with and without an ASD diagnosis was the former group's profile of lower Positive Affect at 12 and 24 months and lower Effortful Control at 24 months. Thus, the combination of higher positive affect and regulation, which we have linked with the meta-trait of Plasticity, was associated with a better outcome, and could be considered a resilience factor in the HR group.

Finally, contrast analyses of temperament profiles in the earlier versus later diagnosed groups revealed that the two differed in the overall intensity of the temperament profile rather than on any one temperament factor. Figure 2 shows similar profiles of lower Positive Affect and Effortful Control relative to Negative Affect, particularly evident at 24 months. Although the lower levels of Positive Affect and Effortful Control at 24 months for the early-diagnosed group were expected, the lower level of Negative Affect was not. This finding indicated that children who were diagnosed earlier with ASD, presumably because their impairments were evident earlier, actually showed less negative affect. It may be that lower levels of affect, regardless of valence, are associated with more severe ASD symptoms, a profile reminiscent of the “aloof” ASD subtype suggested by Wing and Gould (1979).

Temperament as a Possible Endophenotype

Garon et al. (2009) proposed that exploring differences between children with and without ASD in respect to

temperament profiles rather than individual temperament traits could help further our understanding of this heterogeneous disorder. The current findings, combined with those of recent studies exploring early temperament development (e.g., Clifford et al. 2013; Del Rosario et al. 2014) indicate that the search for a viable temperament endophenotype may also involve investigating associations among temperament traits in HR populations, providing clues to the integration of brain systems underlying these temperament traits.

Comparing the development of temperament traits in children from HR samples to that of children from the general population can provide clues to possible mechanisms leading to differences in timing and symptom presentation. For instance, early appearance of ASD symptoms may implicate reactivity and orienting attention systems since reactivity (positive and negative affect) and its regulation by the orienting attention network show a developmental spurt during the first 2 years of life (Posner et al. 2012). Appearance of ASD symptoms in the third year, conversely, may implicate problems in the organization of two attention networks, the orienting system and the executive attention system, which show segregation during this time (Posner et al. 2012).

Limitations and Conclusion

The main limitation of this study is that parent-rated temperament data are subject to report biases. Parents who already have a child with ASD may have perceived their younger children's development differently than other parents. Given this possibility, it will be important to assess temperament using direct observational measures (such as Lab-TAB; Gagne, van Hulle, Aksan, Essex, and Goldsmith 2011). In this regard, however, it should be noted that parent reports and direct observations are not strongly correlated, in part because they reflect different views of the child (Rothbart 2011). Another limitation is the use of item parceling for the CFAs, which has been criticized in the field of SEM (Little 2013). Finally, caution should be used in the interpretation of the group differences (Table 2) as our SEM analysis indicated that the HR and LR groups differ in the structure of temperament factors.

In conclusion, the current results extend findings on temperament and ASD by underscoring the importance of early temperament patterns for understanding the complexity and heterogeneity of ASD. Critically, the findings highlight the importance of examining differences between HR and LR groups, as well as differences within the HR group. A combination of high negative affect and low regulation appeared to characterize the HR group overall. In contrast, two temperament patterns appeared to be associated with an ASD diagnosis. One pattern was that of lower early positive affect combined with lower regulation. The second pattern, overall lower reactivity and lower regulation, was associated with earlier

diagnosis and more severe ASD symptoms. Our findings indicate that the search for a temperament endophenotype for ASD must involve consideration of the hierarchical structure of temperament and personality (DeYoung 2014). Development of temperament involves not just individual temperament traits, but also the integration of these traits. Finally, an important future step will be to explore how the early dispositions of HR infants interact with their social environments to construct core abilities such as emotion regulation and self-control.

This research was supported in part by grants from the Canadian Institutes of Health Research, Autism Speaks Canada, and NeuroDevNet. Dr. Zwaigenbaum was supported by the Stollery Children's Hospital Foundation Chair in Autism Research. Drs. Bryson and Smith were supported by the Joan and Jack Craig Chair in Autism Research, Dr. Szatmari was supported by the Chedoke Health Chair in Child Psychiatry, and Dr. Vaillancourt was supported by a Canada Research Chair in Children's Mental Health and Violence Protection. We also like to thank the children and families who have participated in this project.

Conflict of interest The authors have no conflicts of interest.

