

REPORT

WILEY

Temperament and behaviour problems in children: A multilevel analysis of cross-cultural differences

Allegra X. Campagna¹  | Eric D. Desmarais¹ | Brian French¹ |
 Joshua J. Underwood¹  | Mirjana Majdandžić² |
 Roseriet Beijers³ | Carolina de Weerth⁴ | Eun Gyoung Lee⁵ |
 Blanca Huitron⁶ | Emine Ahmetoglu⁷ | Oana Benga⁸ |
 Katri Raikkonen⁹ | Kati Heinonen⁹ | Carmen Gonzalez-Salinas¹⁰ |
 Helena Slobodskaya¹¹  | Elena Kozlova¹¹  |
 Maria Beatriz Martins Linhares¹² | Felipe Lecannelier¹³ |
 Sara Casalin¹⁴ | Ibrahim Acar¹⁵ | Soile Tuovinen⁹ |
 Zhengyan Wang¹⁶  | Rosario Montiroso¹⁷ | Lorenzo Giusti¹⁷ |
 Seong-Yeon Park¹⁸ | Sae-Young Han¹⁸ | Samuel Putnam¹⁹ |
 Maria A. Gartstein¹

Correspondence

Joshua J. Underwood, PO Box 644820,
 Johnson Tower 233, Washington State
 University, Pullman, WA 99164-4820, USA.
 Email: joshua.underwood@wsu.edu

Abstract

Early temperament attributes have been linked to emerging behaviour problems and significant long-term consequences; however, these relations are rarely examined cross-culturally. The present study addresses this gap, employing multilevel modelling to explain within- and between-culture variances with respect to temperament predicting a spectrum of behaviour problems across 14 nations from the Joint Effort Toddler Temperament Consortium (JETTC). A total of 865 children between 17 and 40 months, with approximately equal age distribution across this developmental period and about equivalent

For affiliations refer to page 15

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Infant and Child Development* published by John Wiley & Sons Ltd.

representation of genders, were recruited from 14 nations. Greater negative emotionality was associated with more internalizing problems, whereas higher surgency and effortful control predicted fewer internalizing difficulties. Controlling for age and gender, temperament explained significant within- and between-culture variances in internalizing and externalizing problems (at the broad-band and fine-grained levels), as well as sleep problems. For internalizing difficulties, temperament accounted for more between-culture variance. In contrast, for externalizing difficulties, temperament accounted more for how individuals within the same culture differed from their same-culture counterparts. The within-culture findings suggest universal patterns of temperament-problem relations, informing cultural adaptation of interventions; between-culture findings enhance understanding of the implications of the cultural niche for normative behaviour and adjustment.

KEYWORDS

behavior problems, cross-cultural research, early development, multilevel modeling, temperament

1 | INTRODUCTION

1.1 | Childhood behaviour problems

Childhood behaviour problems predict significant long-term consequences (Kassing et al., 2019). Internalizing problems manifest as withdrawal, anxiety, fearfulness and depression, whereas externalizing difficulties include aggression, oppositionality, vindictiveness and destructive behaviour (Achenbach, 1991). Early internalizing and externalizing symptoms are associated with later peer relationship problems, lower-quality interactions with parents and delinquency (Weeks et al., 2016). Externalizing difficulties predict compromised social competence and academic underachievement (Brumley & Jaffee, 2016) and youth with externalizing problems were more likely to engage in risky behaviour, affiliate with deviant peers and be rejected by prosocial peers (Brumley & Jaffee, 2016). Early internalizing problems are linked with poor peer relationships and also depression and avoidant tendencies later (Fanti & Henrich, 2010; Korhonen et al., 2018). Internalizing and externalizing problems increase risk of peer victimization, with maladaptive exchanges maintaining these symptoms (Buchanan-Pascall et al., 2018; Reijntjes et al., 2011). Thus, early identification and intervention for behaviour problems in childhood are critical.

1.2 | Temperament

Temperament, defined as constitutionally based individual differences in reactivity and regulation (M. Rothbart & Derryberry, 1981), is an established contributing factor to the emergence of behaviour

problems (Gartstein et al., 2012). The Rothbart psychobiological model defines temperament as individual differences in motor, attentional, emotional reactivity and self-regulation, linked with underlying neuro-behavioral systems and subject to contextual influences (Riem et al., 2017). The psychobiological framework also provides a structural model of temperament, with three overarching factors, each encompassing fine-grained dimensions (Gartstein, Putnam, et al., 2016).

Negative emotionality has been linked conceptually and empirically to the personality trait of neuroticism in adulthood (Evans & Rothbart, 2007). This factor consists of fear, sadness, frustration, discomfort and decreased ability to lower arousal/distress (Gartstein & Rothbart, 2003). Elevations in negative emotionality have been linked to increased activity in the amygdala, as well as other structures associated with appraisal and response to aversive stimuli (e.g., dorsolateral prefrontal cortex and anterior cingulate cortex) (Davis et al., 2019). Overall, negative emotionality poses risk to both internalizing and externalizing disorders in childhood and adolescence (Putnam et al., 2006; Riem et al., 2017).

Surgency manifests as smiling, laughing, displaying pleasure, approaching novel stimuli, sociability and activity/impulsivity (Putnam et al., 2006). Surgency is sometimes used interchangeably with the terms 'Positive emotionality' and 'Extraversion'. Individuals higher in positive affect have the tendency to be engaged with their environment (Olino et al., 2018), presumably due to increased approach tendencies resulting from greater activity of structures such as the medial forebrain bundle and lateral hypothalamus (i.e., behavioural activation system [BAS]) (M. K. Rothbart & Hwang, 2005). Although surgency is often viewed as a protective factor with respect to internalizing problems in contrast to negative emotionality (Dolcini-Catania et al., 2020), elevations in surgency have been associated with externalizing behaviours (Oldehinkel et al., 2004). Additionally, higher surgency was found to be particularly problematic in combination with high negative emotionality (Gartstein et al., 2012).

The third temperament factor, effortful control/regulatory capacity, develops later, and its emergence coincides with the maturation of the anterior attentional mechanisms and executive attention, influenced by cortical (e.g., prefrontal cortex and anterior cingulate) and subcortical (e.g., basal ganglia) regions of the brain (Posner et al., 2012). In general, effortful control is characterized by attentional control, orienting capacity and inhibition of a dominant response in favour of a subdominant, more adaptive reaction (Bornstein et al., 2015; Liu & Bell, 2020). This factor is closely associated with the personality construct of Conscientiousness and is broadly viewed as protective relative to the risks associated with elevated reactivity (Riem et al., 2017; Schwartz et al., 1999). These temperament attributes are important in their own right, and are also widely investigated because of documented links with later occurring symptoms/disorders.

1.3 | Temperament and behaviour problems

Early fear, shyness and emotional reactivity, associated with negative emotionality, reliably predict later internalizing problems (Dougherty et al., 2010; Schwartz et al., 2003). Positive affectivity, an aspect of surgency, predicted lower depressive symptoms (Dougherty et al., 2010; Lonigan et al., 2003). Higher surgency does not appear to be ubiquitously protective however, and was linked with an increased risk for externalizing behaviour problems (Berdan et al., 2008; Stifter et al., 2008). It has been suggested that the latter is largely a function of impulsivity, a component of surgency associated with externalizing symptoms/disorders (e.g., attention deficit hyperactivity disorder [ADHD]; Beauchaine et al., 2010). Effortful control is consistently described as protective, associated with more advanced academic performance (Sánchez-Pérez et al., 2018), prosocial behaviour and social competence (Eisenberg et al., 2003), as well as fewer externalizing problems (Delgado et al., 2018).

