The Temperament Features Associated with Autism Spectrum Disorder in Childhood: A Systematic Review

Carly A. Mallise[a,b], Alison E. Lane[b,c], Alix J. Woolard[a,b], Olivia M. Whalen[a,b], Vanessa E. Murphy[b,d], Frini Karayanidis[a,e] & Linda E. Campbell[a,b*]

[a] School of Psychology, University of Newcastle, Callaghan, Australia

[b] Priority Research Centre GrowUpWell®, University of Newcastle, Australia

[c] School of Health Sciences, University of Newcastle, Callaghan, Australia

[d] VIVA (Viruses, Infections/Immunity, Vaccines and Asthma) Research Group, Hunter Medical Research Institute, New Lambton Heights, Australia

[e] Priority Research Centre for Stroke and Brain Injury, University of Newcastle, Australia

Note: This version has not undergone copyediting by *Research in Developmental Disorders*

*Corresponding Author:

Linda E. Campbell

The University of Newcastle,

Ourimbah NSW 2258

T: +61 2 434 94490, F: +61 2 434 94404

E: linda.e.campbell@newcastle.edu.au

Abstract

Background: Temperament is an important construct that shapes child development. Temperament is suggested to present differently in different groups, such as children with neurodevelopmental disorders. However, it is not known whether there are specific temperament features associated with Autism Spectrum Disorder (ASD).

Aim: This systematic review aimed to synthesise extant literature to determine whether there are temperament features associated with ASD in infancy, toddlerhood and childhood.

Methods and Procedures: Following the PRISMA guidelines for systematic reviews, we searched PsycINFO, CINAHL, Academic Search Ultimate and ProQuest for all available articles from database conception until January 2020. The Joanna Briggs Institute Critical Appraisal checklists were used to assess the methodological quality of included articles.

Outcomes and Results: Twenty-six articles met the selection criteria: (1) reported on the temperament of children (0-12 years of age) diagnosed with ASD, (2) peer-reviewed; and (3) published in English. Articles varied in overall methodological quality. Infants later diagnosed with ASD were found to more frequently being described as having 'easy' temperament features in early infancy, compared to typically developing infants and infants with developmental concerns but not ASD. Once diagnosed, children with ASD were reported to, as a group, display more negative affect, less extraversion and less effortful control than typically developing children.

Conclusions and Implications: The literature suggests that more challenging temperament features are associated with ASD in childhood, but less is known about within group variability. Overall, this review highlights the need for further investigation into the variability of temperament in children with ASD.

Keywords: Temperament, Autism Spectrum Disorder, Individual Differences, Infancy, Childhood, Systematic Review

What this paper adds?

This is the first systematic review to synthesise the literature on the temperament features of children with Autism Spectrum Disorder, prior to and post diagnosis. The findings suggest that some temperament features are associated with ASD in childhood, however, the evidence is weaker prior to diagnosis. Infants later diagnosed with ASD were rated as having 'easier' temperament features in early infancy, as compared to infants without ASD, with more challenging behaviours observed closer to time of diagnosis. Parents of children diagnosed with ASD reported that their children displayed more negative affect, less extraversion, and less effortful control than typically developing peers. This review highlights several gaps in the literature. All the pre-diagnosis studies reported here examined the temperament on infant siblings of children diagnosed with ASD. Future research needs to examine temperament features in other infant cohorts at-risk for ASD. Additionally, only two studies examined variability of temperament features in children with ASD, and both found great heterogeneity among their samples. This suggests that temperament may be a suitable construct for identifying clinically meaningful individual differences within ASD. Therefore, further investigation is needed into the variability of temperament features in children with ASD, pre- and post-diagnosis.

1. Introduction

The study of individual differences is an important approach in understanding human behaviour and development (Kanai & Rees, 2011). A central construct contributing to individual differences is temperament. Temperament has been defined as the observable, individual differences in behavioural style that appears early in life (Thomas, Chess, Birch, Hertzig, & Korn, 1963). There are several conceptual frameworks found in the literature that differ on the level of heritability and continuity of temperament (for further details, see review in Zentner & Bates, 2008). However, all frameworks present temperament as a construct comprised of separate domains, or features, which are bidirectional in nature. Temperament features can be described as 'easy' or 'difficult', depending on whether an individual falls within an extremity of the bidirectional spectrum. These features, taken together, form a temperament profile, which describes an individual's overall behavioural style (Carey, 1970). A temperament profile characterised by positive mood, high approach, quick adaptability to change, predictability of responses and low distractibility is referred to as 'easy', whereas the temperament profile comprised of the opposite features is referred to as 'difficult' (Carey, 1970).

Temperament plays an important role in child development, as it can shape learning (Gartstein, Putnam, & Kliewer, 2016; Studer-Luethi, Bauer, & Perrig, 2016), social functioning (Baer et al., 2015) and attachment (Groh et al., 2017). Further, temperament is closely associated with individual differences in children's responses to stressors, with some children being more likely to display internalised behaviours (e.g., inhibited and over-controlled) while others act out or externalise their behaviours (e.g., aggression, hyperactivity; Atherton, Tackett, Ferrer, & Robins, 2017; Davis, Votruba-Drzal, & Silk, 2015). These behaviours, in turn, can determine the nature of opportunities to learn and socialise within in the home and school environments (Douglas & David, 2005; Hymel,

Rubin, Rowden, & LeMare, 1990). While problematic internalising and externalising behaviours are often apparent in neurotypical children, they are more prominent in clinical populations, such as children with developmental disorders (Bauminger, Solomon, & Rogers, 2010; Mazurek & Kanne, 2010; Volker et al., 2010). However, the severity of such behaviours varies considerably in children with developmental disorders (Vaillancourt et al., 2017), indicating substantial heterogeneity within these populations.

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterised by atypical social communication and interaction, and the presence of restricted, repetitive behaviours, interests or activities (American Psychiatric Association, 2013). Atypical social communication and interaction refers to difficulty with social-emotional reciprocity, nonverbal communicative behaviours, and the development and maintenance of relationships. Restricted and repetitive behaviours can include stereotyped motor movements, insistence on sameness, fixated interests that are atypical in intensity or focus, and hyper- or hypo-reactivity to sensory input. While individuals diagnosed with ASD collectively meet these criteria, functionally, there is a diverse range in the manifestation of symptoms, and their impact on social, behavioural and daily living skills (Weitlauf, Gotham, Vehorn, & Warren, 2014). As a result, there is no intervention approach that suits all individuals diagnosed with ASD. Consequently, researchers in the autism field are seeking methods by which individual differences within ASD can be identified to promote person-centred interventions (U.S. Department of Health and Human Services Interagency Autism Coordinating Committee, 2017).

Temperament is one construct that could help explain why some children with ASD develop emotional, social and behavioural problems, while others do not. Additionally, as temperament provides information on an individual's behavioural style (Iverson & Gartstein, 2018), understanding its variance in ASD may assist caregivers and educators with

anticipating and managing aspects of the social and physical environment that may not fit an individual child's needs. Currently, the assessment of ASD is not primarily focused on identifying potential emotional, social and behavioural problems (Zwaigenbaum & Penner, 2018). However, as part of the assessment, most clinicians use a range of different tools that can be useful to highlight the characteristics of the person. Assessments of temperament, in conjunction with gold standard diagnostic tools, may assist with identifying which intervention would work best for a child, based on how they react to their environment. Little is known, however, about whether there are specific temperament features associated with ASD that could be used to identify individual differences. The purpose of this systematic review is to synthesise the existing literature in order to examine whether there are specific features of temperament associated with ASD, prior to and post diagnosis, in infancy, toddlerhood and childhood. Additionally, this systematic review aims to provide a foundation for further research exploring the temperament of children with ASD by identifying the gaps in the literature.

2. Methods

A systematic review of the literature was conducted using the statement on Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher, Liberati, Tetzlaff, Altman, & Prisma Group, 2009). This review was not registered, and no review protocol was published.

2.1 Search Strategy

A literature search was conducted across six social science and allied health databases: PsycINFO, CINAHL, Academic Search Ultimate, and the Psychology, Nursing & Allied Health, and Social Science databases in ProQuest. The literature was searched by one independent investigator (CM) for all articles published from database inception until January 2020, using search terms related to temperament and ASD in infant and child populations

(Table S1). A manual search identified an additional 70 articles from the reference list of eligible articles and relevant review articles.

2.2 Inclusion and Exclusion Criteria

Articles were screened for eligibility against the selection criteria by three independent investigators (CM, AW, and OW). A fourth independent investigator (LC) resolved conflicts in eligibility status. Articles were included if they (1) reported on the temperament of individuals diagnosed with ASD either (a) post diagnosis, in childhood (3-12 years of age) or (b) prior to diagnosis, in infancy (< 3 years of age); (2) were peer-reviewed; and (3) were published in English. There were no restrictions for inclusion based on study design. Articles were excluded if they (1) reported on adolescents (>13 years) or adults with ASD; (2) included a sample with a combined age range (e.g. 10-16 years) and data could not be extracted for children (< 13 years); (3) used a measure of personality (e.g., Big Five Inventory) rather than a measure of temperament (e.g., Carey Temperament Scales); (4) reported on children with ASD as a co-morbid condition (e.g. Fragile X with ASD); and (5) were a review, conference paper, book, thesis or grey literature.

2.3 Data Extraction and Risk of Bias Assessment

One investigator (CM) independently extracted the following information: study design, country of recruitment, sample size, gender distribution, age of sample (mean and/or range), method of diagnosis, and temperament measure. This information was subsequently checked by another independent investigator (LC, OW or AW). Where possible, we extracted statistically significant mean differences in temperament outcomes between infants later diagnosed with ASD or children with ASD, and typically developing comparison groups. The results are synthesised and discussed in two clusters: (1) Those pertaining to the temperament of infants later diagnosed with ASD (i.e. pre-diagnosis) and (2) those pertaining to the temperament of children diagnosed with ASD (i.e. post-diagnosis). The Joanna Briggs

Institute Critical Appraisal (JBI-CA) checklists were used by one reviewer (CM) to critically assess risk of bias, or methodological quality, of the included studies. All articles were included in the review, regardless of quality assessment outcome. A meta-analysis was not feasible within this review. Rather, we addressed trends descriptively, taking the heterogeneity of reported temperament outcomes into account.

3. Results

The study selection flowchart is presented in Figure 1. The search strategy generated 931 unique articles. During title and abstract screening, 811 articles were excluded (primarily as they did not pertain to ASD or temperament). There were 83 conflicts that occurred during the title and abstract screening stage, which were resolved by an independent investigator (LC). During full-text screening, 94 articles were excluded (see Figure 1 for exclusion reasons). There were 21 conflicts that occurred during the full-text screening stage, which were resolved by an independent investigator (LC). Twenty-six articles met eligibility criteria for inclusion.

3.1 Study Characteristics

The study characteristics of the 26 articles are summarised in Tables 1 and 2. These articles report temperament findings from 18 study cohorts. Ten reported on the temperament of infant siblings of children diagnosed with ASD, who are at high-risk for developing ASD. The other 16 articles reported on the temperament of children diagnosed with ASD. Most articles were published after the year 2010 (69.2%), with 13 articles published within five years prior to the search date. The studies included case-control (n = 9), cohort (n = 11) and cross-sectional (n = 5) designs, with only one case series. The majority of studies were conducted within North America (69.2%). The remaining were conducted in the United Kingdom (n = 3), Israel (n =2), Taiwan (n = 1), the Netherlands (n = 1) and Australia (n = 1). The sample size ranged from nine to 68,197 participants. Gender was unequally distributed,