References

- Adamek, L., Nichols, S., Tetenbaum, S. P., Bregman, J., Ponzio, C. A., & Carr, E. G. (2011). Individual temperament and problem behavior in children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 26, 173–183. doi:10.1177/1088357611405041.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington: American Psychiatric Association.
- Anckarsäter, H., Stahlberg, O., Larson, T., Hakansson, C., Jutblad, S., Niklasson, L., & Rastam, M. (2006). The impact of ADHD and autism spectrum disorders on temperament, character, and personality development. *The American Journal of Psychiatry*, 163, 1239–1244. doi:10.1176/appi.ajp.163.7.1239.
- Arbuckle, J. L. (2006). *Amos (version 7.0) [computer program]*. Chicago: SPSS.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57, 289–300.
- Bradley, E. A., & Isaacs, B. J. (2006). Inattention, hyperactivity, and impulsivity in teenagers with intellectual disabilities, with and without autism. *The Canadian Journal of Psychiatry / La Revue Canadienne de Psychiatrie*, 51, 598–606.
- Brock, M. E., Freuler, A. A., Baranek, G. T., Watson, L. R., Poe, M. D., & Sabatino, A. A. (2012). Temperament and sensory features of children with autism. *Journal of Autism and Developmental Disorders*, 42, 2271–2284. doi:10.1007/s10803-012-1472-5.
- Bryson, S. E., Zwaigenbaum, L., Brian, J., Roberts, W., Szatmari, P., Rombough, V., & McDermott, C. (2007). A prospective case series of high-risk infants who developed autism. *Journal of Autism and Developmental Disorders*, 37, 12–24. doi:10.1007/s10803-006-0328-2.

- Capps, L., Kasari, C., Yirmiya, N., & Sigman, M. (1993). Parental perception of emotional expressiveness in children with autism. *Journal of Consulting and Clinical Psychology, 61*, 475–484. doi:10.1037/0022-006X.61.3.475.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. doi:10.1080/1070551070130183410.1080/107055107013018342007-12387-005.
- Chess, S., & Thomas, A. (1990). The New York longitudinal study (NYLS): the young adult periods. *The Canadian Journal of Psychiatry / La Revue Canadienne De Psychiatrie, 35*, 557–561.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.
- Clifford, S. M., Hudry, K., Elsabbagh, M., Charman, T., & Johnson, M. H. (2013). Temperament in the first 2 years of life in infants at high-risk for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*, 673–686. doi:10.1007/s10803-012-1612-y.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*, 16–29. doi:10.1037/1082-989X.1.1.16.
- del Rosario, M., Gillespie-Lynch, K., Johnson, S., Sigman, M., & Hutman, T. (2014). Parent-reported temperament trajectories among infant siblings of children with autism. *Journal of Autism and Developmental Disorders, 44*, 381–393. doi:10.1007/s10803-013-1876-x.
- DeYoung, C. G. (2006). Higher-order factors of the Big five in a multi-informant sample. *Journal of Personality and Social Psychology, 91*, 1138–1151. doi:10.1037/0022-3514.91.6.1138.
- DeYoung, C. G. (2014). Cybernetic big five theory. *Journal of Research in Personality, 48*, 10.1016/j.jrp.2014.07.004.
- Doyle, A. E., Faraone, S. V., Seidman, L. J., Willcutt, E. G., Nigg, J. T., Waldman, I. D., & ... Biederman, J. (2005). Are endophenotypes based on measures of executive functions useful for molecular genetic studies of ADHD? *Journal of Child Psychology and Psychiatry, 46*, 778–803. doi:10.1111/j.1469-7610.2005.01476.x
- Eilam-Stock, T., Xu, P., Cao, M., Gu, X., Van Dam, N. T., Anagnostou, E., & ... Fan, J. (2014). Abnormal autonomic and associated brain activities during rest in autism spectrum disorder. *Brain: A Journal of Neurology, 137*, 153–171. doi:10.1093/brain/awt294
- Evans, D. E., & Rothbart, M. K. (2009). A two-factor model of temperament. *Personality and Individual Differences, 47*, 565–570. doi:10.1016/j.paid.2009.05.010.
- Filliter, J. H., Longard, J., Lawrence, M. A., Zwaigenbaum, L., Brian, J., Garon, N., & ... Bryson, S. E. (2015). Positive affect in infant siblings of children diagnosed with autism spectrum disorder. *Journal of Abnormal Child Psychology, 43*, 567–575. doi:10.1007/s10802-014-9921-6
- Gagne, J. R., Van Hulle, C. A., Aksan, N., Essex, M. J., & Goldsmith, H. (2011). Deriving childhood temperament measures from emotion-eliciting behavioral episodes: scale construction and initial validation. *Psychological Assessment, 23*, 337–353. doi:10.1037/a0021746.
- Garon, N., Bryson, S., Zwaigenbaum, L., Smith, I.M., Brian, J., Roberts, W., & Szatmari, P. (2009). Temperament and its relationship to autistic symptoms in a high-risk infant sib cohort. *Journal of Abnormal Child Psychology, 37*, 59–78. doi:10.1007/s10802-008-9258-0.
- Goldsmith, H. H. (1996). *Toddler behavior assessment questionnaire*. Eugene: University of Oregon, Department of Psychiatry.
- Goodwin, M. S., Groden, J., Velicer, W. F., Lipsitt, L. P., Baron, M., Hofmann, S. G., & Groden, G. (2006). Cardiovascular arousal in individuals with autism. *Focus on Autism and Other Developmental Disabilities, 21*, 100–123. doi:10.1177/10883576060210020101.
- Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *The American Journal of Psychiatry, 160*, 636–645. doi:10.1176/appi.ajp.160.4.636.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. doi:10.1080/10705519909540118.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.
- Komsi, N., Rääkkönen, K., Pesonen, A., Heinonen, K., Keskiavaara, P., Järvenpää, A., & Strandberg, T. E. (2006). Continuity of temperament from infancy to middle childhood. *Infant Behavior & Development, 29*, 494–508. doi:10.1016/j.infbeh.2006.05.002.
- Konstantareas, M., & Stewart, K. (2006). Affect regulation and temperament in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 36*, 143–154. doi:10.1007/s10803-005-0051-4.
- Kushki, A., Brian, J., Dupuis, A., & Anagnostou, E. (2014). Functional autonomic nervous system profile in children with ASD. *Molecular Autism, 5*, 2040–2392.
- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guilford Press.
- Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods, 18*, 285–300. doi:10.1037/a0033266.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview—revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685. doi:10.1007/BF02172145.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. R., Leventhal, B. L., DiLavore, P. C., & ... Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205–223. doi:10.1023/A:1005592401947
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule* (2nd ed.). Torrance: Western Psychological Services.
- Loveland, K. A., Tunali-Kotoski, B., Pearson, D. A., Brelsford, K. A., Ortegon, J., & Chen, R. (1994). Imitation and expression of facial affect in autism. *Development and Psychopathology, 6*, 433–444. doi:10.1017/S0954579400006039.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*, 519–530.
- Mullen, E. M. (1995). *Mullen scales of early learning - AGS edition*. Circle Pines: AGS Publishing.
- 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.
- Osborne, J. W. (2010). Improving your data transformations: applying Box-Cox transformations as a best practice. *Practical Assessment, Research, and Evaluation, 15*, 1–9.
- Ozonoff, S., Garcia, N., Clark, E., & Lainhart, J. E. (2005). MMPI-2 personality profiles of high-functioning adults with autism spectrum disorders. *Assessment, 12*, 86–95. doi:10.1177/1073191104273132.
- Porges, S. W. (2005). The vagus: A mediator of behavioral and visceral features associated with autism. In M. L. Bauman & T. L. Kemper (Eds.), *The neurobiology of autism* (pp. 65–78). Baltimore: Johns Hopkins University Press.
- Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2012). Control networks and neuromodulators of early development. *Developmental Psychology, 48*, 827–835. doi:10.1037/a0025530.
- Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory

- activity. *International Journal of Psychophysiology*, 61, 5–18. doi:10.1016/j.ijpsycho.2005.10.024.
- Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, 52, 569–578. doi:10.2307/1129176.
- Rothbart, M. K. (2011). *Becoming who we are: Temperament and personality in development*. New York: Guilford Press.
- Rothbart, M. K., & Bates, J. E. (2006). Temperament. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 3, social, emotional, and personality development* (6th ed., pp. 99–166). Hoboken: Wiley.
- Rothbart, M. K., & Posner, M. I. (2006). Temperament, attention, and developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol 2: Developmental neuroscience* (2nd ed., pp. 465–501). Hoboken: Wiley.
- Rothbart, M. K., Ellis, L. K., Rueda, M., & Posner, M. I. (2003). Developing mechanisms of temperamental effortful control. *Journal of Personality*, 71, 1113–1143. doi:10.1111/1467-6494.7106009.
- Shiner, R. L., & DeYoung, C. G. (2013). The structure of temperament and personality traits: A developmental perspective. In P. D. Zelazo & P. D. Zelazo (Eds.), *The Oxford handbook of developmental psychology* (Self and other, Vol. 2, pp. 113–141). New York: Oxford University Press.
- Soderstrom, H., Rastam, M., & Gillberg, C. (2002). Temperament and character in adults with Asperger syndrome. *Autism*, 6(3), 287–297. doi:10.1177/1362361302006003006.
- Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neuroscience and Biobehavioral Reviews*, 36, 747–756. doi:10.1016/j.neubiorev.2011.11.009.
- Wainer, A. L., Ingersoll, B. R., & Hopwood, C. J. (2011). The structure and nature of the broader autism phenotype in a non-clinical sample. *Journal of Psychopathology and Behavioral Assessment*, 33, 459–469. doi:10.1007/s10862-011-9259-0.
- Whittle, S., Allen, N. B., Lubman, D. I., & Yücel, M. (2006). The neurobiological basis of temperament: towards a better understanding of psychopathology. *Neuroscience and Biobehavioral Reviews*, 30, 511–525. doi:10.1016/j.neubiorev.2005.09.003.
- Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, S. G. West, K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington: American Psychological Association. doi:10.1037/10222-009.
- Wiggins, L. D., Reynolds, A., Rice, C. E., Moody, E. J., Bernal, P., Blaskey, L., & ... Levy, S. E. (2014). Using standardized diagnostic instruments to classify children with autism in the study to explore early development. *Journal of Autism and Developmental Disorders*, 10.1007/s10803-014-2287-3
- Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, 9, 11–29. doi:10.1007/BF01531288.
- Zwaigenbaum, L., Bryson, S. E., Szatmari, P., Brian, J., Smith, I. M., Roberts, W., & ... Roncadin, C. (2012). Sex differences in children with autism spectrum disorder identified within a high-risk infant cohort. *Journal of Autism and Developmental Disorders*, 42, 2585–2596. 10.1007/s10803-012-1515-y