As noted, temperament has also been shown to predict more narrowly defined symptom sets. For instance, children presenting within the dysregulation profile, that is, showing a pattern of elevated scores on the aggression, attention problems and depressive/anxious subscales of the Child Behaviour Checklist (CBCL) had higher scores of negative emotionality, higher surgency and lower effortful control (Kim et al., 2012). Effortful control and regulation

were also protective with respect to sleep problems, as well as their consequences (e.g., academic difficulties) throughout childhood/adolescence (Deater-Deckard et al., 2018; Fischer & Schwartz, 2011; Ostlund et al., 2017). Sánchez-Pérez et al. (2020) showed that attentional focusing and motor activation were significantly associated with ADHD scores, whereas toddlers' frustration significantly related to Oppositional Defiant Disorder (ODD) scores. According to Kozlova et al. (2020), toddler effortful control was uniquely protective with respect to emotional symptoms (i.e., depression/anxiety) in the school-age period (around 8 years of age). Together, these studies demonstrate predictive utility of temperament in explaining narrowly defined behaviour problems/symptoms in samples from different countries.

1.4 | Temperament, behaviour problems and culture

Consistent patterns of differences in temperament and behaviour problems/symptoms emerge in cross-cultural research. For example, infants from the United States were rated higher on aspects of surgency and regulatory capacity and lower on negative emotionality than infants from Russia, and Japan (Slobodskaya et al., 2013), and higher on several surgency dimensions compared with Italian children (Montirosso et al., 2011). Compared with the United States, Russian and Japanese toddlers scored lower on surgency (Slobodskaya et al., 2013). Gartstein, Putnam, et al. (2016) compared expressions of fear for infants from Spain and the United States (Gartstein, Carranza, et al., 2016). Significant differences emerged in laboratory observations, with higher fearfulness documented in Spain.

A comparison of children aged 6–11 in 12 cultures revealed significant variability in the number of externalizing and internalizing scores. Higher internalizing scores were reported for Puerto Rican and Jamaican children, with greater externalizing problems in the United States, Sweden and Germany (Crijnen et al., 1997). Although a review of CBCL scores in 11 cultures revealed minimal differences (Verhulst & Achenbach, 1995), considerable cross-cultural variability has been noted elsewhere (Achenbach & Rescorla, 2010). Rescorla et al. (2007) reported on children aged 6–16 years in 31 societies (including 10 countries examined herein) and later compared children aged 1.5–5 from 24 societies (including 11 cultures considered in this study) (Desmarais, Brown, et al., 2021; Desmarais, French, et al., 2021; Rescorla et al., 2011). Spanish children scored low in externalizing and internalizing problems, with Chilean children receiving high scores in both domains. Korean preschoolers scored low in externalizing, while Romanian and Turkish youngsters scored high in internalizing difficulties.

While notable differences in temperament and behaviour problems across cultures are evident in the existing literature, it was also observed that variance within cultures is often greater than variance between them (Deater-Deckard et al., 2018; Fischer & Schwartz, 2011). Thus, while individuals within the same culture vary substantially, shared values and experiences can be expected to influence the development of temperament and behaviour problems, resulting in reliable cross-cultural differences. In fact, although age and gender are clearly important factors contributing emotional development and outcomes, effect sizes associated with culture are often greater (Achenbach & Rescorla, 2007; Rescorla et al., 2007). Given that temperament uniquely predicts later behavioural functioning, it is critical to understand how variability in temperament within and between cultures contributes to social-emotional difficulties and differences in these presentations across cultures.

Interpretation of cross-cultural differences in temperament and behaviour problems is typically contextual in nature, although biological factors may vary across cultures, for example, different allele frequencies have been reported (Putnam & Gartstein, 2017). Nonetheless, between-culture variability is typically attributed to differences in socialization and, more broadly, the 'developmental niche'. The developmental niche is a framework that incorporates a child's physical and social environment, child-rearing practices/socialization and caregiver psychology (Majdandžić et al., 2019; Super & Harkness, 1986), capturing how culture influences these critical elements of the child's daily life, contributing to temperament development and risk for behaviour problems. The developmental niche is not just a theoretical construct, as a number of studies document connections between elements of the

niche (e.g., parenting) and child social-emotional development, along with cross-cultural differences in these domains (Bornstein, 2013; Ren et al., 2023).

Relations between temperament and behaviour problems have also been examined across cultures (Gartstein et al., 2018). Positive correlations between negative emotionality and both internalizing and externalizing problems as well as a negative correlation between effortful control and both problem sets were reported consistently across different culture groups/countries. Surgency was positively correlated with externalizing difficulties. Positive correlations between negative emotionality and internalizing problems, as well as between surgency and externalizing problems, also emerged in analyses relying on culture-level scores. This correlational approach is associated with significant limitations, violating the assumption of independent observations, conflating between- and within-culture effects and reducing meaningful variance collapsing observations within each culture into a single mean value. The present study addresses these limitations, while also broadening the scope.

1.5 | The present study

The present study utilizes multilevel modelling to explain both within- and between-culture differences in temperament factors and their links to behaviour problems, examined in terms of broad-band and narrowly defined indicators.

High negative emotionality was expected to confer risk for emotional reactivity, anxious/depressed symptoms and withdrawn behaviour. Effortful control was hypothesized to be related to lower symptom severity, especially for externalizing problems (i.e., reduced sleep problems, aggression, emotional reactivity and inattention). We expected higher surgency to predict fewer internalizing but more externalizing problems. Multilevel modelling ascertains the amount of variance in behaviour problems attributable to within- and between-cultural differences. We hypothesized that temperament would explain within-culture differences (i.e., accounting for the extent to which individuals within cultures vary) as well as significant between-culture variance previously reported for behaviour problems (Achenbach & Rescorla, 2010). Thus, temperamental similarities for individuals sharing a culture were expected to account for within-culture similarities in behaviour problems and differences relative to others.

2 | METHODS

2.1 | Participants

The Joint Effort Toddler Temperament Consortium (JETTC; Gartstein et al., 2018) included data from 14 nations, United States, Brazil, Spain, Mexico, Italy, Russia, Finland, Romania, Belgium, the Netherlands, China, South Korea, Turkey and Chile, to examine mean-level differences between cultures in both behaviour problems and temperament, measured via the CBCL (Achenbach & Rescorla, 2000) and Early Childhood Behaviour Questionnaire (ECBQ; Putnam et al., 2006), respectively.