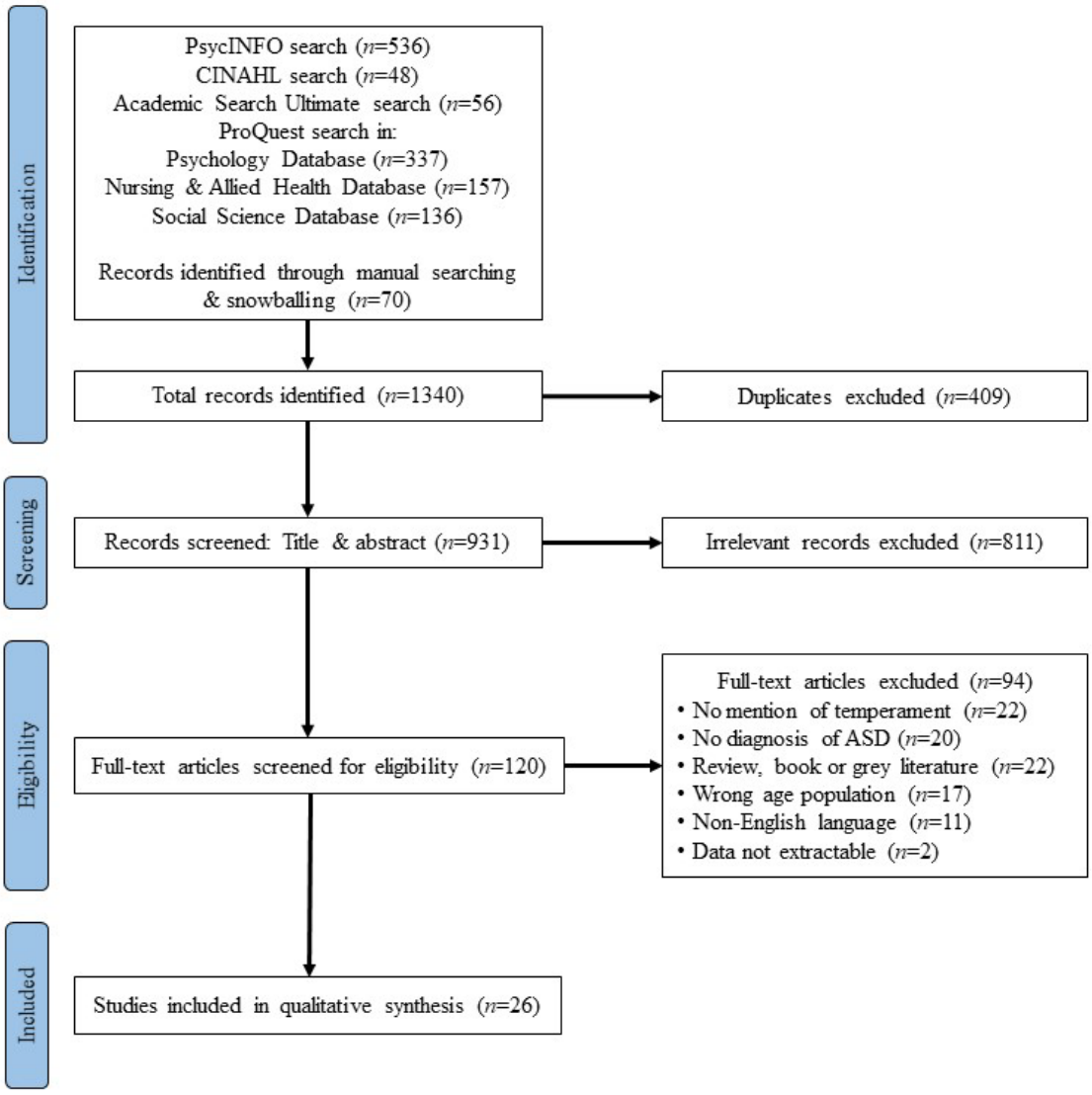

Figure 1. PRISMA flow chart of study selection.

with more than half of the articles reporting 80% or more males in their sample. The majority of articles utilised parent-report temperament measures, with only two studies using a laboratory measure.

3.2 Risk of Bias of Included Studies

Table 3 reports the percentage of JBI-CA items that met methodological requirements within each article. The JBI-CA checklists containing each item can be found on the JBI website (https://joannabriggs.org/). All 26 articles used appropriate statistical analyses of the

data. Of the cross-sectional studies, four out of five described participants and setting in sufficient detail (e.g., demographic variables, recruitment method), and used a valid and reliable measure of temperament (Adamek et al., 2011; Bailey et al., 2000; Brock et al., 2012; Hepburn & Stone, 2006).

All nine case-control studies used a valid and reliable assessment of ASD (Barger et al., 2019; Chuang et al., 2012; Hirschler-Guttenberg et al., 2015; Kasari & Sigman, 1997; Konstantareas & Stewart, 2006; Macari et al., 2017; Macari et al., 2018; Ostfeld-Etzion et al., 2016; Reyes et al., 2019). Five of the nine case-control studies appropriately matched cases with controls on appropriate confounds, including gender (Hirschler-Guttenberg et al., 2015; Kasari & Sigman, 1997; Konstantareas and Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016), chronological age (Konstantareas & Stewart, 2006), mental age (Hirschler-Guttenberg et al., 2015; Ostfeld-Etzion et al., 2016), ethnicity (Kasari & Sigman, 1997) and/or family demographics (Hirschler-Guttenberg et al., 2015; Kasari & Sigman, 1997; Ostfeld-Etzion et al., 2016). In addition, Macari et al. (2017) identified developmental quotient as a confound and included it as a covariate in their statistical analyses. However, four of the nine case-control studies did not match on gender (Barger et al., 2019; Chuang et al., 2012; Macari et al., 2018; Reyes et al., 2019). Chuang et al. (2012), Macari et al. (2018) and Reyes et al. (2019) only matched participants on age (i.e., chronological or mental). However, Macari et al. (2018) and Reyes et al. (2019) identified gender as a confounder and included it as a covariate in their statistical analyses. Further, Barger et al. (2015) did not match participants on any appropriate confounds. Regarding the assessment of temperament, eight of the nine case-control studies used valid and reliable tools (Barger et al., 2019; Chuang et al., 2012; Hirschler-Guttenberg et al., 2015; Kasari & Sigman, 1997; Konstantareas & Stewart, 2006; Macari et al., 2018; Ostfeld-Etzion et al., 2016; Reyes et al., 2019). It was

unclear whether the temperament measure used by Macari et al. (2017) was valid for their participants' age range.

All cohort studies, except for Bischof et al. (2018) and Clifford et al. (2013), reported using valid and reliable measures of ASD and temperament (Bryson et al., 2017; Del Rosario et al., 2014; Garon et al., 2009; Garon et al., 2016; Hendry et al., 2018; Øien et al., 2018; Paterson et al., 2019; Pijl et al., 2019; Zwaigenbaum et al., 2005). Seven of the 11 cohort studies reported that groups were from similar populations (Bischof et al., 2018; Bryson et al., 2017; Del Rosario et al., 2014; Garon et al., 2009; Garon et al., 2016; Paterson et al., 2019; Zwaigenbaum et al., 2005). Additionally, nine of the 11 cohort studies identified confounding factors, including gender (Bryson et al., 2017; Del Rosario et al., 2014; Garon et al., 2009; Garon et al., 2016; Øien et al., 2018; Paterson et al., 2019; Pijl et al., 2019; Zwaigenbaum et al., 2005), chronological age (Bryson et al., 2017; Del Rosario et al., 2014; Pijl et al., 2019; Zwaigenbaum et al., 2005), birth order (Bryson et al., 2017; Zwaigenbaum et al., 2005), birth weight (Bryson et al., 2017), gestational age (Bryson et al., 2017; Hendry., 2018; Pijl et al., 2019) and intelligence/developmental quotient (Garon et al., 2009; Paterson et al., 2019). These covariates were addressed by either matching participants on the factors (Bryson et al., 2017; Zwaigenbaum et al., 2005), excluding participants based on the factors (Bryson et al., 2017; Hendry., 2018; Pijl et al., 2019) or including the factors as covariates in analyses (Del Rosario et al., 2014; Garon et al., 2009; Garon et al., 2016; Øien et al., 2018; Paterson et al., 2019; Pijl et al., 2019). The case series met nine out of ten JBI-CA criteria (Bryson et al., 2007), with the item not met pertaining to a lack of clearly reported temperament outcomes.

Some methodological concerns, however, were identified. Of the cross-sectional studies, three of the five did not clearly state how ASD was assessed (Adamek et al., 2011; Bagnato & Neisworth, 1999; Hepburn & Stone, 2006). Further, four of the five cross-

sectional studies did not identify or control for any of the main confounding factors, including gender and age (Adamek et al., 2011; Bagnato & Neisworth, 1999; Brock et al., 2012; Hepburn & Stone, 2006). Seven of the nine case-control studies did not demonstrate that the same eligibility criteria were used for cases and controls (Chuang et al., 2012; Hirschler-Guttenberg et al., 2015; Kasari & Sigman, 1997; Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016; Reyes et al., 2019), or that ASD was assessed in the same manner across groups (Barger et al., 2019; Chuang et al., 2012; Hirschler-Guttenberg et al., 2015; Kasari & Sigman, 1997; Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016). Lastly, four of the 10 cohort studies did not clearly report on whether the groups were from the same population (Clifford et al., 2013; Hendry et al., 2018; Øien et al., 2018; Pijl et al., 2019). Clifford et al. (2013) clearly reported on only two of the seven JBI-CA items, which pertained to using a similar method of assessing ASD to assign participants into the groups and appropriate statistical analyses. Additionally, it was unclear whether Clifford et al. (2013) used valid and reliable methods of assessing ASD and temperament, and the authors did not identify or address any confounding factors (e.g., age, gender) within their study.

3.3 Associations between Temperament and Autism Spectrum Disorder Pre-Diagnosis

Ten articles reported on the temperament characteristics of infant siblings (hereafter, infant-sibs) of children diagnosed with ASD (Table 4; Bryson et al., 2007; Bryson et al., 2017; Clifford, Hudry, Elsabbagh, Charman, Johnson, et al., 2013; Del Rosario, Gillespie-Lynch, Johnson, Sigman, & Hutman, 2014; Garon et al., 2009; Garon et al., 2016; Hendry et al., 2018; Paterson et al., 2019; Pijl et al., 2019; Zwaigenbaum et al., 2005). These articles compared the temperament of infant-sibs later diagnosed with ASD, prior to diagnosis, to typically developing controls or infant-sibs without ASD. Within these articles, infant-sibs and typically developing controls had been assessed for ASD by 36 months of age. Infant-

sibs who received a diagnosis of ASD will be referred to as *infant-sibs with ASD,* with age of diagnosis reported afterwards in brackets. Infant-sibs who did not receive a diagnosis of ASD, but may have had some developmental concerns, will be referred to as *infant-sibs without ASD*. Infant-sibs who proceeded to develop typically will be referred to as *typically developing infant-sibs* in order to differentiate them from *typically developing controls* (i.e., infants with no family history of ASD).

Results will be synthesised in groups by age: early infancy, later infancy and toddlerhood. Additionally, the three articles (Del Rosario et al., 2014; Paterson et al., 2019; Pijl et al., 2019) that reported on longitudinal trajectories of temperament will be synthesised in a fourth section. Five of the infant-sib articles (Bryson et al., 2007; Bryson et al., 2017; Garon et al., 2009; Garon et al., 2016; Zwaigenbaum et al., 2005) reported data from the same Canadian cohort study and three other articles (Clifford et al., 2013; Hendry et al., 2018; Pijl et al., 2019) reported data from the same UK cohort study. As such, results from articles based on the same cohort data will be considered together with the exception of Bryson et al. (2007), who did not describe temperament outcomes in a way that enabled them to be included in the synthesis (i.e., utilising temperament domain names).

3.3.1 Temperament Profile: Early Infancy (6 – 8 months)

Six articles reported on the temperament of infant-sibs later diagnosed with ASD in early infancy, at six (Bryson et al., 2017; Del Rosario et al., 2014; Paterson et al., 2019; Zwaigenbaum et al., 2005), seven (Clifford, Hudry, Elsabbagh, Charman, Johnson, et al., 2013) or eight months of age (Pijl et al., 2019). Group differences emerged in the temperament domains of *activity, approach, adaptability, effortful control* and *negative affect* (from 19 domains assessed; Table 5) between infants with and without a later diagnosis of ASD. Zwaigenbaum et al. (2005) found that 6-month infant-sibs with ASD (at 24 months), had lower levels of activity compared to 6-month typically developing controls and 6-month

infant-sibs without ASD. Using the same cohort, Bryson et al. (2017) provide descriptive comparisons (i.e., no statistical comparisons were conducted) and reported that 6-month infant-sibs with ASD (at 36 months) demonstrated more frustration, more fear, more smiling and laughter, a higher activity level, a shorter duration of visual orientation, and were more easily soothed than typical developing 6-month infants.

Unlike Zwaigenbaum et al. (2005), Del Rosario et al. (2014) found no differences in activity level between 6-month infant-sibs with ASD (at 36 months) and 6-month typically developing infant-sibs. Differences in findings between these two articles may be explained by the use of different temperament questionnaires. Del Rosario et al. (2-14) found, however, that 6-month infant-sibs with ASD (at 36 months) were less withdrawn and more adaptive to change, compared to infants who proceeded to develop typically. Conversely, Clifford et al. (2013) found 7-month infant-sibs with ASD (at 36 months) were more withdrawn than infants who proceeded to develop typically. This may be attributed to the two studies again using different measures of temperament. However, it is also possible that discrepancies were due to the lack of controlling for confounders in Clifford et al. (2013). Lastly, Paterson et al. (2019) and Pijl et al. (2019) found that infant-sibs with ASD (at 24 and 36 months, respectively) had poorer effortful control at 6 months and displayed more negative affect at 8 months, respectively, than typically developing controls. Overall, findings are mixed; some articles reported that infant-sibs with ASD had more manageable temperament features in early infancy (i.e., lower activity levels, less withdrawal, more adaptability; Del Rosario et al., 2014; Zwaigenbaum et al., 2005), while others reported on temperament features that are considered to be more challenging (e.g., more withdrawal, more negative affect, less effortful control; Clifford et al., 2013; Paterson et al., 2019; Pijl et al., 2019). As a result, there is limited evidence to suggest that children with ASD have specific temperament features prior to diagnosis in early infancy.

3.3.2 Temperament Profile: Late Infancy (12 months – 14 months)

Seven articles (four cohorts) reported on the temperament of infant-sibs with a later diagnosis of ASD in late infancy, at 12 (Bryson et al., 2017; Del Rosario et al., 2014; Garon et al., 2016; Paterson et al., 2019; Zwaigenbaum et al., 2005) or 14 months of age (Clifford, Hudry, Elsabbagh, Charman, Johnson, et al., 2013; Pijl et al., 2019). Together, they assessed 20 temperament domains, with differences emerging in eight: *distress to limitations, cuddliness, duration of orienting, smiling and laughter, positive affect, negative affect, surgency,* and *effortful control*. Zwaigenbaum et al. (2005) found 12-month infants with ASD (at 24 months) demonstrated more distress to limitations (i.e., frustration) and a longer duration of visual orientation towards objects than 12-month typically developing controls and 12-month infant-sibs without ASD. Within the same cohort, Garon et al. (2016) reported that 12-month infants with ASD (at 36 months) showed less *positive affect* (discriminant function comprised of *smiling and laughter*, *soothability* and reversed *fear*) than infant-sibs without ASD. Bryson et al. (2017), again using the same cohort, descriptively found that 12-month infant-sibs with ASD (at 36 months) demonstrated more frustration, more fear, less smiling and laughter, a higher activity level, a shorter duration of visual orientation, and were more easily soothed than typical developing 6-month infants.

Del Rosario et al. (2014) reported that 12-month infant-sibs with ASD (at 36 months) were more adaptive to change in routine, compared to typically developing infant-sibs. Clifford et al. (2013) found that 14-month infant-sibs with ASD (at 36 months) were lower in *cuddliness* and *smiling and laughter* than typically developing controls. Both Paterson et al. (2019) and Pijl et al. (2019) reported that infant-sibs with ASD (at 24 and 36 months, respectively) displayed more *negative affect*, less *surgency* (i.e., extraversion) and less *effortful control* than typically developing controls at 12 months and 14 months, respectively.

Overall, findings were mixed, with no more than two articles reporting the same results (i.e., Paterson et al, 2019; Pijl et al., 2019). However, most articles solely reported that infant-sibs with ASD in later childhood displayed temperament features that are considered to be more challenging, compared to controls and infant-sibs without ASD. Despite this, the findings from the current synthesis suggest that children with ASD have specific temperament features prior to diagnosis in late infancy is limited.

3.3.3 Temperament Profile: Toddlerhood (24 months – 36 months)

Seven articles (three cohorts) reported on the temperament of infant-sibs later diagnosed with ASD during toddlerhood, at 24 (Clifford, Hudry, Elsabbagh, Charman, Johnson, et al., 2013; Del Rosario et al., 2014; Garon et al., 2009; Garon et al., 2016; Pijl et al., 2019; Zwaigenbaum et al., 2005) or 36 months of age (Del Rosario et al., 2014; Hendry et al., 2018). Differences emerged in 12 out of 17 temperament domains assessed (six of which were not assessed earlier) between infant-sibs with ASD, and typically developing controls and infant-sibs without ASD: *soothability*, *sadness*, *cuddliness, shyness, approach, inhibitory control, low intensity pleasure, adaptability, attentional shifting, positive anticipation, effortful control and negative affect***.** Zwaigenbaum et al. (2005) found that 24-month infant-sibs with ASD (at 24 months) were reported to have less inhibitory control, a lower ability to shift attention and display less excitement about pleasurable activities than typically developing controls and infant-sibs without ASD. Within the same prospective infant-sib cohort, the researchers later found that 24-month infant-sibs with ASD (at 36 months) displayed less behavioural approach and less effortful control compared to typically developing controls (Garon et al., 2009), and less positive affect than infant-sibs without ASD (Garon et al., 2016).

Clifford, Hudry, Elsabbagh, Charman, Johnson, et al. (2013) found that 24-month infant-sibs with ASD (at 36 months) were harder to soothe, displayed more sadness, sought

less physical comfort, were shyer, and participated less in low intensity activities than typically developing controls. In the same cohort, it was also found that infant-sibs with ASD (at 36 months) displayed more negative affect at 24 months (Pijl et al., 2019), and less effortful control at 24 and 36 months (Hendry et al., 2018; Pijl et al., 2019). Lastly, Del Rosario et al. (2014) found that infant-sibs with ASD (at 36 months) were more withdrawn and less adaptable to change than typically developing infant-sibs at 24 and 36 months of age.

Overall, the articles reported that infant-sibs with ASD in later childhood had temperament features that were more challenging, compared to controls or infant-sibs without ASD. Specifically, three articles, across two cohorts, reported that infant-sibs with ASD in later childhood had less effortful control than typically developing controls (Garon et al., 2009; Hendry et al., 2018; Pijl et al., 2019). However, most articles did not report differences between groups in the same temperament domains. As a result, there is limited evidence suggesting that children with ASD have specific temperament features prior to diagnosis in toddlerhood.

3.3.4 Longitudinal Trajectories

Three articles reported on statistical analyses examining the developmental trajectories of temperament features of infant-sibs (ASD, atypical and typical) and typically developing controls (Del Rosario et al., 2014; Paterson et al., 2019; Pijl et al., 2019). The earliest article reported that infant-sibs with ASD (at 36 months) differed from typically developing infant-sibs in their trajectory of *activity*, *adaptability* and *approach* (Del Rosario et al., 2014). From 6 months to 36 months, infant-sibs with ASD (at 36 months) became more active, more withdrawn and less adaptable with change over time, whereas typically developing infant-sibs were stable in their level of activity and approach, and more adaptable with change over time.

The latter two articles examined the trajectories of *surgency*, *negative affect* and *effortful control* (Paterson et al., 2019; Pijl et al., 2019). Paterson et al. (2019) found a difference in the trajectory of surgency from 6 months to 24 months between infant-sibs with ASD (at 24 months), and infant-sibs without ASD and typically developing controls. Infant-sibs with ASD (at 24 months) initially lowered in their level of surgency from 6 months to 12 months but then increased from 12 months to 24 months. Typically developing controls displayed the opposite trajectory while infant-sibs without ASD were stable. No differences in the trajectory of negative affect and effortful control were found, as all groups were stable over time. Pijl et al. (2019) alternatively found differences in the trajectories of *surgency* and *effortful control* between infant-sibs with ASD (at 36 months), and typically developing infant-sibs and typically developing controls. Infant-sibs with ASD (at 36 months) decreased in their level of surgency and effortful control from 8 months to 14 months and remained stable from 14 months to 24 months. Alternatively, typically developing infant-sibs were stable in surgency, and controls surgency and effortful control, across time. Lastly, Pijl et al. (2019) found that all groups were stable over time for negative affect. Overall, there are few articles that reported on the trajectory of temperament features in infant-sibs with later ASD. Emerging research does indicate, however, that infant-sibs with later ASD have less stable temperament features, compared to infant-sibs without ASD and typically developing controls.

3.4 Associations between Temperament and Autism Spectrum Disorder Post-Diagnosis

Sixteen articles reported on the temperament of children diagnosed with ASD (mean age between 2-6 years; Table 4). Of these, two descriptively characterised the temperament of children with ASD, three compared temperament scores to normative reference samples, and 11 compared temperament scores to typically developing or general population controls. For ease of synthesis, results are grouped by three overarching temperament factors: (1)

Negative Affectivity (domains pertaining to negative affect, e.g., sadness, fear), (2) Extraversion/Surgency (domains pertaining to positive affect, e.g., activity level, approach), and (3) Effortful Control (domains pertaining to attention and inhibition, e.g., inhibitory control and distractibility). Domains that do not fall within these factors will be discussed in a fourth section. The definitions of the temperament domains can be found in Table 5. All differences included in the synthesis were reported as statistically significant. For studies where statistical comparisons were not conducted (Bagnato & Neisworth, 1999; Hepburn & Stone, 2006), findings are reported descriptively.

3.4.1 Negative Affectivity

Four articles reported on the domain of *Soothability* or *Falling Reactivity*, or how easily the child can be soothed after a peak in reactivity. Three articles (Konstantareas & Stewart, 2006; Macari, Koller, Campbell, & Chawarska, 2017; Ostfeld-Etzion, Feldman, Hirschler-Guttenberg, Laor, & Golan, 2016) reported that children with ASD were harder to soothe than typically developing controls, while one article found no difference compared to a normative reference group (Adamek et al., 2011). The discrepancy found here can be potentially explained by the lack of controlling for confounders in the cross-sectional study (i.e., Adamek et al., 2011). Six articles (four case-control, one cohort, one cross-sectional) reported on the *anger* and *frustration* of children with ASD across four cohorts (Adamek et al., 2011; Hirschler-Guttenberg, Feldman, Ostfeld-Etzion, Laor, & Golan, 2015; Konstantareas & Stewart, 2006; Macari et al., 2017; Macari et al., 2018; Ostfeld-Etzion et al., 2016). None of the case-control or cohort studies observed differences between groups (Hirschler-Guttenberg et al., 2015; Konstantareas & Stewart, 2006; Macari et al., 2017; Macari et al., 2018; Ostfeld-Etzion et al., 2016). Only one article (Adamek et al., 2011) reported that their sample of children with ASD had a higher level of anger, compared to a normative reference group. Inconsistencies in results found within the *anger* domain may be

attributed to the lack of controlling for confounders in the cross-sectional study (i.e., Adamek et al., 2011).

Four articles reported on the level of *discomfort* of children with ASD, with two finding no significant difference between children with ASD and typically developing controls (Macari et al., 2017; Ostfeld-Etzion et al., 2016). One study (Konstantareas & Stewart, 2006) found children with ASD were reported to display a higher level of discomfort than typically developing controls. In contrast, Adamek et al. (2011) found children with ASD were reported to exhibit less discomfort than the normative reference group. These inconsistences in results between articles that assessed *discomfort* may be due to varying methodological quality (e.g., controlling vs not controlling for confounders) or the use of different temperament measures. Of note, only one article reported differences related to *fear*, whereby children with ASD displayed a milder intensity of fear than typically developing children on the lab-based measure (Macari et al., 2018). The other four articles that examined fear found no differences (Hirschler-Guttenberg et al., 2015; Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016). However, three of these articles used a parent-report measure (Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016). Further, the other article that used a lab-based measure only reported findings on a single item, rather than an average of the entire battery (Hirschler-Guttenberg et al., 2015).

Of the four articles that reported on *mood* (Bailey, Hatton, Mesibov, Ament, & Skinner, 2000; Brock et al., 2012; Chuang, Tseng, Lu, & Shieh, 2012; Reyes et al., 2019), only one found a difference; children with ASD were rated as more negative in mood than typically developing children (Reyes et al., 2019). The inconsistencies between articles in findings may be the result of varying methodological quality, due to study design (e.g., controlling vs not controlling for confounders). Further, differences between Chuang et al. (2012) and Reyes et al., 2019 may be cultural, as two different populations of children with

ASD were sampled (i.e., Taiwan and the United States). However, Hepburn and Stone (2006) found that whilst over half of their sample fell within the average range for mood, one third were reported to be primarily negative in mood (i.e. scores > 1 standard deviation from normative mean). Similarly, Chuang et al. (2012) reported that nearly one fifth of their ASD sample scored within the difficult range for mood. Kasari and Sigman (1997) constructed a difficultness score using five domains (Rhythmicity, Approach, Adaptability, Intensity and Mood) from the Carey Temperament Scales and found that children with ASD were more temperamentally difficult compared to typically developing controls. One case-control study reported on *crying*, with children with ASD reported to have a higher tendency to cry than typically developing controls (Barger et al., 2019). Lastly, no articles reported significant differences related to *sadness* (Adamek et al., 2011; Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016).

Overall, eight domains related to negative affectivity were examined across 13 articles. Results were mixed for four domains (i.e., anger/frustration, discomfort, fear & mood) and two domains were each only examined in one article (i.e., difficultness & crying). This shows that there is limited evidence for differences in temperament features related to negative affectivity. However, there is some evidence to suggest that children with ASD are more difficult to soothe compared to typically developing children, yet there are no differences in sadness.

3.4.2 Extraversion/Surgency

Eleven of the 16 articles reported on the domain of *activity level*, with seven finding no difference between children with ASD and typically developing controls (Barger et al., 2019; Konstantareas & Stewart, 2006; Macari et al., 2017; Øien et al., 2018; Reyes et al., 2019) or normative reference groups (Adamek et al., 2011; Bailey et al., 2000). However, three studies found that children with ASD had a significantly higher level of activity than

typically developing controls (Chuang et al., 2012; Ostfeld-Etzion et al., 2016) and norms (Brock et al., 2012). Discrepancies observed here may be due to different measures of temperament, culture, methodological quality (e.g., controlling vs not controlling for confounders) or varying age ranges of participants. Furthermore, Hepburn and Stone (2006) reported that over half of the children with ASD fell within the average range for level of activity, whilst Chuang et al. (2012) reported that one third fell within the difficult range.

The domain of *high intensity pleasure* refers to the frequency of a child engaging in recreational activities with a high physical component (e.g., bike riding). Of the four articles reporting on this feature, one reported that children with ASD engaged in more high intensity activities than the normative reference group (Adamek et al., 2011). The other three articles, which used a case-control design, reported no significant differences (Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016). Considering that Adamek et al. (2011) had poor methodological quality overall, this may explain the inconsistencies between their findings and the other articles. Only one study reported on a related domain, *positive anticipation*, finding that children with ASD were significantly less excitable in anticipation of pleasurable activities than typically developing children (Macari et al., 2017).