The JETTC sites recruited an average of 61 families, with samples from individual countries ranging from 49 (Chile) to 112 (the Netherlands) for a total sample of $N = 865$. Children were between 17 and 40 months of age ($M = 26.88$ months, $SD = 5.65$ months), with an approximately equal distribution of ages across this developmental period and about an equivalent representation of genders (52% male). Although the ECBQ is optimally designed for use with children 18–36 months of age, a small subset of children between 15 and 18 months ($n = 22$, ~2% of the overall sample) and 37–40 months of age ($n = 13$, ~1% of overall sample) were included in the study. Mild expansion of age ranges is typical for childhood temperament instruments as items remain developmentally appropriate (Putnam et al., 2014). For all but two of the JETTC nations, data were collected in a single site. In the Netherlands and the United States, data collected from two locations were combined following analyses, demonstrating no

significant difference between these sites in either country. Recruitment strategies (e.g., distributing advertisements through social media, birthing classes, day care centres and hospitals) varied between sites, as is common in cross-cultural research (Keller et al., 2006) since not all approaches are equally viable in all cultures. All samples were considered representative of the communities in which recruitment took place (demographic information presented in Table 1); however, these are not necessarily reflective of the entire nation/culture taking part in JETTC. Overall, families in this study represent a range of occupations, primarily reflecting middle socio-economic status (Revised Duncan Sociometric Index [RDSI]) (Stevens & Featherman, 1981).

The JETTC countries were selected for theoretical and practical reasons. First and foremost, our selection provided an opportunity to examine temperament as a risk or protective factor with respect to psychopathology in a set of countries with a wide range in terms of cultural orientation. Extensive research addressing cultural orientation/values dimensions (Hofstede, 1980, 2011) shows that selected countries vary with respect to individualism/collectivism, masculinity/femininity, power distance, uncertainty avoidance, long-term orientation and indulgence/restraint. For example, for countries higher in individualism (the United States, the Netherlands, Belgium, Italy and Finland), prioritization of personal achievement, success and self-reliance can be expected rather than group well-being and achievement, emphasized in more collectivistic settings (South Korea, China, Chile, Romania, Mexico, Turkey, Brazil and Russia), with mid-scoring countries (e.g., Spain) demonstrating balance between these two cultural values also included. There is notable variability on other cultural orientation dimensions: masculinity, power distance, uncertainty avoidance, long/short-term orientation and indulgence versus restraint (Supplementary Table 1). Although cultural orientation scores speak to culture-level effects, in our work, these have been consistently associated with child temperament and parenting variables (Gartstein & Putnam, 2018; Putnam & Gartstein, 2017).

2.2 | Measures

2.2.1 | Temperament

The ECBQ (Putnam et al., 2006) was used to measure temperament. The ECBQ is a fine-grained assessment tool based on the psychobiological definition of temperament that relies on parent ratings of observable behaviour patterns associated with temperament constructs. Factor analysis of this measure has identified a hierarchical structure, with three overarching factors encompassing 18 fine-grained scales, in turn, based on 201 items. The items are rated on a seven-point Likert scale with responses ranging from 'Never' to 'Always'. The first factor, labelled as surgency, consists of five subscales: impulsivity, activity level, high-intensity pleasure, sociability and positive anticipation. Negative emotionality includes eight subscales: discomfort, fear, motor activation, sadness, perceptual sensitivity, shyness, soothability and frustration scales. Effortful control comprises the following five subscales: inhibitory control, attention shifting, low-intensity pleasure, cuddliness and attention focusing. The ECBQ has demonstrated cross-cultural reliability and validity (Gartstein, 2019; Slobodskaya et al., 2019), consistency across time (Carranza Carnicero et al., 2000) as well as associations with behaviour problems (Gartstein et al., 2012). Translations of the ECBQ were carried out by principal investigators at the JETTC sites, with an author of the original ECBQ providing feedback on back-translated items.

ECBQ scales were computed as previously described (Putnam et al., 2006), with each fine-grained dimension calculated as an arithmetic mean of the scale items and each broad factor as a mean of the associated subscales. Internal consistency was examined separately for each culture, and items were subsequently dropped one-by-one across cultures to maximize the number of scales with $\alpha > 0.60$. Examining the internal consistency of multi-item measures is a necessary practice for assessing the validity of scales. As described by Van de Vijver and Hambleton (1996), removal of items that compromise validity in one or more cultures is a common step for equalizing measures (Van de Vijver & Hambleton, 1996). Regarding the ECBQ specifically, a similar process was followed in the

TABLE 1 Sample demographics by culture.

Culture	Child gender			Child age (in months)			Family socio-economic status (RDSI) ^a			Marital status (in percent) ^b					Maternal education (in years)			Maternal age (in years)			Number of children in the household						
	Female	Male	Range	M	SD	Range	M	SD	Range	M	SD	Ma	Lt	Di	Si	Range	M	SD	Range	M	SD	Range	M	SD	Range	M	SD
US	49	39	17-36	25.6	5.8	10-97	50.3	26.2	92	7	1	0	0	0	9-24	17.2	2.3	23-46	33.1	4.47	1-6	1.7	1				
Belgium	21	27	17-41	25.7	5.3	10-97	63.8	21.1	56	38	12	4	4	4	10-32	18.0	2.9	27-38	32.26	2.67	1-5	1.9	1				
Brazil	23	28	18-38	29.4	5.6	15-96	56.9	24.2	82	12	0	6	6	6	11-37	18.3	4.9	22-43	32.90	4.55	1-3	1.4	1				
Chile	21	28	17-41	27.3	7.2	10-97	49.7	28.3	62	15	2	21	21	21	12-28	18.1	4.9	17-41	28.54	7.11	1-4	1.8	1				
China	30	24	19-36	26.4	4.7	15-97	58.7	29.9	87	13	0	0	0	0	8-23	15.6	3.6	21-40	30.11	3.99	1-2	1.2	1				
Finland	24	31	18-40	27.6	5.7	10-97	61.6	20.8	62	30	2	6	6	6	12-26	17.7	2.6	24-41	33.57	3.87	1-4	1.5	1				
Italy	24	28	17-36	26.6	4.9	15-97	61.9	20.6	77	23	0	0	0	0	11-25	17.2	3.1	30-48	37.15	3.72	1-5	1.7	1				
Mexico	25	29	18-36	26.4	5.6	10-97	38.3	29.8	69	24	6	1	1	1	9-25	16.8	3.8	17-43	32.35	5.89	1-5	1.6	1				
Netherlands	55	64	16-40	26.6	5.8	10-87	56.6	22.3	53	40	2	5	5	5	5-25	17.7	3.7	20-41	31.99	4.27	1-3	1.6	1				
Romania	30	28	17-38	21.2	6.4	15-97	72.4	19.4	98	2	0	0	0	0	12-29	18.1	6.4	23-41	32.91	3.93	1-3	1.4	1				
Russia	26	25	17-36	27.0	5.6	15-93	62.8	19.0	77	21	2	0	0	0	10-22	14.9	2.1	21-43	29.37	5.20	1-8	1.6	1				
Spain	27	35	18-35	26.1	5.1	10-97	58.2	27.3	74	18	1	7	7	7	8-21	15.6	4.2	29-43	35.88	3.55	1-4	1.8	1				
S. Korea	26	27	17-35	28.0	4.8	15-96	51.6	24.5	100	0	0	0	0	0	7-18	15.3	2.2	29-44	34.58	3.45	1-3	1.9	1				
Turkey	25	34	16-36	27.7	5.6	10-97	50.5	26.1	92	7	1	0	0	0	9-24	14.4	3.9	19-46	31.78	5.46	1-4	1.4	1				

Note: Table adapted with permission from Gartstein et al. (2018).

Abbreviations: M, mean, SD, standard deviation.