Eight articles reported on the bidirectional domain of *Approach-Withdrawal* (Bailey et al., 2000; Barger et al., 2019; Brock et al., 2012; Chuang et al., 2012; Hepburn & Stone, 2006; Konstantareas & Stewart, 2006; Ostfeld-Etzion et al., 2016; Reyes et al., 2019). This domain relates to whether a child approaches new people, objects and environments or tends to withdraw from novel stimuli. Five studies found that children with ASD were reported to be significantly more withdrawn than typically developing controls (Barger et al., 2019; Chuang et al., 2012; Reyes et al., 2019) and normative reference groups (Bailey et al., 2000; Brock et al., 2012); however two case-control studies found no differences between groups (Konstantareas & Stewart, 2006; Ostfeld-Etzion et al., 2016). Again, Hepburn and Stone

(2006) reported that over half of their ASD sample fell within the average range for this domain, and Chuang et al. (2012) found approximately one third to fall within the difficult range. On a related domain, *shyness*, three articles (Konstantareas & Stewart, 2006; Øien et al., 2018; Ostfeld-Etzion et al., 2016) found children with ASD to exhibit more shyness than their typically developing controls, whereas Adamek et al. (2011) found no difference with a normative reference group. More specifically, Øien et al. (2018) found gender-based differences in *shyness,* whereby girls with ASD were rated as less shy, and boys with ASD as shyer, than typically developing peers. Further, Øien et al. (2018) was the only article to report on *sociability*, finding that boys with ASD were less sociable than their typically developing counterparts.

Overall, six domains related to surgency were examined across 11 articles. Results were mixed for two domains (i.e., activity level, high intensity pleasure) and two other domains were only examined in one article (i.e., positive anticipation & sociability). This shows that there is limited evidence for differences in temperament features related to positive emotions. However, there is some evidence to suggest that children with ASD are more withdrawn and shyer compared to typically developing children.

3.4.3 Effortful Control

Four of the five articles that reported on *inhibitory control* found that children with ASD were significantly less able to focus on relevant stimuli when irrelevant stimuli were present, compared to typically developing controls (Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016) or a normative reference group (Adamek et al., 2011). Using a prospective cohort design, the fifth article (Bischof et al., 2018) found no differences between children with ASD and typically developing peers in inhibition. However, it should be noted that all participants in the study reported by Bischof et al. (2018)

were part of an intervention for inhibited children; thus group differences would not be expected.

In contrast, the four studies that assessed for differences in *distractibility* found children with ASD were significantly less distractible compared to typically developing controls (Chuang et al., 2012; Reyes et al., 2019) or normative reference groups (Bailey et al., 2000; Brock et al., 2012). Hepburn and Stone (2006) reported that over half of their sample fell within the average range for the domain of distractibility. However, one third of their sample was reported to be very difficult to distract, a proportion similar to that reported by Chuang et al. (2012).

Eight articles reported on *persistence* or *attentional focusing*, finding that children with ASD were less able to pursue tasks in the face of obstacles, compared to typically developing children (Chuang et al., 2012; Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016; Reyes et al., 2019) or normative reference groups (Adamek et al., 2011; Bailey et al., 2000; Brock et al., 2012). In addition, Hepburn and Stone (2006) and Chuang et al. (2012) reported that over 50% and approximately 42% of children with ASD were non-persistent, respectively. One of the 16 articles examined a related domain, *social inattention*, finding that children with ASD were reported to be more socially inattentive than controls (Barger et al., 2019).

Six studies investigated the *adaptability* domain, which refers to how well a child copes with change to routine. Five of the six studies found that children with ASD were significantly less adaptable to change in routine than typically developing children (Barger et al., 2019; Chuang et al., 2012; Reyes et al., 2019) or normative reference groups (Bailey et al., 2000; Brock et al., 2012). Furthermore, Hepburn and Stone (2006) and Chuang et al. (2012) reported that two thirds and approximately one third of their sample were non-adaptable, respectively.

Five articles reported on *low intensity pleasure* or *quiet persistence* (i.e., engaging in recreational activities with a low physical component, e.g., reading), with inconsistent findings. Of the four case-control studies, three (Barger et al., 2019; Macari et al., 2017; Ostfeld-Etzion et al., 2016) reported that children with ASD engaged less in low intensity pleasure compared to typically developing controls. In contrast, Adamek et al. (2011) found that children with ASD were reported as engaging more in low intensity pleasure, compared to the normative reference group. Furthermore, Konstantareas and Stewart (2006) found no difference between children with ASD and typically developing children. Discrepancies observed here may be due to different measures of temperament, methodological quality (e.g., controlling vs not controlling for confounders) or varying age ranges of participants.

Three studies (Konstantareas & Stewart, 2006; Macari et al., 2018; Ostfeld-Etzion et al., 2016) reported on the related domains of *smiling and laughter* and *joy*, with one (Konstantareas & Stewart, 2006) finding that children with ASD displayed less smiling and laugher than typically developing controls. Seven of the eight studies reporting on the related domains of *threshold*, *perceptual sensitivity* or *environmental sensitivity* found that children with ASD had a significantly lower sensory threshold than typically developing children (Barger et al., 2019; Chuang et al., 2012; Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016) or normative reference groups (Bailey et al., 2000; Brock et al., 2012). Additionally**,** Chuang et al. (2012) found a significantly smaller proportion of children with ASD falling within the difficult range on the threshold domain, compared to typically developing peers (7.5% vs 25%). However, one case-control study (Reyes et al., 2019) found no differences between groups. Three studies reported on the domain of *attentional shifting*, with all reporting that children with ASD had a significantly lower ability to shift their attention than typically developing children (Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016). Notably, no articles reported significant differences in the

domain of *impulsivity* (Adamek et al., 2011; Konstantareas & Stewart, 2006; Ostfeld-Etzion et al., 2016).

Overall, ten domains related to effortful control were examined across 13 articles. Results were mixed for only two domains (i.e., low intensity pleasure, smiling & laughter/joy) and one domain was only examined in one article (i.e., social inattention). Further, there were unanimously no differences in impulsivity. However, in all other six domains, most articles suggest that children with ASD are less able to focus on relevant stimuli, less able to switch attention, less able to pursue tasks in the face of obstacles, less adaptable to change, and more sensitive to sensory stimulation, compared to typically developing children and normative reference groups. This shows that there is substantial evidence indicating that there are differences in temperament features related to effortful control.

3.4.4 Other Domains

There were four articles reporting on the domain of *intensity*. This domain refers to the energy level a child uses in expressing their mood, regardless of direction. For example, one child may cry when they fall over, whereas another may display a sad facial expression. Three studies found that children with ASD were less intense when exhibiting their mood than typically developing children (Reyes et al., 2019) or normative reference groups (Bailey et al., 2000; Brock et al., 2012), whereas Chuang et al. (2012) found no significant difference compared to typically developing controls. Hepburn and Stone (2006), and Chuang et al. (2012) alike, found that the majority of their sample of children with ASD were reported to be mild in emotional intensity. Differences in findings between Chuang et al. (2012) and the other articles may be due to culture or controlling for different confounders.

Of the six studies that reported temperament outcomes related to the domain of *rhythmicity*, four found that children with ASD were significantly less predictable in their

biological functions than the normative reference group (Bailey et al., 2000; Brock et al., 2012) or typically developing peers (Barger et al., 2019; Reyes et al., 2019), with one case-control study finding no difference (Chuang et al., 2012). Again, differences in findings between Chuang et al. (2012) and the other articles may be due to culture or controlling for different confounders. The fourth study (Hepburn & Stone, 2006) found that more than half of the children with ASD fell within the average range for rhythmicity. One study examined *food openness* and found that children with ASD were less open to new foods than typically developing controls (Barger et al., 2019). In an additional study (Bagnato & Neisworth, 1999), children with ASD were reported to be more detached (i.e., disconnected from daily routines), hyper-sensitive (i.e., highly active, inconsolable), underreactive (i.e., unresponsive) and dysregulated (i.e., state disorganization) than the norm.

Overall, four domains not related to a previous overarching factor were examined across seven articles. Results were mostly congruent for two of the domains, intensity and rhythmicity. This shows that there is some evidence indicating that there are differences in temperament features related to intensity of emotions and rhythmicity of biological functions. However, five other domains were each only examined by one article (i.e., food openness, detached, hyper-sensitive/active, underreactive, dysregulated).

4. Discussion

The aim of this systematic review was to synthesise the literature reporting on the temperament features of children diagnosed with ASD, prior to and after diagnosis. Of the 26 included articles, ten were within infancy (i.e., pre-diagnosis) and 16 were within childhood (i.e., post-diagnosis). The quality of the articles that reported temperament outcomes within infancy was overall high. The main concern of the quality of the articles, overall, was for those that used data from the same UK-based infant-sib cohort (Clifford et al., 2013; Hendry et al., 2018; Øien et al., 2018; Pijl et al., 2019); it was not clear whether the high-risk infants

were from the same population as the low-risk infants. The 10 articles suggest that infant-sibs with later ASD had temperament features in early infancy that were mixed; some articles reported that they had more manageable features (e.g., better adaptability) whilst others reported on more challenging features (e.g., more withdrawal). By late infancy, however, these infants had distinct challenging temperament features (i.e., more negative affect, less surgency & less effortful control). These findings suggest that infants later diagnosed with ASD have an unclear temperament profile in early infancy that becomes more challenging as infants near time of diagnosis. However, evidence of an association between specific temperament features and later ASD diagnosis is weak, as no more than three articles (across two cohorts) reported on the same temperament outcome.

Alternative explanations of the findings within infancy need to be considered, given the reliance on parent-report measurement methods and infant-sib cohorts. One explanation is that parents are unconsciously comparing the temperament of their older child with ASD to their participating infant, and thus there are some reports of them being less challenging in early infancy. As time progresses and behaviours related to ASD emerge, parents rate their infants differently. The three articles that assessed for longitudinal differences reported that infant-sibs with later had less stable temperament features, compared to infant-sibs without ASD and typically developing controls. However, as the majority of articles were cross-sectional, any conclusions about changes in temperament over time cannot be made.

Later in childhood, 16 articles reported on the temperament of children diagnosed with ASD. The quality of the articles that reported temperament outcomes within childhood was mixed. The case-control studies were of higher quality overall, compared to the cross-sectional studies. The main limitations of these articles were that they did not control for important covariate (e.g., gender), use the same eligibility criteria for cases and controls, and/or assess ASD in the same manner across groups. Our synthesis found sufficient

evidence to suggest that some temperament features are associated with a diagnosis of ASD in childhood. Across nine of the 16 articles were reports of a temperament profile that is characterised by more difficulty in soothing (Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016), more withdrawal and shyness (Barger et al., 2019; Bailey et al., 2000; Brock et al., 2012; Chuang et al., 2012; Konstantareas & Stewart, 2006; Øien et al., 2018; Ostfeld-Etzion et al., 2016; Reyes et al., 2019), less intensity of emotions (Bailey et al., 2000; Brock et al., 2012; Reyes et al., 2019) and more arrythmia of biological functions (Barger et al., 2019; Bailey et al., 2000; Brock et al., 2012; Reyes et al., 2019). Further, 80-100% of articles that examined domains related to the overarching factor of effortful control found differences between children with and without ASD (Adamek et al., 2011; Bailey et al., 2000; Barger et al., 2019; Brock et al., 2012; Chuang et al., 2012; Konstantareas & Stewart, 2006; Macari et al., 2017; Ostfeld-Etzion et al., 2016; Reyes et al., 2019). These articles found that children with ASD were less able to focus on relevant stimuli (inhibitory control), less able to switch attention (distract/shifting), less able to pursue tasks in the face of obstacles (persistence), slower to adapt to change (adaptability) and more sensitive to sensory stimuli (threshold), compared to typically developing children or normative reference groups. These differences are also supported by another included article, where more than half of their sample were reported to have challenges with adapting to change and persisting with difficult tasks (Hepburn and Stone, 2006). These differences are not surprising, considering that the diagnostic criteria of ASD involves an insistence on sameness and an inflexible adherence to routines, and, in some cases, hyper-reactivity to sensory input (American Psychiatric Association, 2013).

While the majority of articles reported more challenging temperament characteristics for children with ASD, compared to comparison groups, it is important to consider studies that found temperament features that are considered to be more manageable within their

sample. These temperament features include reports of less discomfort (Adamek et al., 2011), less fear (Macari et al., 2018), greater pleasure from low- and high-intensity activities (Adamek et al., 2011), less shyness (girls; Øien et al., 2018), less intense emotional response (Bailey et al., 2000; Brock et al., 2012; Reyes et al., 2019), and less distractibility (Bailey et al., 2000; Brock et al., 2012; Chuang et al., 2012; Reyes et al., 2019). Of note, Hepburn and Stone (2006) found that more than half of their sample fell within the average range for rhythmicity, mood, activity, approach and distractibility. This highlights that many children with ASD are generally considered to have temperament features and profiles that are 'easy' rather than 'difficult'.

To conclude, the results from this systematic review supports the notion that there are some temperamental features that are more frequently present in ASD. However, these behaviours are not necessarily viewed as more challenging. It is also clear that there are significant heterogeneities in expressed behaviours among infants and children with ASD. Therefore, each child diagnosed with ASD is not necessarily going to demonstrate the same behavioural style as another.

4.1 Limitations and Directions for Future Research

While the included articles provide important information regarding the temperament profile of children with ASD, they are not without limitations. All studies, with the exception of Hirschler-Guttenberg, et al. (2015) and Macari et al. (2018), used parent-report temperament questionnaires. Given this reliance on parent-report questionnaires, temperament may not have been accurately measured. Parental biases may have been present that could potentially explain the differences, or lack thereof, between children with and without ASD. Future research would therefore benefit from including a more objective assessment of ASD (e.g., the LAB-TAB) or alternatively include a questionnaire measure based on multiple informants (e.g., parents & teachers).

Most of the cross-sectional and case-control studies did not clearly report on how cases of ASD were confirmed in participants. This is an important confound that future studies need to consider, as false cases (i.e., children without ASD) within the ASD group may result in under-detection of between-group differences. The cross-sectional studies also poorly identified confounding factors and thus did not account for them within their study design. Additionally, only two out of the nine case-control studies were able to clearly demonstrate that the same eligibility criteria were used for both cases and controls (Barger et al., 2019; Macari et al., 2018). The most likely confounding factors include age, ethnicity and particularly gender, which is known to influence both temperament and the presentation of ASD (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006; Loomes, Hull & Mandy, 2017). It is important for future research to control and/or match for these criteria, as it is possible that factors other than diagnosis of ASD could be contributing to differences in temperament features. These methodological concerns have implications for the generalisability of study findings, specifically, to gender. The majority of included studies had predominantly male children in their samples. Further, only one study examined sex differences in temperament (Øien et al., 2018). As there are known differences in temperament based on gender in typically developing populations (for review, see Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006), future research should further investigate whether the temperament profiles of children with ASD also differ as a function of gender.

Regarding the trajectory of temperament features in infants later diagnosed with ASD, only three studies directly addressed changes in temperament over time (Del Rosario et al., 2014; Paterson et al., 2019; Pijl et al., 2019). As a result, there is not enough evidence to establish temporal relationships between temperament features and a diagnosis of ASD. The knowledge of how temperament features change over time prior to diagnosis of ASD may help inform clinical practice surrounding diagnosis and intervention targets. Thus, there is a

need for future research to use prospective longitudinal designs, in order to determine whether temperament features have predictive ability of a later ASD diagnosis.

It is possible that the assessments of temperament within the included studies were measuring symptoms of ASD, rather than independent features of temperament. While temperament is proposed to be a construct independent of ASD symptoms, it currently cannot be determined. However, research pertaining to the variability of temperament features in children with ASD may shed light on this issue. Within the included articles, only two (Chuang et al., 2012; Hepburn & Stone, 2006) reported on the distribution of scores (i.e., % within/above/below 1 standard deviation of the normative mean). Most studies indicated that children with ASD as a group had more challenging temperament features. However, both Hepburn and Stone (2006) and Chuang et al. (2012) highlight that some children with ASD had temperament features comparable to normative samples. These studies suggest that temperament is a, potentially related, but largely independent construct from the diagnostic criteria for ASD, as there was significant variability among temperament profiles within the ASD groups; however, evidence is limited. Therefore, there is a need to further explore temperamental profiles (i.e. distribution of scores) to examine the variability in temperament among children with ASD, prior to and post diagnosis.

4.2 Conclusions

To the authors' knowledge, this is the first systematic review to synthesise the literature investigating the temperament of children with ASD, prior to and after diagnosis. Our review found sufficient evidence to suggest that some specific temperament features are associated with ASD in childhood. Children diagnosed with ASD were reported by their parents to be more difficult to soothe, more withdrawn and shyer, milder in emotional intensity, more arrhythmic in biological functions, and less able to control their attention and inhibition. Our review also indicates that infant siblings later diagnosed with ASD were rated

as displaying a mix of behaviours generally deemed to be 'easier' and 'challenging' in early infancy, with more challenging behaviours observed closer to time of diagnosis. However, due to the relatively small number of studies in this area, these findings should be interpreted with caution. All of articles that reported on the temperament of infants later diagnosed with ASD were conducted with infant siblings of children diagnosed with ASD. Further, only three articles assessed changes in temperament over time. Thus, our review highlights the need for future research to investigate how temperament features present and change in other infant cohorts at-risk for ASD.

Acknowledgements

The authors would like to thank Professor Lonnie Zwaigenbaum and Professor Susan L. Hepburn for providing clarification of some aspects within their included article(s), through personal correspondence with CM.

Declaration of Interest

CM, AW and OW were supported by a Research Training Program Stipend Scholarship provided by the Australian Government. VM is supported by a Career Development Fellowship from the National Health and Medical Research Council (1084816) and a Gladys M Brawn Memorial Career Development Fellowship from the University of Newcastle. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors declare that they have no conflicts of interest to disclose. This review did not require ethics approval.

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*(3), 173-183. doi:10.1177/1088357611405041

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Arlington, VA: American Psychiatric Association.

Atherton, O. E., Tackett, J. L., Ferrer, E., & Robins, R. W. (2017). Bidirectional pathways between relational aggression and temperament from late childhood to adolescence. *Journal of Research in Personality, 67*, 75-84. doi:10.1016/j.jrp.2016.04.005

Baer, J., Schreck, M., Althoff, R., Rettew, D., Harder, V., Ayer, L., . . . Hudziak, J. (2015). Child temperament, maternal parenting behavior, and child social functioning. *Journal of Child & Family Studies, 24*(4), 1152-1162. doi:10.1007/s10826-014-9924-5

Bagnato, S. J., & Neisworth, J. T. (1999). Normative detection of early regulatory disorders and autism: Empirical confirmation of DC:0-3. *Infants and young children, 12*(2), 98-106. doi:10.1097/00001163-199507000-00005

Bailey, D. B., Jr., Hatton, D. D., Mesibov, G., Ament, N., & Skinner, M. (2000). Early development, temperament, and functional impairment in autism and fragile X syndrome. *Journal of Autism and Developmental Disorders, 30*(1), 49-59. doi:10.1023/A:1005412111706

Barger, B., Moody, E. J., Ledbetter, C., D'Abreu, L., Hepburn, S., & Rosenberg, S. A. (2019). Temperament similarities and differences: A comparison of factor structures from the behavioral style questionnaire in children with and without autism spectrum disorder.

Journal of Autism & Developmental Disorders, 49(5), 1749-1762. doi:10.1007/s10803-018-03866-1

Bauminger, N., Solomon, M., & Rogers, S. J. (2010). Externalizing and internalizing behaviors in ASD. *Autism Research, 3*(3), 101-112. doi:10.1002/aur.131

Bischof, N. L., Rapee, R. M., Hudry, K., & Bayer, J. K. (2018). Acceptability and caregiver-reported outcomes for young children with autism spectrum disorder whose parents attended a preventative population-based intervention for anxiety: A pilot study. Autism Research, 11(8), 1166-1174. doi: 10.1002/aur.1963

Brock, M., Freuler, A., Baranek, G., Watson, L., Poe, M., & Sabatino, A. (2012). Temperament and sensory features of children with autism. *Journal of Autism and Developmental Disorders, 42*(11), 2271-2284. doi:10.1007/s10803-012-1472-5

Bryson, S., Garon, N., McMullen, T., Brian, J., Zwaigenbaum, L., Armstrong, V., . . . Szatmari, P. (2018). Impaired disengagement of attention and its relationship to emotional distress in infants at high-risk for autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology, 40*(5), 487-501. doi: 10.1080/13803395.2017.1372368

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*(1), 12-24. doi:10.1007/s10803-006-0328-2

Carey, W. B. (1970). A simplified method for measuring infant temperament. *The Journal of Pediatrics, 77*(2), 188-194. doi:10.1016/S0022-3476(70)80322-5

Chuang, I. C., Tseng, M.-H., Lu, L., & Shieh, J.-Y. (2012). Sensory correlates of difficult temperament characteristics in preschool children with autism. *Research in Autism Spectrum Disorders, 6*(3), 988-995. doi:10.1016/j.rasd.2012.01.002

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*(3), 673-686. doi:10.1007/s10803-012-1612-y

Clifford, S. M., Hudry, K., Elsabbagh, M., Charman, T., Johnson, M. H., & Team, B. (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*(3), 673-686. doi:10.1007/s10803-012-1612-y

Davis, S., Votruba-Drzal, E., & Silk, J. S. (2015). Trajectories of internalizing symptoms from early childhood to adolescence: Associations with temperament and parenting. *Social Development, 24*(3), 501-520. doi:10.1111/sode.12105

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*(2), 381-393. doi:10.1007/s10803-013-1876-x

Douglas, L., & David, V. (2005). Explaining girls' advantage in kindergarten literacy learning: Do classroom behaviors make a difference? *The Elementary School Journal, 106*(1), 21-38. doi:10.1086/496905

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychol Bull, 132*(1), 33-72. doi:10.1037/0033-2909.132.1.33

Garon, N., Bryson, S. E., 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. *J Abnorm Child Psychol, 37*(1), 59-78. doi:10.1007/s10802-008-9258-0

Garon, N., Zwaigenbaum, L., Bryson, S., Smith, I. M., Brian, J., Roncadin, C., . . . Roberts, W. (2016). Temperament and its association with autism symptoms in a high-risk population. *J Abnorm Child Psychol, 44*(4), 757-769. doi:10.1007/s10802-015-0064-1

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

Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development, 26*(1), 64-86. doi:10.1016/S0163-6383(02)00169-8

Goldsmith, H. H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., . . . McCall, R. B. (1987). Roundtable: What is temperament? Four approaches. *Child Development*, Roundtable, 505-529. doi:10.2307/1130527

Groh, A. M., Narayan, A. J., Bakermans-Kranenburg, M. J., Roisman, G. I., Vaughn, B. E., Fearon, R. M. P., & Van Ijzendoorn, M. H. (2017). Attachment and temperament in the early life course: A meta-analytic review. *Child Development, 88*(3), 770-795. doi:10.1111/cdev.12677

Hendry, A., Jones, E. J. H., Bedford, R., Gliga, T., Charman, T., & Johnson, M. H. (2018). Developmental change in look durations predicts later effortful control in toddlers at familial risk for ASD. *Journal of Neurodevelopmental Disorders, 10*. doi: 10.1186/s11689-017-9219-4

Hepburn, S. L., & Stone, W. L. (2006). Using carey temperament scales to assess behavioral style in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*(5), 637-642. doi:10.1007/s10803-006-0110-5

Hirschler-Guttenberg, Y., Feldman, R., Ostfeld-Etzion, S., Laor, N., & Golan, O. (2015). Self- and co-regulation of anger and fear in preschoolers with autism spectrum

disorders: The role of maternal parenting style and temperament. *Journal of Autism and Developmental Disorders, 45*(9), 3004-3014. doi:10.1007/s10803-015-2464-z

Hymel, S., Rubin, K. H., Rowden, L., & LeMare, L. (1990). Children's peer relationships: longitudinal prediction of internalizing and externaliziing problems from middle to late childhood. *Child Development, 61*(6), 2004-2021. doi:10.2307/1130854

Iverson, S. L., & Gartstein, M. A. (2018). Applications of temperament: A review of caregiver-focused temperament-driven interventions. *Early Education and Development, 29*(1), 31-52. doi:10.1080/10409289.2017.1341805

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience, 12*(4), 231-242. doi:10.1038/nrn3000

Kasari, C., & Sigman, M. (1997). Linking parental perceptions to interactions in young children with autism. *Journal of Autism and Developmental Disorders, 27*(1), 39-57. doi:10.1023/A:1025869105208

Konstantareas, M., & Stewart, K. (2006). Affect regulation and temperament in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 36*(2), 143-154. doi:10.1007/s10803-005-0051-4

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(6), 466-474. doi:10.1016/j.jaac.2017.03.013

Macari, S., DiNicola, L., Kane-Grade, F., Prince, E., Vernetti, A., Powell, K., . . . Chawarska, K. (2018). Emotional expressivity in toddlers with autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 57*(11), 828-836. doi: 10.1016/j.jaac.2018.07.872

Macari, S. L., Koller, J., Campbell, D. J., & Chawarska, K. (2017). Temperamental markers in toddlers with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 819-828. doi:10.1111/jcpp.12710

Mazurek, M. O., & Kanne, S. M. (2010). Friendship and internalizing symptoms among children and adolescents with ASD. *Journal of Autism and Developmental Disorders, 40*(12), 1512-1520. doi:10.1007/s10803-010-1014-y

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine, 6*(7), 1-6. doi:10.1371/journal.pmed.1000097

Oien, R. A., Schjolberg, S., Volkmar, F. R., Shic, F., Cicchetti, D. V., Nordahl-Hansen, A., . . . Chawarska, K. (2018). Clinical features of children with autism who passed 18-month screening. *Pediatrics, 141*(6), 1-10. doi: 10.1542/peds.2017-3596

Ostfeld-Etzion, S., Feldman, R., Hirschler-Guttenberg, Y., Laor, N., & Golan, O. (2016). Self-regulated compliance in preschoolers with autism spectrum disorder: The role of temperament and parental disciplinary style. *Autism, 20*(7), 868-878. doi:10.1177/1362361315615467

Paterson, S. J., Wolff, J. J., Elison, J. T., Winder-Patel, B., Zwaigenbaum, L., Estes, A., . . . Gu, H. (2019). The importance of temperament for understanding early manifestations of autism spectrum disorder in high-risk infants. *Journal of Autism and Developmental Disorders, 49*(7), 2849-2863. doi: 10.1007/s10803-019-04003-2

Pijl, M. K. J., Bussu, G., Charman, T., Johnson, M. H., Jones, E. J. H., Pasco, G., . . . Buitelaar, J. K. (2019). Temperament as an early risk marker for autism spectrum disorders? A longitudinal study of high-risk and low-risk infants. *Journal of Autism & Developmental Disorders, 49*(5), 1825-1836. doi:10.1007/s10803-018-3855-8

Reyes, N., Walsh, C., Soke, N., & Hepburn, S. (2019). Changes in temperament over time in young children with autism spectrum disorder and other developmental delays: A follow-up comparison study. *Journal of Intellectual Disability Research*, 63(10), 1248-1261. doi: 10.1111/jir.12648

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development, 72*(5), 1394-1408. doi:10.1111/1467-8624.00355

Studer-Luethi, B., Bauer, C., & Perrig, W. J. (2016). Working memory training in children: Effectiveness depends on temperament. *Memory & Cognition, 44*(2), 171-186. doi:10.3758/s13421-015-0548-9

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

U.S. Department of Health and Human Services Interagency Autism Coordinating Committee. (2017). *2016-2017 Interagency Autism Coordinating Committee Strategic Plan For Autism Spectrum Disorder*. Retrieved from https://iacc.hhs.gov/publications/strategic-plan/2017/.