^aRDSI, Revised Duncan Sociometric Index—An occupation-based measure of social prestige, based on maternal occupations (Stevens & Featherman, 1981).

^bMa, married; Lt, living together; Di, divorced; Si, single.

development of the short form of the ECBQ, as only items contributing to internal consistency across multiple samples were retained for the abbreviated form (Putnam et al., 2010).

After consulting values of α with individual items deleted, three items were eliminated from the activity level scale, as were two items from attention focusing and impulsivity. One item from attention shifting, low-intensity pleasure and shyness was also removed. Internal consistency for impulsivity remained below 0.60 in eight countries (i.e., further deletion of individual items did not result in an improved α). However, impulsivity items resulting in the most optimal internal consistency were utilized in computing the surgency factor score to enhance the comparability of the current findings and those obtained in prior research. ECBQ factor scores were utilized in this study, with good internal consistency across all 14 countries: Negative Affectivity (77 items) α range 0.84–0.93, average $\alpha = 0.89$; surgency (48 items) α range 0.80–0.88, average $\alpha = 0.85$ and Effortful control (57 items) α range 0.82–0.93, average $\alpha = 0.88$.

2.2.2 | Behaviour problems

The CBCL (Achenbach & Rescorla, 2000) an established parent-report questionnaire addressing behaviour problems between 18 months and 5 years of age was administered. This version of the CBCL contains 100 items that utilize the 3-point rating scale (0—not true; 1—somewhat/sometimes true; 2—very/often true) common to all Achenbach System of Empirically Based Assessment (ASEBA) instruments. The internalizing narrow-band dimensions of the scale include; emotionally reactive, anxious/depressed, somatic complaints and withdrawn, whereas attention problems and aggressive behaviour scales are associated with the externalizing domain of behaviour problems. This instrument also assesses sleep difficulties as a stand-alone scale. ASEBA instruments have seen wide-ranging international applications with translations relying on a standard procedure, for example, translating and back-translating items (Carranza Carnicero et al., 2000; Van de Vijver & Hambleton, 1996). The standard scoring approach recommended by Achenbach and Rescorla (summing symptom/problem scores for the items relevant to each scale of interest) was followed in deriving the internalizing and externalizing composite scores used as dependent variables. Reliability and validity of the CBCL scores have been established, with adequate criterion-related validity (Achenbach & Rescorla, 2000) as well as inter-rater reliability (r ranging from 0.40 to 0.75) and test–retest reliability ($r = 0.80$ s and 0.90 s). The CBCL has also been extensively utilized and validated in prior cross-cultural research (Rescorla et al., 2007). Good internal consistency for internalizing and externalizing behaviour problem indicators was demonstrated for the present sample(s), with overall satisfactory reliability for the narrow-band scales Table 2. The somatic complaints syndrome scale was not included due to poor internal consistency across cultures.

2.3 | Analytic strategy

Analyses were conducted in STATA[®] version 14. Data were analysed using a linear multilevel modelling approach. Child age and gender were included as covariates in all models to enhance the generalizability of findings to existing cross-cultural temperament studies. Models were constructed in three phases for each CBCL behaviour problem variable (i.e., three broad-band factors and five narrow dimensions) included in the present study.

First, a null model was estimated to partition within- and between-level variances and provide a standard to compare with subsequent models. Notably, the ratio of between-culture to total variance, the intraclass correlation (ICC), is also calculated in this model. For the present study, the general null model can be noted as

$$\text{Behaviour Problems}_{ij} = \gamma_{00} + u_{0j} + r_{ij}, \quad (1)$$

TABLE 2 Internal consistency reliability for Child Behaviour Checklist subscales: Cronbach's α 's.

Scale	Brazil	South Korea	Spain	Mexico	Russia	Italy	Belgium	Finland	Netherlands	Romania	China	United States	Turkey	Chile
Emotional	0.74	0.73	0.67	0.63	0.63	0.68	0.70	0.62	0.73	0.52	0.83	0.63	0.78	0.65
Anxious	0.71	0.58	0.67	0.61	0.53	0.65	0.57	0.28	0.54	0.52	0.78	0.59	0.76	0.66
Somatic	0.32	0.47	0.30	0.60	0.51	0.64	0.28	0.51	0.56	0.47	0.81	0.49	0.56	0.47
Withdrawal	0.67	0.66	0.50	0.66	0.52	0.58	0.39	0.55	0.56	0.69	0.83	0.59	0.76	0.48
Sleep	0.70	0.65	0.64	0.58	0.70	0.63	0.69	0.73	0.75	0.70	0.69	0.69	0.75	0.65
Attention	0.59	0.60	0.63	0.53	0.68	0.56	0.72	0.62	0.61	0.49	0.57	0.67	0.51	0.68
Aggression	0.90	0.89	0.85	0.84	0.86	0.87	0.86	0.80	0.91	0.90	0.89	0.87	0.85	0.94

where γ_{00} is the sample grand mean, u_{0j} is the variation of group j from the grand mean and r_{ij} is the error term associated with individual i in group j . Subsequently, child age and gender were included as covariates (Equation 2), and surgency, negative emotionality and effortful control were added to the final model (Equation 3). Importantly, because the primary interest was in assessing individual-level temperament predictors, and these were first and foremost expected to explain within-culture variability in behaviour problems, temperament factors were grand-mean-centred. This also aids in interpretation, as a value of zero is not possible for temperament ratings. Thus, regression coefficients indicate the expected change in behaviour problems given a one-unit increase in a temperament variable while fixing other temperament variables at the grand mean.

$$\text{Behaviour Problems}_{ij} = \gamma_{00} + \gamma_{01}(\text{Age}) + \gamma_{02}(\text{Gender}) + u_{0j} + r_{ij}, \quad (2)$$

$$\text{Behaviour Problems}_{ij} = \gamma_{00} + \gamma_{01}(\text{Age}) + \gamma_{02}(\text{Gender}) + \gamma_{03}(\text{Negative Emotionality}) + \gamma_{04}(\text{Surgency}) + \gamma_{05}(\text{Effortful Control}) + u_{0j} + r_{ij}. \quad (3)$$

Models were compared using a variety of fit indices, including reduction of both between- and within-level variances (i.e., change in R^2), Akaike's information criterion (AIC) (Akaike, 1987), Bayesian information criterion (BIC) (Schwarz, 1978) and chi-square difference tests. Restricted maximum likelihood (RML) estimation techniques were used for parameter estimations, as these are more realistic and accurate, particularly with a smaller number of level-2 variables (i.e., cultures; $J = 14$). However, RML is functionally limited when comparing fixed-effect parameters (Hox et al., 2017); the chi-square difference test cannot be applied to models of fixed effects using RML. As such, models were also estimated using full maximum likelihood only to calculate chi-square values to assess model fit, used to compare the deviance statistic.

3 | RESULTS

Descriptive statistics are presented in Table 3, and Table 4 provides summary statistics for all models, including variance estimates and standardized coefficients (δ), interpreted in the same manner as a beta coefficient for traditional

TABLE 3 Descriptive statistics.