Vaillancourt, T., Haltigan, J. D., Smith, I., Zwaigenbaum, L., Szatmari, P., Fombonne, E., . . . Bryson, S. (2017). Joint trajectories of internalizing and externalizing problems in preschool children with autism spectrum disorder. *Development and psychopathology, 29*(01), 203-214. doi:10.1017/s0954579416000043

Volker, M. A., Lopata, C., Smerbeck, A. M., Knoll, V. A., Thomeer, M. L., Toomey, J. A., & Rodgers, J. D. (2010). BASC-2 PRS profiles for students with high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders, 40*(2), 188-199. doi:10.1007/s10803-009-0849-6

Weitlauf, A. S., Gotham, K. O., Vehorn, A. C., & Warren, Z. E. (2014). Brief report: DSM-5 "levels of support": A comment on discrepant conceptualizations of severity in ASD. *Journal of Autism and Developmental Disorders, 44*(2), 471-476. doi:10.1007/s10803-013-1882-z

Zentner, M., & Bates, J. E. (2008). Child temperament: An integrative review of concepts, research programs, and measures. *International Journal of Developmental Science*, 2(1-2), 7-37. doi:10.3233/dev-2008-21203

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience, 23*(2–3), 143-152. doi:10.1016/j.ijdevneu.2004.05.001

Zwaigenbaum, L., & Penner, M. (2018). Autism spectrum disorder: advances in diagnosis and evaluation. BMJ: British Medical Journal, 361, k1674. doi:10.1136/bmj.k1674

Table 1. The characteristics of the included studies

Article	Country	Design	Sample Size	Eligibility	ASD Diagnosis	Chronological Age (Months) M ± SD (range)	% Male
1. Kasari and Sigman (1997)	USA	Case-Control	ASD = 28 TD = 28	ASD group met criteria for at least 2/3 of the diagnostic tools	• Clinician diagnosis using DSM-IV by a psychiatrist or psychologist • CARS score ≥32 • ABC score ≥70	ASD = 42.39 ± 11.61 TD = 20.29 ± 8.26	ASD = 92.9 TD = 85.7
2. Bagnato and Neisworth (1999)	USA	Cross-sectional	ASD = 36	Not reported	• Clinician diagnosis by early intervention provider or psychologist	15-45	Not reported
3. Bailey et al. (2000)	USA	Cross-sectional	ASD = 31	No suspected or confirmed diagnosis of ASD with FXS, enrolment in the state-wide autism program	• Clinician diagnosis using CARS and DSM-IV, direct observation of the child, parental report, medical records, school observations	64.1 (36 – 95)	100
4. Zwaigenbaum et al. (2005)	Canada	Prospective Cohort	T1: Autism-Sibs = 1, LR = 12 T2: Autism-Sibs = 4, LR = 19 T3: Not reported	Infant-sibs were recruited by 6 months, LR group had no $1^{st}/2^{nd}$ degree relatives with ASD, had term gestation and a birth weight >2500g	• ADOS classification of autism	Age is for all infant-sibs T1: Autism-Sibs = 6.44 ± .50 LR = 6.15 ± .43 T2: Autism-Sibs = 12.50 ± .75 LR = 12.81 ± .77 T3: Not reported	Not reported

5. Hepburn and Stone (2006)	USA	Cross-sectional	ASD = 110	Documented diagnosis of AD, PDD-NOS or Asperger's, CA between 36 and 96 months, and absence of severe motor, sensory or medical conditions	• Clinician diagnosis	57.3 ± 15.4 (23 - 94)	86
6. Konstantareas and Stewart (2006)	Canada	Case-Control	ASD = 19 TD = 23	Not reported	• Clinician diagnosis using CARS, ADOS or ADI-R	ASD = 6.16 yrs (3 - 10) TD = 6.37 yrs	ASD = 63 TD = Not reported
7. Bryson et al. (2007)	Canada	Case Series	ASD-Sibs = 9	No neurological conditions, genetic conditions and severe sensory and motor impairments	• Clinician diagnosis using ADI-R, ADOS and DSM-IV-TR	Not reported	66.7
8. Garon et al. (2009)	Canada	Prospective Cohort	ASD-Sibs = 34 Non-ASD Sibs = 104 LR = 73	Term gestation, LR group had no $1^{st}/2^{nd}$ degree relative with ASD, probands had no genetic/chromosomal/neurological disorders	• Clinician diagnosis using ADI-R, ADOS and DSM-IV-TR	Not reported	ASD-Sibs = 64.7 Non-ASD Sibs = 49.0 LR = 47.9
9. Adamek et al. (2011)	USA	Cross-Sectional	ASD = 111	Diagnosis of ASD	• School assessment of ASD by Department of Health–approved evaluator or	4.2 ± 1.5 yrs (2 - 8)	82

					• Clinician diagnosis using ADOS		
10. Brock et al. (2012)	USA	Cross-sectional	ASD = 54	No known genetic conditions, seizure disorders, epilepsy, uncorrected hearing or visual impairments, significant dysmorphic features or physical impairments	• Clinician diagnosis • ADI-R and/or ADOS • DSM-IV	56.17 ± 13.67 (36 - 84)	83.3
11. Chuang et al. (2012)	Taiwan	Case-Control	ASD = 67 TD = 44	Not reported	• Clinician diagnosis using DSM IV-TR • Catastrophic Illness Card with autism diagnosis	ASD = 64.21 ± 9.01 TD = 63.59 ± 10.14	ASD = 85.1 TD = 50.0
12. Clifford et al. (2013)	UK	Retrospective Cohort	ASD-Sibs = 17 LR = 48	Infant-sibs had older sibling with author-confirmed diagnosis of ASD, TD group were full-term, had normal birth weight and had no 1st degree relative with ASD	• Researcher diagnosis using ICD-10	Age not reported separately T1 = 7.2 ± 1.1 T2 = 13.7 ± 1.5 T3 = 23.7 ± 1.0	ASD-Sibs = 64.7 LR = 41.7
13. Del Rosario et al. (2014)	USA	Prospective Cohort	T1: ASD-Sibs = 11, TD-Sibs = 7	Infant-sibs of children with PDD-NOS and	• Clinician diagnosis using ADOS, MSEL,	T1: ASD-Sibs = 6.5 ± .9 TD-Sibs = 6.0 ± .4	ASD-Sibs = 85.7

			T2: ASD-Sibs = 16, TD-Sibs = 13 T3: ASD-Sibs = 10, TD-Sibs = 15 T4: ASD-Sibs = 10, TD-Sibs = 18 T5: ASD-Sibs = 10, TD-Sibs = 27	Asperger's disorder. Infants without ASD but with developmental concerns were excluded.	VABS, SCQ and DSM-V	T2: ASD-Sibs = 12.4 ± .6 TD-Sibs = 12.6 ± .6 T3: ASD-Sibs = 18.4 ± .4 TD-Sibs = 18.6 ± .6 T4: ASD-Sibs = 24.4 ± .6 TD-Sibs = 24.6 ± .6 T5: ASD-Sibs = 37.8 ± 4.0 TD-Sibs = 36.6 ± .5	TD-Sibs = 51.5
14. Hirschler-Guttenberg et al. (2015)	Israel	Case-Control	ASD = 39 TD = 40	Excluded from ASD group if failed to meet ASD criteria. Excluded from TD group if neuro-psychiatric diagnoses present	• Clinician diagnosis using DSM-V, confirmed by authors using ADOS	ASD = 63.38 ± 12.35 TD = 53.56 ± 13.83	ASD = 87.2 TD = 85.0
15. Garon et al. (2016)	Canada	Prospective Cohort	ASD-Sibs = 95 Non-ASD Sibs = 278	Term gestation, and no chromosomal or neurological disorders	• Clinical judgement using DSM-V and ADI-R	Not reported	ASD-Sibs = 69.5 Non-ASD Sibs = 50.4
16. Ostfeld-Etzion et al. (2016)	Israel	Case-Control	ASD = 25 TD = 32	ASD group met ASD criteria, TD group had no	• Clinician diagnosis using DSM-V,	ASD = 63.38 ± 12.35 (36 - 82)	ASD = 80.0 TD = 81.3

				neuro-psychiatric diagnoses	confirmed by authors using ADOS	TD = 53.56 ± 13.83 (29 - 78)	
17. Macari et al. (2017)	USA	Case-Control	ASD = 165 TD = 92	Not reported	• Clinical best estimate diagnosis using parent interview, developmental and medical history, MSEL, VABS, ADOS and DSM-IV	ASD = 26.46 ± 5.77 TD = 24.88 ± 5.57	ASD = 81.8 TD = 76.1
18. Bryson et al. (2017)	Canada	Prospective Cohort	ASD-Sibs = 16 Non-ASD Sibs = 67 LR = 53	HR group: Older sibling diagnosed with ASD, no known genetic or chromosomal syndrome in infant-sib or proband LR group: No known first- or second-degree relatives with ASD All infants were ≥37 weeks gestation and ≥2500g at birth.	• Clinician diagnosis using ADI-R, ADOS and DSM-IV-TR	Not reported	ASD-Sibs = 43.8 Non-ASD Sibs = 55.2 LR = 47.2

19. Bischof et al. (2018)	Australia	Prospective Cohort	ASD-Inhibited = 26 TD-Inhibited = 515	All children were: in their final pre-school year, screened as inhibited, and randomised into the cool kids trial. Children were excluded if presence of ASD was unclear.	• Parent report of ASD diagnosis	ASD-Inhibited = 4.7 yrs (median) TD-Inhibited = 4.6 yrs (median)	ASD-Inhibited = 73.1 TD-Inhibited = 50.7
20. Hendry et al. (2018)	UK	Prospective Cohort	ASD-Sibs = 16 Non-ASD Sibs = 92 LR = 24	HR: ≥1 older sibling with an ASD diagnosis LR: gestational age 38–42 weeks, ≥1 older sibling, no first-degree relatives with a diagnosis of ASD No infants had a known medical or developmental condition	• Clinical researcher diagnosis using ADOS-2, ADR-R, SCQ & DSM-5	ASD-Sibs = 37.87 ± 2.56 (30–42) Non-ASD Sibs = 39.03 ± 2.16 (35–50) LR = 38.58 ± 1.38 (6–41)	ASD-Sibs = 88 Non-ASD Sibs = 51 LR = 58
21. Macari et al. (2018)	USA	Case-Control	ASD = 43 TD = 40	No history of prematurity or known genetic abnormalities	• Clinical best estimate diagnosis using DSM-5	ASD = 21.9 ± 3.0 TD = 20.8 ± 3.9	ASD = 88 TD = 50

22. Øien et al. (2018)	NL	Retrospective Cohort	ASD-False Negative = 228 Non-ASD-True Negative = 67,969	Not reported	• Diagnoses obtained through the Norwegian Patient Registry • Diagnosis using from ADOS and ADI-R	ASD-False Negative = 18.51 ± 0.55 Non-ASD -True Negative = 18.53 ± 0.62	ASD-False Negative = 84.2 Non-ASD - True Negative = 50.8
23. Barger et al. (2019)	USA	Case-Control	ASD = 649 TD = 866	All children were: • born in study catchment area and resided there at time of study contact • lived with knowledgeable caregiver • aged between 30-68 months at time of developmental assessment	• SCQ used to assign children into ASD or TD workflow group at study entry • ADOS and ADR-R used to assign final study group classification for ASD workflow only	ASD = 4.9 ± 0.6 TD = 4.9 ± 0.6	ASD = 82 TD = 54
24. Paterson et al. (2019)	USA	Prospective Cohort	ASD-Sibs = 61 Non-ASD Sibs = 221 LR = 114	HR infants had an older sibling with confirmed ASD. LR infants had no older siblings or other first-degree relatives with ASD. Infants from both groups were excluded if they: • had a genetic condition	• Clinical judgment using ADR-I, ADOS and DSM-IV-TR	T1: HR-ASD = 6.45 ± 0.59 HR-No ASD = 6.6 ± 0.74 LR = 6.68 ± 0.74 T2: HR-ASD = 12.66 v 0.62 HR-No ASD = 12.55 v 0.60 LR = 12.61 ± 0.71 T3:	ASD-Sibs = 78.7 Non-ASD Sibs = 57 LR = 59.6

				• were born premature or with low birth weight • had a perinatal brain injury • were from a non-English speaking family • had a family history of severe mental illness		HR-ASD = 24.43 ± 0.12 HR-No ASD = 24.47 ± 0.08 LR = 24.62 ± 0.14	
25. Pijl et al. (2019)	UK	Retrospective Cohort	ASD-Sibs = 24 Atypical-Sibs = 34 Typical-Sibs = 75 LR = 66	HR infants had at least one older sibling with confirmed diagnosis of ASD but no other medical conditions. LR infants has no first-degree family members with ASD.	• Diagnoses by clinical researchers using ADOS-2, ADI-R, SCQ, and DSM-5 or ICD-10	T1: HR-ASD = 8.3 ± 1.4 HR-Atypical = 8.6 ± 1.0 HR-Typical = 8.5 ± 1.3 LR = 8.3 ± 1.4 T2: HR-ASD = 14.8 ± 1.6 HR-Atypical = 14.7 ± 1.4 HR-Typical = 14.9 ± 1.3 LR = 14.7 ± 1.3 T3: HR-ASD = 25.4 ± 2.8 HR-Atypical = 25.4 ± 2.1	ASD-Sibs = 75 Atypical-Sibs = 47.1 Typical-Sibs = 41.3 LR = 42.4

						HR-Typical = 26.0 ± 1.9 LR = 24.7 ± 1.0 T4: HR-ASD = 38.0 ± 2.0 HR-Atypical = 38.0 ± 2.8 HR-Typical = 38.5 ± 1.8 LR = 38.4 ± 2.7	
26. Reyes et al. (2019)	USA	Case-Control	ASD = 37 TD = 27	Not Reported	Diagnosis by clinical psychologist using ADOS	T1: ASD = 34.11 (26.00–46.00) TD = 19.33 (12.00–35.00) T2: ASD = 57.27 (48.00–72.00) TD = 32.65 (25.00–43.00)	ASD = 78.4 TD = 37.0

Note: Only groups of interested (i.e., ASD, TD, HR, LR) are reported in this table. Age is reported in months, unless otherwise specified.
Abbreviations: ADI-R = Autism Diagnostic Interview-Revised; ADOS = Autism Diagnostic Observation Schedule; ADOS-2 = Autism Diagnostic Observation Schedule-Second Edition; ASD = Autism Spectrum Disorder; ASD-sibs = infant-sibs later diagnosed with ASD; CARS = Childhood Autism Rating Scale; DSM = Diagnostic and Statistical Manual of Mental Disorders; ICD-10 = International Statistical Classification of Diseases and Related Health Problems, 10th revision; infant-sibs = infant siblings of children with ASD; LR = low-risk controls; MSEL = Mullen Scales of Early Learning; NL = The Netherlands; Non-ASD sibs = infant-sibs not diagnosed with ASD but may have other developmental concerns; PDD-NOS = Pervasive Developmental Disorder-Not Otherwise Specified; SCQ = Social Communication Questionnaire; T = time point; TD = typically developing; UK = United Kingdom; USA = United States of America; VABS = Vineland Adaptive Behavior Scales.

Table 2. The temperament measures, assessors and outcomes of the included studies

Article	Temperament Measure(s)	Temperament Assessor	Comparison Group	Temperament Outcome(s)
1. Kasari and Sigman (1997)	Behavioral Style Questionnaire	Parents; % mothers not specified	Typically developing	• Higher scores on *difficultness*[a]
2. Bagnato and Neisworth (1999)	Temperament and Atypical Behavior Scale	Not specified	Normative reference	• Higher[b] scores on Detached, Hyper-sensitive/active, Underactive, and Dysregulated
3. Bailey et al. (2000)	Behavioral Style Questionnaire	Mothers	Normative reference	• Higher scores on Adaptability, Persistence, Approach and Rhythmicity • Lower scores on Intensity, Distractibility and Threshold
4. Zwaigenbaum et al. (2005)	Infant Behavior Questionnaire Toddler Behavior Assessment Questionnaire	Parents; % mothers not specified	Typically developing Infant-sibs without ASD	6 mo • Lower score on Activity Level 12 mo • Higher scores on Distress to Limitations and Duration of Orienting 24 mo • Lower scores on Attentional Shifting, Inhibitory Control and Positive Anticipation
5. Hepburn and Stone (2006)	Behavioral Style Questionnaire	Mothers	None; descriptive	• Over half of sample were in the average range for Activity, Rhythmicity, Approach, Mood and Distractibility • Over half were in the difficult range for Persistence • Two thirds were in the difficult range for Adaptability • One third was in the difficult range for Mood and Distractibility

6. Konstantareas and Stewart (2006)	Children's Behavior Questionnaire	Parents; % mothers not specified	Typically developing	• Higher scores on Discomfort and Shyness • Lower scores on Attentional Focusing*, Soothability*, Inhibitory Control*, Attentional Shifting*, Perceptual Sensitivity, and Smiling and Laughter
7. Bryson et al. (2007)	Infant Behavior Questionnaire Toddler Behavior Assessment Questionnaire	Not specified	None; descriptive	• Temperament outcomes were not clearly described
8. Garon et al. (2009)	Toddler Behavior Assessment Questionnaire-Revised	Parents; % mothers not specified	Typically developing Infant-sibs without ASD	• Lower on Behavioural Approach[a] and Emotion Regulation[a]
9. Adamek et al. (2011)	Children's Behavior Questionnaire-Short Form	Parents; 87% mothers	Normative reference	• Higher scores on Anger/Frustration, High Intensity Pleasure and Low Intensity Pleasure • Lower scores on Discomfort, Inhibitory Control and Attentional Focusing
10. Brock et al. (2012)	Behavioral Style Questionnaire	Parents; % mothers not specified	Normative reference	• Lower scores on Intensity, Threshold and Distractibility • Higher scores on Activity, Approach, Adaptability, Rhythmicity and Persistence
11. Chuang et al. (2012)	Behavioral Style Questionnaire (Chinese version)	Parents; % mothers not specified	Typically developing	• Higher scores on Activity, Approach, Adaptability and Persistence • Lower scores on Threshold and Distractibility
12. Clifford et al. (2013)	Infant Behavior Questionnaire-Revised Early Childhood Behavior Questionnaire	Parents; % mothers not specified	Typically developing	7 mo • Lower scores on Approach 14 mo • Lower scores on Smiling and laughter, and Cuddliness

				24 mo • Lower scores on Soothability, Low-Intensity Pleasure and Cuddliness 24 mo • Higher scores on Sadness and Shyness
13. Del Rosario et al. (2014)	Revised Infant Temperament Questionnaire Toddler Temperament Scale Behavioral Style Questionnaire	Parents; % mothers not specified	Infant-sibs with typical development	6 mo • Lower scores on Adaptability and Approach 12 mo • Lower scores on Adaptability 24 mo • Higher scores on Adaptability and Approach 36 mo • Higher scores on Adaptability and Approach Longitudinal • Trajectory from 6mo to 36mo differed on Activity, Adaptablity and Approach.
14. Hirschler-Guttenberg et al. (2015)	Laboratory Temperament Assessment Battery	Researchers	Typically developing	• No differences were found on the two administered tasks, Fear and Anger/Frustration
15. Garon et al. (2016)	Infant Behavior Questionnaire Toddler Behavior Assessment Questionnaire-Revised	Parents; % mothers not specified	Infant-sibs without ASD	12 mo • Lower scores on Positive Affect 24 mo: • Lower scores on Positive Affect and Effortful Control
16. Ostfeld-Etzion et al. (2016)	Children's Behavior Questionnaire	Mothers	Typically developing	• Higher scores on Activity Level and Shyness • Lower scores on Attention Focusing*, Attention Shifting, Soothability, Inhibitory Control*, Perceptual Sensitivity* and Low-Intensity Pleasure

17. Macari et al. (2017)	Toddler Behavior Assessment Questionnaire - Supplement	Parents; 74.5% mothers for ASD group, % mothers for TD group not specified	Typically developing	• Lower scores on Attentional Focusing, Attentional Shifting, Inhibitory Control, Low-Intensity Pleasure, Soothability, Positive Anticipation and Perceptual Sensitivity
18. Bryson et al. (2017)	Infant Behavior Questionnaire	Parents; % mothers not specified	Typically developing Infant-sibs without ASD	6mo • Higher[b] scores on Activity Level, Distress to Limitations, Fear, Smiling & Laughter, and Soothability • Lower[b] scores on Duration of Orienting 12mo • Higher[b] scores on Activity Level, Distress to Limitations, Fear and Soothability • Lower[b] scores on Duration of Orienting, and Smiling & Laughter
19. Bischof et al. (2018)	Short Temperament Scale for Children	Parents; 94% mothers for total sample	Typically developing-Inhibited	• No differences were found on Inhibition.
20. Hendry et al. (2018)	Children's Behavior Questionnaire-Very Short Form	Parents; % mothers not specified	Typically developing Infant-sibs without ASD	36mo • Lower scores on Effortful Control
21. Macari et al. (2018)	Laboratory Temperament Assessment Battery	Researchers	Typically developing	• Lower scores on Fear
22. Øien et al. (2018)	Emotionality Activity Sociability Temperament Survey	Parents; % mothers not specified	Non-ASD (General Population)	• Higher scores on Shyness (females) • Lower scores on Sociability and Shyness (males)

23. Barger et al. (2019)	Behavioral Style Questionnaire	Parents; % mothers not specified	Non-ASD (General Population)	• Higher scores on Maladaptive[a], Social Inattention[a], Crying[a] • Lower scores on Environmental Sensitivity[a], Quiet Persistence[a], Social Approach[a], Rhythmicity[a] and Food Openness[a]
24. Paterson et al. (2019)	Infant Behavior Questionnaire-Revised Early Childhood Behavior Questionnaire	Parents; % mothers not specified	Typically developing Infant-sibs without ASD	6mo • Lower scores on Effortful Control 12mo • Higher scores on Negative Affect • Lower scores on Surgency and Effortful Control Longitudinal • Trajectory from 6mo to 24mo differed on Surgency
25. Pijl et al. (2019)	Infant Behavior Questionnaire-Revised Early Childhood Behavior Questionnaire	Parents; % mothers not specified	Typically developing Infant-sibs without ASD Infant-sibs with typical development	8mo • Higher scores on Negative Affect 14mo • Higher scores on Negative Affect • Lower scores on Surgency and Effortful Control 24mo • Higher scores on Negative Affect • Lower scores on Effortful Control Longitudinal • Trajectory from 8mo to 24mo differed on Surgency, Negative Affect and Effortful Control

26. Reyes et al. (2019)	Carey Temperament Scales; versions not specified	Parents; % mothers not specified	Typically developing	T1 (~2yo) • Higher scores on Approach, Adaptability, Intensity and Mood • Lower scores on Distractibility T2 (~4yo) • Higher scores on Rhythmicity, Approach, Mood and Persistence • Lower scores on Distractibility

Note: Unless otherwise stated, all group differences presented are statistically significant.
*Statistically significant after multiple comparisons analysis
[a]Author-constructed domain
[b] Descriptive difference

Table 3. Outcome of Methodological Quality Assessment: JBI Critical Appraisal Checklist

	Item Number											% of Items Meeting Requirements
	1	2	3	4	5	6	7	8	9	10	11	
Case Control												
Kasari and Sigman (1997)	Yes	Yes	No	Yes	No	Yes	Yes	Yes	N/A	Yes	***	78
Konstantareas and Stewart (2006)	Yes	Yes	Unclear	Yes	No	Yes	Yes	Yes	N/A	Yes	***	78
Chuang et al. (2012)	Unclear	No	Unclear	Yes	No	No	No	Yes	N/A	Yes	***	33
Hirschler-Guttenberg et al. (2015)	Yes	Yes	Unclear	Yes	No	Yes	Yes	Yes	N/A	Yes	***	78
Ostfeld-Etzion et al. (2016)	Yes	Yes	Unclear	Yes	No	Yes	Yes	Yes	N/A	Yes	***	78
Macari et al. (2017)	Yes	Yes	Unclear	Yes	No	Yes	Yes	Unclear	N/A	Yes	***	67
Macari et al. (2018)	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	N/A	Yes	***	89
Barger et al. (2019)	No	No	Yes	Yes	No	Yes	Yes	Yes	N/A	Yes	***	67
Reyes et al. (2019)	No	No	Unclear	Yes	Yes	Yes	Yes	Yes	N/A	Yes	***	60
Case Series												
Bryson et al. (2007)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	***	90
Analytical Cross Sectional												
Bagnato and Neisworth (1999)	No	No	Unclear	Unclear	No	No	Unclear	Yes	***	***	***	13
Bailey et al. (2000)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	***	***	***	100
Hepburn and Stone (2006)	Yes	Yes	Unclear	Unclear	No	No	Yes	Yes	***	***	***	50
Adamek et al. (2011)	No	Yes	Unclear	No	No	No	Yes	Yes	***	***	***	38
Brock et al. (2012)	Yes	Yes	Yes	Yes	No	No	Yes	Yes	***	***	***	75
Cohort												
Zwaigenbaum et al. (2005)	Yes	Yes	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	100
Garon et al. (2009)	Yes	Yes	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	100
Clifford et al. (2013)	Unclear	Yes	Unclear	No	No	N/A	Unclear	N/A	N/A	N/A	Yes	29
Del Rosario et al. (2014)	Yes	Yes	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	100
Garon et al. (2016)	Yes	Yes	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	100
Bryson et al. (2017)	Yes	Yes	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	100
Bischof et al. (2018)	Yes	Yes	No	Yes	No	N/A	Yes	N/A	N/A	N/A	Yes	71
Hendry et al. (2018)	Unclear	Yes	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	86
Øien et al. (2018)	Unclear	Yes	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	86
Paterson et al. (2019)	Yes	Yes	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	100
Pijl et al. (2019)	Unclear	No	Yes	Yes	Yes	N/A	Yes	N/A	N/A	N/A	Yes	71