Variable	M	SD	Skew	Kurtosis
ECBQ negative emotionality	2.98	0.53	0.35	3.19
ECBQ surgency	4.90	0.54	-0.09	3.33
ECBQ effortful control	4.68	0.53	0.04	3.16
CBCL internalizing	8.40	6.74	2.06	12.94
CBCL emotional	2.61	2.37	1.24	5.40
CBCL anxiety	2.35	2.19	1.32	5.72
CBCL withdrawal	1.43	1.75	2.20	11.60
CBCL externalizing	13.14	7.65	0.48	3.00
CBCL aggression	10.50	6.45	0.53	2.93
CBCL attention	2.64	1.77	0.49	2.98
CBCL sleep	3.18	2.52	0.79	3.23

Note: CBCL raw scores are presented.

Abbreviations: CBCL, Child Behaviour Checklist; ECBQ, Early Childhood Behaviour Questionnaire; M, mean, SD, standard deviation.

TABLE 4 Summary statistics for final models.

CBCL scale	Model 1 τ_{00}	Model 1 σ^2	Model 2 τ_{00}	Model 2 σ^2	ΔR^{2a} between (%)	ΔR^{2b} within (%)	Negative emotionality		Surgency/extraversion		Effortful control	
							γ	δ	γ	δ	γ	δ
Internalizing problems	4.73	40.79	2.16	30.41	54.33	25.45	6.11	0.48	-0.76	-0.06	-1.79	0.14
Emotional problems	0.32	5.35	0.28	4.21	12.50	21.31	1.88	0.42			-0.76	-0.17
Anxiety	0.68	4.12	0.31	3.21	54.41	22.09	1.93	0.47	-0.33	-0.08	-0.31	-0.07
Withdrawal	0.16	2.92	0.09	2.46	43.75	15.75	1.07	0.32	-0.49	-0.15	-0.61	-0.18
Externalizing problems	4.80	54.04	3.92	37.01	18.33	31.51	3.89	0.27	2.67	0.19	-5.78	-0.40
Attention	0.14	2.97	0.11	2.25	7.50	24.24	0.35	0.11	0.77	0.23	-1.35	-0.40
Aggression	3.56	38.24	3.09	27.27	13.20	28.69	3.54	0.29	1.90	0.16	-4.43	-0.36
Sleep problems	0.19	6.13	0.07	5.81	63.16	5.22	0.93	0.20			-0.58	-0.12

Note: γ = unstandardized coefficient, δ = standardized coefficient, τ_{00} = between-level variance, σ^2 = within-level variance. 'Model 1' refers to models with age and gender covariates and 'Model 2' refers to models including temperament factors. Only significant coefficients (i.e., $p < 0.05$) from the final model are presented.

^a'Between-culture variance (ΔR^2 between) reflects a proportional reduction in remaining variance attributed to temperament after controlling for age and gender covariates.

^b'Within-culture variance (ΔR^2 within) reflects a proportional reduction in remaining variance attributed to temperament after controlling for age and gender covariates.

regression models (i.e., in units of standard deviation, see Snijders & Bosker, 2012). Additionally, supplementary tables provide detailed model comparisons, including confidence intervals (Supplementary Tables S2–S9). For all outcomes, chi-square, AIC and BIC estimates were reduced when temperament variables were added to models, including only age and gender covariates. The proportion of variance occurring at the culture level (reflected by the null model ICC) varied across behaviour problems domains. For internalizing problems, 10.57% of the variance occurred at the culture level. There was proportionally less culture-level variance to explain for the emotional problems (5.60%) and withdrawal (5.30%) subscales but more between-culture variance in anxiety (14.28%). For externalizing problems, 8.19% of the variance occurred at the culture level. ICCs for attention and aggression scales were 4.36% and 8.65%, respectively. Finally, the ICC for sleep problems was 3.12%.

Table 4 also provides the change in R^2 when comparing the model, including age and gender covariates, to the model that accounts for temperament variables. The interpretation of R^2 differs in multilevel modelling relative to multiple regressions, reflecting the relative (i.e., proportional) difference in between-level variance statistics. Thus, the change in R^2 values discussed herein reflects the percentage reduction in between- and within-culture variances when adding temperament predictors to the previous model, which included only age and gender covariates. For example, inclusion of temperament variables reduced the between-culture variance in internalizing problems by approximately 54.33% and the within-culture variance by 25.45% compared with the model including only age and gender. Thus, temperament ratings explained over half of the between-culture variance and about a quarter of the within-culture variance that remained in the model after accounting for age and gender effects. Similarly, inclusion of temperament variables reduced the between- and within-level variances for all models, such as the externalizing problems model, reducing between-culture variance by approximately 18.33% and the within-culture variance by 31.51%. However, it should be noted that surgency did not account for significant variance in either emotional problems or sleep problems scales on the CBCL. Also of note, the only significant effect for age and gender was a positive relationship between age and aggression (i.e., older children were rated higher in aggression).

4 | DISCUSSION

Multilevel modelling was conducted with data from 14 JETTC countries addressing within- and between-cultural differences with regard to associations between temperament and behaviour problems, largely supporting a priori hypotheses. Temperament explained relatively more between- than within-culture variances in sleep problems and internalizing symptoms with the same pattern of results observed for anxious/depressed and withdrawal problems components of the internalizing broad-band score (but not emotional behaviours). In contrast, temperament variables consistently explained more within-culture variance than between-culture variance for externalizing symptoms, aggression and attention problems. In other words, temperament attributes explained how children from different cultures varied in terms of internalizing difficulties. In contrast, for externalizing difficulties, temperament accounted more for how individuals within the same culture differed from their same-culture counterparts.

Consistent with hypotheses and prior research, higher surgency predicted greater externalizing, attention and aggression problems (Gartstein et al., 2012). Similarly, exuberant children (i.e., those high in surgency) faced an increased risk for externalizing problems and disruptive behaviour disorders (de la Osa et al., 2014). However, the literature is mixed (El-Sheikh & Buckhalt, 2005; Watson et al., 2015) and additional studies addressing risk and protective functions of surgency are needed, examining these relations across cultures.

This study further supports the literature, indicating that early negative affectivity predicts internalizing problems (Kozlova et al., 2020). Higher negative emotionality also conferred risk for sleep difficulties, in line with investigations linking poor sleep with 'difficult' temperament (including distress-proneness) and behaviour problems (Gregory & Sadeh, 2016).

Lower effortful control predicted more internalizing problems, consistent with the documenting protection resulting from advanced self-regulation (Raines et al., 2019; Dolcini-Catania et al., 2020). However, the pattern of

associations between effortful control and emotional/behavioural problems is somewhat nuanced (Liu et al., 2018). More advanced attention shifting reduced the risk of anxiety problems, whereas high levels of inhibitory control conferred risk for behavioural inhibition/fear (White et al., 2011), presumably due to over-control. Thus, components of effortful control in relation to internalizing should be evaluated in the future. The association between lower effortful control and sleep problems was consistent with studies demonstrating sleep quality relations with self-regulation (Astill et al., 2012).

Our primary goal was to quantify relations between temperament and behaviour problems via a more robust statistical approach than that previously employed. This approach enabled us to identify the extent to which individual differences in temperament predict variance in behaviour problems at both the individual and cultural levels. Interestingly, the ICCs for our null models, which reflect the amount of between-culture variance associated with behaviour problems prior to introducing predictors and covariates, were similar to estimates reported by Rescorla (Rescorla et al., 2011). This pattern of results has implications related to universality versus a culture-specific nature of emerging behavioural/emotional difficulties (Bornstein, 2017) which should be considered when providing early intervention/preventative services. It may be that behaviour problem dimensions with less between-country variability, thus more universal, have stronger biological bases or are related to more widely shared socialization/parenting practices. On the other hand, problem sets with greater variability between countries, such as anxiety and attention problems, may arise largely due to culture-specific factors (e.g., unique approaches to discipline or the acceptability of certain displays of emotion).

Our results were generally consistent with the existing literature in terms of the variance explained by temperament (Rescorla et al., 2011; White et al., 2011). Importantly, individual differences in temperament explain both within- and between-culture variances, despite greater variability in behaviour problems occurring within cultures. That is, individuals within cultures share similar temperamental traits, which confer shared risk or protection with respect to various symptom sets and also explain variability in behaviour problems at the culture level. This pattern of results could be largely anticipated based on considerations related to the developmental niche framework (Majdandžić et al., 2019; Super & Harkness, 1986). Specifically, cultural values (e.g., Hofstede's cultural orientation dimensions) influence child-rearing practices, caregiver psychology and the social and physical environment. For example, parents prioritizing autonomy as a socialization goal for their children and among parental ethnotheories, prescriptions regarding optimal parenting, tended to come from cultures with higher indulgence ratings. With respect to discipline, requiring the child to think about their misbehaviour was most common in high power distance and collectivist cultures. Asking the child to 'repair damage' was positively associated with Short-term Orientation and indulgence, whereas shouting was reported more commonly in countries with low indulgence (Gartstein & Putnam, 2018).

The latter (e.g., child-rearing practices and caregiver psychology), shaped by culture, in turn influence the pathways of social-emotional development, including temperament and emerging behaviour problems, and their associations; therefore, people within a culture are more likely to share a developmental niche with others in the same culture than those outside of it. The more elements of the niche (value placed on autonomy, access to non-parental caregivers, etc.) look alike, the greater the similarity expected for these pathways. As noted, greater similarity can be expected for cultures that rank near each other on the Hofstede dimensions (e.g., individualism/collectivism, power distance; Hofstede, 2011; Hofstede & McCrae, 2004).

This study makes an important contribution to the existing literature but is not without limitations. Although multilevel modelling represents a powerful statistical technique, the analyses in the current study were somewhat limited by a relatively low number of level-2 units (i.e., countries). Including more countries in future studies would increase power and allow for analyses of random slopes to determine if the effects of temperament variables differ across cultures. Moreover, future studies should rely on the current findings for establishing a priori models to test specific random slopes, thus providing necessary replication and extension. Another limitation involved our exclusive reliance on parent-report measures in the assessment of temperament. Unfortunately, the vast majority of the cross-cultural literature is based on parent-report instruments, with notable exceptions (e.g., Gartstein, Putnam,

et al., 2016; Rubin et al., 2006). Future studies should include observational indicators, which could be obtained in the home and/or the laboratory. Finally, culturally influenced parenting practices were not considered herein and likely influence temperament development and conferred symptom risk (Lansford et al., 2018; Pham et al., 2022), thus should be considered in the future.

The present study demonstrated the feasibility of the multilevel approach and the advantages of assessing variance at both the between and within levels. Future research should employ these and similar methods (e.g., multilevel structural equation modelling and growth curve modelling). Identifying the role of culture in links between temperament and emerging behaviour problems/symptoms is essential to expanding our understanding of mechanisms resulting in risk versus protection, advancing developmental science and cross-cultural research.

5 | SUMMARY

Early temperament attributes are associated with the emergence of behaviour problems and significant long-term consequences later in life. However, these relations are rarely examined across cultures. This study addressed a critical gap, employing multilevel modelling to explain both within- and between-culture variances with respect to temperament predicting behaviour problems across 14 nations from the JETTC (Gartstein & Putnam, 2019). Controlling for the effects of child age and gender, temperament explained significant individual- and culture-level variances in sleep problems and both internalizing and externalizing problems (at the broad-band and fine-grained level). Greater negative emotionality was associated with more extensive internalizing problems, whereas higher surgency and effortful control predicted fewer internalizing difficulties. Higher negative emotionality and lower effortful control predicted more significant sleep problems in toddlerhood. Importantly, between-culture variance estimates varied among problem sets, consistent with prior research, suggesting that difficulties including anxiety and inattention may be more influenced by cultural factors than others, such as sleep difficulties. Cross-cultural studies provide unique insight into the environmental and contextual factors setting the stage for social-emotional development. Our findings have implications for prevention and intervention efforts, potentially facilitating culturally sensitive services beyond the United States.

AUTHOR CONTRIBUTIONS

Allegra Campagna: Conceptualization; formal analysis; investigation; methodology; writing – original draft; writing – review and editing. **Eric Desmarais:** Conceptualization; formal analysis; investigation; methodology; writing – original draft; writing – review and editing. **Brian French:** Formal analysis; methodology; writing – original draft; writing – review and editing. **Joshua J. Underwood:** Writing – original draft; writing – review and editing. **Mirjana Majdandžić:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Roseriet Beijers:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Carolina de Weerth:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Eun Gyoung Lee:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Blanca Huitron:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Emine Ahmetoglu:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Oana Benga:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Katri Raikkonen:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Kati Heinonen:** Conceptualization; data curation; methodology; writing – original draft; writing – review and editing. **Carmen Gonzalez-Salinas:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Helena R. Slobodskaya:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Elena Kozlova:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Maria Beatriz Martins Linhares:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Felipe Lecannelier:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Sara Casalin:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Ibrahim Acar:**

Conceptualization; data curation; investigation; methodology; writing – review and editing. **Soile Tuovinen:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Zhengyan Wang:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Rosario Montiroso:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Lorenzo Giusti:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Seong-Yeon Park:** Conceptualization; data curation; methodology; writing – review and editing. **Sae-Young Han:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Samuel Putnam:** Conceptualization; data curation; investigation; methodology; writing – review and editing. **Maria A. Gartstein:** Conceptualization; data curation; formal analysis; investigation; methodology; supervision; writing – original draft; writing – review and editing.

AFFILIATIONS

¹Department of Psychology, Washington State University, Pullman, Washington, USA

²Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, the Netherlands

³Developmental Psychology, Radboud University and Radboud University Medical Center, Nijmegen, the Netherlands

⁴Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, the Netherlands

⁵Ewha Social Science Research Institute, Ewha Womans University, Seoul, Republic of Korea

⁶Department of Psychology, National Autonomous University of Mexico, Mexico City, Mexico

⁷Department of Early Childhood Education, Trakya University in Edirne, Edirne, Turkey

⁸Department of Psychology, Babes Bolyai University, Cluj-Napoca, Romania

⁹Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland

¹⁰Department of Developmental and Educational Psychology, University of Murcia, Murcia, Spain

¹¹Research Institute of Physiology and Basic Medicine, Novosibirsk State University, Novosibirsk, Russia

¹²Neurosciences and Behavior, University São Paulo, Ribeirão Preto, Brazil

¹³Faculty of Medical Sciences, Universidad de Santiago, Santiago, Chile

¹⁴Department of Psychology, University of Leuven, Leuven, Belgium

¹⁵Department of Psychology, Özyeğin University, Istanbul, Turkey

¹⁶College of Psychology, Capital Normal University, Beijing, China

¹⁷O-3 Center for the at-Risk Infant, Scientific Institute IRCCS Eugenio Medea, Lecco, Italy