Note: The following items were removed from the forms as they were not applicable to any of the assessed studies: *Case Control* (item 9) and *Cohort* (items 6, 8, 9 & 10).
***Denotes that item does not exist for checklist version

Table 4. Summary of findings on temperament traits in children with ASD

Age Group	ASD vs. other infant-sibs[a]			ASD vs. typically developing controls/reference samples		
	Negative Affectivity	Surgency/Extraversion	Effortful Control	Negative Affectivity	Surgency/Extraversion	Effortful Control
Infancy (6-36 months)	> negative affect at 8, 14 and 24 mo[25] < Rate of recovery from arousal at 6 mo[25] < soothable at 24 mo[24]	> approach at 6 mo[13] < positive affect at 12 and 24 mo[15] > withdrawal at 6 mo[25], 24 mo[13] and 36 mo[13] < surgency at 6 mo[24], 12 mo[24] and 14 mo[25] < smiling and laughter at 6 and 12 mo[25] < vocalisations at 6 and 12 mo[25] < impulsivity at 24 mo[25] < excitability at 24 mo[25] < sociability at 24 mo[25]	> adaptable to change at 6 and 12 mo[13] < adaptable to change at 24 and 36 mo[13] < effortful emotion regulation at 6 mo[24], 14 mo[25], 24 mo[15,25] and 36 mo[20] < low-intensity pleasure at 6 and 24 mo[25] < attentional focusing at 24 mo[24] < attentional shifting at 24 mo[24] < inhibitory control at 24 mo[24]	> frustration at 12 mo[4] and 24 mo[25] < soothable at 24 mo[12,24] > sadness at 12 mo[25] and 24 mo[12,24] > negative affect at 8 mo[25], 12 mo[24], 14 mo[25] and 24 mo[25] < Rate of recovery from arousal at 6 and 12 mo[25] > discomfort at 24 mo[25]	< activity level at 6 mo[4] < smiling and laughter at 6 mo[25] and 14 mo[12] > smiling and laughter at 12 mo[25] < cuddliness at 14 and 24 mo[12] > shyness at 24 mo[12] < excitability at 24 mo[4,24] < behavioural approach at 24 mo[8] < surgency at 12 mo[24] and 14 mo[25] < vocalisations at 6 mo[25] > vocalisations at 12 mo[25] > withdrawal at 6 mo[25] < sociability at 24 mo[25]	> duration of orienting at 12 mo[4] < low-intensity pleasure at 6 mo[25] and 24 mo[12,24] < inhibitory control at 24 mo[4,24] < attentional shifting at 24 mo[4] < effortful emotion regulation at 6 mo[24], 12 mo[24], 24 mo[8] 14 mo[25] 24 mo[25] and 36 mo[20]
Childhood (mean age 2-6 years)				> difficult[1] < soothable[6,16,17] > discomfort[6] < discomfort[9] > arrhythmia[3,10,23,26] > anger/frustration[9] > detached[2] < intensity of fear[21] > crying[23] > negative mood[26]	< smiling and laughter[6] > shyness[6,16,22(boys)] < shyness[22(girls)] > activity level[10,11,16] < excitability[19]7 > high-intensity pleasure[9] < emotional intensity[3,10] > emotional intensity[26] > hyper-sensitive/active[2] > underreactive[2] > withdrawal[3,10,11,23,26] < sociability[22(boys)] < open to new food[23]	< attentional shifting[5,16,17] < low-intensity pleasure[16,17] > low-intensity pleasure[9] < adaptable to change[3,10,11,23,26] < persistent[3,10,11,23,26] < distractable[3,10,11,26] < attentional focusing[6,9,16,17] < social attention[23] < inhibitory control[6,9,16,17] < perceptual sensitivity[6,16,17] > sensory threshold[3,10,11] < environmental sensitivity[23] > dysregulated[2]

Note: > symbol = more or greater. < symbol = less, lower or slower. Results from Bryson et al. (2007), Hepburn and Stone (2006), and Bryson et al. (2017) are not reported within this table, as they did not assess group differences. Hirschler-Guttenberg et al. (2015) and Bischof et al. (2018) are not reported within this table, as they did not report any significant group differences.
[a]Infant-sibs classified as typically developing or as non-ASD.

1. Kasari and Sigman (1997)
2. Bagnato and Neisworth (1999)
3. Bailey et al. (2000)
4. Zwaigenbaum et al. (2005)
5. Hepburn and Stone (2006)
6. Konstantareas and Stewart (2006)
7. Bryson et al. (2007)
8. Garon et al. (2009)

9. Adamek et al. (2011)
10. Brock et al. (2012)
11. Chuang et al. (2012)
12. Clifford et al. (2013)
13. del Rosario et al. (2014)
14. Hirschler-Guttenberg et al. (2015)
15. Garon et al. (2016)
16. Ostfeld-Etzion et al. (2016)
17. Macari et al. (2017)
18. Bryson et al. (2017)
19. Bischof et al. (2018)
20. Hendry et al. (2018)
21. Macari et al. (2018)
22. Øien et al. (2018)
23. Barger et al. (2019)
24. Paterson et al. (2019)
25. Pijl et al. (2019)
26. Reyes et al. (2019)

Table 5. Definitions of selected temperament domains, by framework

Temperament Framework	Domain	Description
Carey & Colleagues[a]	Activity	The motor component present in a given child's functioning, and the diurnal proportion of active and inactive periods.
	Rhythmicity	The predictability and/or the unpredictability…[of behaviour related to]…the sleep-wake cycle, hunger, feeding pattern, and elimination schedule.
	Approach	The nature of the response to a new stimulus, be it a new food, new toy, or new person.
	Adaptability	Responses to new or altered situations. One is not concerned with the nature of the initial responses, but with the frequency with which they were successfully modified in desired directions.
	Intensity	The energy level of response, irrespective of its quality or direction.
	Mood	The amount of pleasant, joyful, and friendly behavior, as contrasted with unpleasant, crying, and unfriendly behavior.
	Distractibility	The effectiveness of extraneous environmental stimuli in interfering with, or in altering the direction of, the ongoing behavior.
	Persistence	The continuation of an activity in the face of obstacles to the maintenance of the activity direction.
	Threshold	The intensity level of stimulation that is necessary to evoke a discernible response, irrespective of the specific form that the response might take or the sensory modality affected. The behaviors utilised are those concerning reactions to sensory stimuli, environmental objects, and social contacts.
Rothbart & Colleagues[b]	Activity Level	Gross motor activity, including rate and extent of locomotion.
	Anger	Negative affectivity related to interruption of ongoing tasks or goal blocking.

Attentional Shifting	Capacity to maintain attentional focus on task-related channels.
Cuddliness	Expression of enjoyment and molding of the body to being held by a caregiver.
Discomfort	Negative affectivity related to sensory qualities of stimulation, including intensity; rate; or complexities of light, movement, sound, and texture.
Distress to Limitations/Frustration	Fussing, crying or showing distress while (a) in a confining place or position; (b) in caretaking activities; (c) unable to perform a desired action.
Duration of Orienting/Attentional Focusing	Capacity to maintain attentional focus on task-related channels.
Fear	Negative affectivity, including unease, worry, or nervousness, which is related to anticipated pain or distress and/or potentially threatening situations.
High Intensity Pleasure	Pleasure or enjoyment related to situations involving high stimulus intensity, rate, complexity, novelty, and incongruity.
Impulsivity	Speed of response initiation.
Inhibitory Control	Capacity to plan and to suppress inappropriate approach responses under instructions or in novel or uncertain situations.
Low Intensity Pleasure	Pleasure or enjoyment related to situations involving low stimulus intensity, rate, complexity, novelty, and incongruity.
Perceptual Sensitivity	Detection of slight, low-intensity stimuli from the external environment.
Positive Anticipation/Approach	Amount of excitement and anticipation for expected pleasurable activities.
Sadness	Negative affectivity and lowered mood and energy related to exposure to suffering, disappointment, and object loss.
Shyness	Slow or inhibited (versus rapid) speed of approach and discomfort (versus

		comfort) in social situations.
	Smiling & Laughter	Positive affect in response to changes in stimulus intensity, rate, complexity, and incongruity.
	Soothability	Rate of recovery from peak distress, excitement, or general arousal.
Bagnato & Colleagues[c]	Detached	Aloof, self-absorbed, disconnected from daily routines (active avoidance).
	Hyper-sensitive/active	Overreactive, low sensory threshold, highly active, impulsive, inconsolable, negative, and defiant.
	Underreactive	Unresponsive, high sensory threshold, poor awareness, low alertness (passive avoidance suggesting neurophysiological origins).
	Dysregulated	State disorganization and dyscontrol (suggesting a neurophysiological basis).
Buss & Plomin[d]	Emotionality	Emotionality…is equivalent to distress. The dimension varies from an almost stoic lack of reaction to intense emotional reactions that are out of control. Examples of the high extreme are crying, tantrums, difficulty in being soothed, a low threshold for the aversive stimuli that trigger distress, and intense activation of the sympathetic division of the autonomic nervous system. Emotionality clearly involves emotional arousal and, to a lesser extent, behavioral arousal.
	Activity	Activity…[has]…two major components of which are tempo and vigor. Individuals vary from lethargy to an almost hypomanic push of energetic behavior. Activity involves behavioral arousal, specifically, elevated amplitude and rate of responses. Such behavioral arousal is different from the physiological and experiential arousal that occurs in emotionality.
	Sociability	Sociability…is the preference for being with others rather than being alone. No normal person is expected to be a hermit, but there are wide variations in the need to be with others. Sociable individuals seek to share activities, to receive attention from others, and to be involved in the back-and-forth

responsivity that characterizes social interaction.

[a] Descriptions are quoted from Thomas et al. (1963), p. 40-42
[b] Descriptions are quoted from Gartstein and Rothbart (2003), p. 72; Rothbart, Ahadi, Hershey, and Fisher (2001), p. 1406
[c] Descriptions are quoted from Bagnato and Neisworth (1999), p. 102
[d] Descriptions are quoted from Goldsmith et al. (1987), p. 512

Table S1. Database search terms

	AND:	AND:	AND:
autism	diagnos*	TI (temperament OR temperamental)	child*
OR	OR	OR	OR
ASD	high-risk	AB (temperament OR temperamental)	infan*
OR	OR	OR	OR
autism spectrum disorder	genetic risk	SU (temperament OR temperamental)	toddler*
OR	OR		OR
autistic disorder	familial risk		baby
OR	OR		OR
PDD-NOS	prematur*		babies
OR	OR		OR
Pervasive Developmental Disorder-Not Otherwise Specified	sibling*		newborn
OR	OR		
Asperger Syndrome	disorder		
OR	OR		
Asperger's	(autis* OR Pervasive Developmental Disorder-Not Otherwise Specified OR Asperger* OR Childhood Disintegrative Disorder) adj5 symptom*		
OR	OR		
Childhood Disintegrative Disorder	(autis* OR Pervasive Developmental Disorder-Not Otherwise Specified OR Asperger* OR Childhood Disintegrative Disorder) adj5 sign*		
	OR		
	(autis* OR Pervasive Developmental Disorder-Not Otherwise Specified OR Asperger* OR Childhood Disintegrative Disorder) adj5 marker*		
	OR		
	(autis* OR Pervasive Developmental Disorder-Not Otherwise Specified OR Asperger* OR Childhood Disintegrative Disorder) adj5 trait*		
	OR		
	(autis* OR Pervasive Developmental Disorder-Not Otherwise Specified OR Asperger* OR Childhood Disintegrative Disorder) adj5 indicator*		

	OR (autis* OR Pervasive Developmental Disorder-Not Otherwise Specified OR Asperger* OR Childhood Disintegrative Disorder) adj5 precursor* OR (autis* OR Pervasive Developmental Disorder-Not Otherwise Specified OR Asperger* OR Childhood Disintegrative Disorder) adj5 risk

Note: *Indicates that the search included the term and any other possible terms stemming from it (e.g., diagnos* captured diagnosis, diagnostic, diagnose, diagnosed). '(autis* OR Pervasive Developmental Disorder-Not Otherwise Specified OR Asperger* OR Childhood Disintegrative Disorder) adj5…' is a search phrase that requires the word autis* (or Pervasive Developmental Disorder-Not Otherwise Specified or Asperger* or Childhood Disintegrative Disorder) to be within 5 words of the following search term in a section of text (e.g., 'autis* adj5 marker*' could detect 'temperament may act as a behavioural marker for autism').