¹⁸Department of Child Development, Ewha Womans University, Seoul, Republic of Korea

¹⁹Psychology Department, Bowdoin College, Brunswick, Maine, USA

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/icd.2443>.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Allegra X. Campagna  <https://orcid.org/0000-0002-9079-008X>

Joshua J. Underwood  <https://orcid.org/0000-0002-1941-3097>

Helena Slobodskaya  <https://orcid.org/0000-0003-2076-0691>

Elena Kozlova  <https://orcid.org/0000-0003-4843-4839>

Zhengyan Wang  <https://orcid.org/0000-0003-0178-804X>

REFERENCES

- Achenbach, T. M. (1991). Manual for the Child Behavior Checklist/4-18. *Journal of Abnormal Child Psychology*, 15, 629–650.
- Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. University of Vermont, Research Center for Children, Youth, & Families.
- Achenbach, T. M., & Rescorla, L. A. (2007). *Multicultural understanding of child and adolescent psychopathology: Implications for mental health assessment*. Guilford Press.
- Achenbach, T. M., & Rescorla, L. A. (2010). *Multicultural supplement to the manual for the ASEBA preschool forms & profiles*. University of Vermont, Research Center for Children, Youth & Families.
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52(3), 317–332. <https://doi.org/10.1007/BF02294359>
- Astill, R. G., Van der Heijden, K. B., Van Ijzendoorn, M. H., & Van Someren, E. J. W. (2012). Sleep, cognition, and behavioral problems in school-age children: A century of research meta-analyzed. *Psychological Bulletin*, 138(6), 1109–1138. <https://doi.org/10.1037/a0028204>
- Beauchaine, T. P., Hinshaw, S. P., & Pang, K. L. (2010). Comorbidity of attention-deficit/hyperactivity disorder and early-onset conduct disorder: Biological, environmental, and developmental mechanisms. *Clinical Psychology: Science and Practice*, 17(4), 327–336. <https://doi.org/10.1111/j.1468-2850.2010.01224.x>
- Berdan, L. E., Keane, S. P., & Calkins, S. D. (2008). Temperament and externalizing behavior: Social preference and perceived acceptance as protective factors. *Developmental Psychology*, 44, 957–968.
- Bornstein, M. H. (2013). Parenting and child mental health: A cross-cultural perspective. *World Psychiatry*, 12(3), 258–265.
- Bornstein, M. H. (2017). The specificity principle in acculturation science. *Perspectives on Psychological Science*, 12(1), 3–45. <https://doi.org/10.1177/1745691616655997>
- Bornstein, M. H., Putnick, D. L., Gartstein, M. A., Hahn, C. S., Auestad, N., & O'Connor, D. L. (2015). Infant temperament: Stability by age, gender, birth order, term status, and socioeconomic status. *Child Development*, 86(3), 844–863. <https://doi.org/10.1111/cdev.12367>
- Brumley, L. D., & Jaffee, S. R. (2016). Defining and distinguishing promotive and protective effects for childhood externalizing psychopathology: A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, 51(6), 803–815. <https://doi.org/10.1007/s00127-016-1228-1>
- Buchanan-Pascall, S., Gray, K. M., Gordon, M., & Melvin, G. A. (2018). Systematic review and meta-analysis of parent group interventions for primary school children aged 4–12 years with externalizing and/or internalizing problems. *Child Psychiatry and Human Development*, 49(2), 244–267. <https://doi.org/10.1007/s10578-017-0745-9>
- Carranza Carnicero, J. A., Pérez-López, J., González Salinas, M. D. C., & Martínez-Fuentes, M. T. (2000). A longitudinal study of temperament in infancy: Stability and convergence of measures. *European Journal of Personality*, 14(1), 21–37. [https://doi.org/10.1002/\(SICI\)1099-0984\(200001/02\)14:1<21::AID-PER367>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1099-0984(200001/02)14:1<21::AID-PER367>3.0.CO;2-A)
- Crijnen, A. A. M., Achenbach, T. M., & Verhulst, F. C. (1997). Comparisons of problems reported by parents of children in 12 cultures: Total problems, externalizing, and internalizing. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(9), 1269–1277. <https://doi.org/10.1097/00004583-199709000-00020>
- Davis, M. M., Miernicki, M. E., Telzer, E. H., & Rudolph, K. D. (2019). The contribution of childhood negative emotionality and cognitive control to anxiety-linked neural dysregulation of emotion in adolescence. *Journal of Abnormal Child Psychology*, 47(3), 515–527. <https://doi.org/10.1007/s10802-018-0456-0>
- de la Osa, N., Granero, R., Penelo, E., Domènech, J. M., & Ezpeleta, L. (2014). The short and very short forms of the Children's Behavior Questionnaire in a community sample of preschoolers. *Assessment*, 21(4), 463–476. <https://doi.org/10.1177/1073191113508809>
- Deater-Deckard, K., Godwin, J., Lansford, J. E., Bacchini, D., Bombi, A. S., Bornstein, M. H., Chang, L., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Alampay, L. P., Uribe Tirado, L. M., Zelli, A., & Al-Hassan, S. M. (2018). Within- and between-person and group variance in behavior and beliefs in cross-cultural longitudinal data. *Journal of Adolescence*, 62, 207–217. <https://doi.org/10.1016/j.adolescence.2017.06.002>
- Delgado, B., Carrasco, M. A., González-Peña, P., & Holgado-Tello, F. P. (2018). Temperament and behavioral problems in young children: The protective role of extraversion and effortful control. *Journal of Child and Family Studies*, 27(10), 3232–3240. <https://doi.org/10.1007/s10826-018-1163-8>
- Desmarais, E., Brown, K., Campbell, K., French, B. F., Putnam, S. P., Casalin, S., Linhares, M. B. M., Lecannelier, F., Wang, Z., Raikkonen, K., Heinonen, K., Tuovinen, S., Montiroso, R., Provenzi, L., Park, S. Y., Han, S. Y., Lee, E. G., Huitron, B., de Weerth, C., ... Gartstein, M. A. (2021). Links between television exposure and toddler dysregulation: Does culture matter? *Infant Behavior and Development*, 63(March), 101557. <https://doi.org/10.1016/j.infbeh.2021.101557>
- Desmarais, E. E., French, B. F., Ahmetoglu, E., Acar, I., Gonzalez-Salinas, C., Kozlova, E., Slobodskaya, H., Benga, O., Majdandžić, M., Beijers, R., de Weerth, C., Huitron, B., Lee, E. G., Han, S. Y., Park, S. Y., Giusti, L., Montiroso, R., Tuovinen, S., Heinonen, K., ... Gartstein, M. A. (2021). Cultural contributors to negative emotionality: A multilevel

- analysis from the Joint Effort Toddler Temperament Consortium. *International Journal of Behavioral Development*, 45, 545–552. <https://doi.org/10.1177/01650254211020128>
- Dolcini-Catania, L. G., Byrne, M. L., Whittle, S., Schwartz, O., Simmons, J. G., & Allen, N. B. (2020). Temperament and Symptom Pathways to the Development of Adolescent Depression. *Journal of Abnormal Child Psychology*, 48(6), 839–849. <https://doi.org/10.1007/s10802-020-00638-3>
- Dougherty, L. R., Klein, D. N., Durbin, C. E., Hayden, E. P., & Olino, T. M. (2010). Temperamental positive and negative emotionality and children's depressive symptoms: A longitudinal prospective study from age three to age ten. *Journal of Social and Clinical Psychology*, 29(4), 462–488. <https://doi.org/10.1521/jscp.2010.29.4.462>
- Eisenberg, N., Valiente, C., Fabes, R. A., Smith, C. L., Reiser, M., Shepard, S. A., Losoya, S. H., Guthrie, I. K., Murphy, B. C., & Cumberland, A. J. (2003). The relations of effortful control and ego control to children's resiliency and social functioning. *Developmental Psychology*, 39(4), 761–776. <https://doi.org/10.1037/0012-1649.39.4.761>
- El-Sheikh, M., & Buckhalt, J. A. (2005). Vagal regulation and emotional intensity predict children's sleep problems. *Developmental Psychobiology*, 46(4), 307–317. <https://doi.org/10.1002/dev.20066>
- Evans, D. E., & Rothbart, M. K. (2007). Developing a model for adult temperament. *Journal of Research in Personality*, 41(4), 868–888. <https://doi.org/10.1016/j.jrp.2006.11.002>
- Fanti, K. A., & Henrich, C. C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: Findings from the national institute of child health and human development study of early child care. *Developmental Psychology*, 46(5), 1159–1175. <https://doi.org/10.1037/a0020659>
- Fischer, R., & Schwartz, S. (2011). Whence differences in value priorities?: Individual, cultural, or artifactual sources. *Journal of Cross-Cultural Psychology*, 42(7), 1127–1144. <https://doi.org/10.1177/00220221110381429>
- Gartstein, M. A., Carranza, J. A., González-Salinas, C., Ato, E., Galián, M. D., Erickson, N. L., & Potapova, N. (2016). Cross-cultural comparisons of infant fear: A multi-method study in Spain and the United States. *Journal of Cross-Cultural Psychology*, 47(9), 1178–1193. <https://doi.org/10.1177/0022022116663851>
- Gartstein, M. A., & Putnam, P. (2018). *Toddlers, parents, and culture: Findings from the joint effort toddler temperament consortium*. Routledge.
- Gartstein, M. A., Putnam, S., Aron, E. N., & Rothbart, M. K. (2016). Temperament and Personality. In *The Oxford Handbook of Treatment Processes and Outcomes in Counseling Psychology* (1st ed., pp. 11–41). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199739134.013.2>
- Gartstein, M. A., Putnam, S. P., & Rothbart, M. K. (2012). Etiology of preschool behavior problems: Contributions of temperament attributes in early childhood. *Infant Mental Health Journal*, 33(2), 197–211. <https://doi.org/10.1002/imhj.21312>
- Gartstein, M. A., Putnam, S. P., Slobodskaya, H. R., & de Weerth, C. (2018). The Joint Effort Toddler Temperament Consortium (JETTC): Origins, aims, and hypotheses. In M. A. Gartstein & S. P. Putnam (Eds.), *Toddlers, parents, and culture: Findings from the Joint Effort Toddler Temperament Consortium*. Routledge.
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the Revised Infant Behavior Questionnaire. *Infant Behavior and Development*, 26(1), 64–86. [https://doi.org/10.1016/S0163-6383\(02\)00169-8](https://doi.org/10.1016/S0163-6383(02)00169-8)
- Gregory, A. M., & Sadeh, A. (2016). Annual research review: Sleep problems in childhood psychiatric disorders—A review of the latest science. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 57(3), 296–317. <https://doi.org/10.1111/jcpp.12469>
- Hofstede, G. (1980). Culture and Organizations. *International Studies of Management & Organization*, 10(4), 15–41. <https://doi.org/10.1080/00208825.1980.11656300>
- Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture*, 2(1), 2307. <https://doi.org/10.9707/2307-0919.1014>
- Hofstede, G., & McCrae, R. (2004). Personality and culture revisited: Linking traits and dimensions of culture. *Cross-Cultural Research*, 38, 52–88.
- Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315650982>
- Kassing, F., Godwin, J., Lochman, J. E., Coie, J. D., Bierman, K. L., Dodge, K. A., Greenberg, M. T., McMahon, R. J., & Pinderhughes, E. E. (2019). Using early childhood behavior problems to predict adult convictions. *Journal of Abnormal Child Psychology*, 47(5), 765–778. <https://doi.org/10.1007/s10802-018-0478-7>
- Keller, H., Lamm, B., Abels, M., Yovsi, R., Borke, J., Jensen, H., Papaligoura, Z., Holub, C., Lo, W., Tomiyama, A. J., Su, Y., Wang, Y., & Chaudhary, N. (2006). Cultural models, socialization goals, and parenting ethnotheories: A multicultural analysis. *Journal of Cross-Cultural Psychology*, 37(2), 155–172. <https://doi.org/10.1177/0022022105284494>
- Kim, J., Carlson, G. A., Meyer, S. E., Bufferd, S. J., Dougherty, L. R., Dyson, M. W., Lupton, R. S., Olino, T. M., & Klein, D. N. (2012). Correlates of the CBCL-dysregulation profile in preschool-aged children. *The Journal of Child Psychology and Psychiatry*, 53(9), 918–926.
- Korhonen, M., Luoma, I., Salmelin, R., Siirtola, A., & Puura, K. (2018). The trajectories of internalizing and externalizing problems from early childhood to adolescence and young adult outcome. *Journal of Child and Adolescent Psychiatry*, 2(3), 7–12.

- Kozlova, E. A., Slobodskaya, H. R., & Gartstein, M. A. (2020). Early Temperament as a Predictor of Child Mental Health. *International Journal of Mental Health and Addiction*, 18, 1493–1506. <https://doi.org/10.1007/s11469-019-00181-3>
- Lansford, J. E., Rothenberg, W. A., Jensen, T. M., Lippold, M. A., Bacchini, D., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Uribe Tirado, L. M., Alampay, L. P., & Al-Hassan, S. M. (2018). Bidirectional relations between parenting and behavior problems from age 8 to 13 in nine countries. *Journal of Research on Adolescence*, 28(3), 571–590. <https://doi.org/10.1111/jora.12381>
- Liu, R., & Bell, M. A. (2020). Fearful temperament and the risk for child and adolescent anxiety: The role of attention biases and effortful control. *Clinical Child and Family Psychology Review*, 23(2), 205–228. <https://doi.org/10.1007/s10567-019-00306-z>
- Liu, R., Calkins, S. D., & Bell, M. A. (2018). Fearful inhibition, inhibitory control, and maternal negative behaviors during toddlerhood predict internalizing problems at age 6. *Journal of Abnormal Child Psychology*, 46(8), 1665–1675. <https://doi.org/10.1007/s10802-018-0419-5>
- Lonigan, C. J., Phillips, B. M., & Hooe, E. S. (2003). Relations of positive and negative affectivity to anxiety and depression in children: Evidence from a latent variable longitudinal study. *Journal of Consulting and Clinical Psychology*, 71(3), 465–481. <https://doi.org/10.1037/0022-006X.71.3.465>
- Majdandžić, M., Kirchoff, M. C., Räikkönen, K., Benga, O., & Ahmetoglu, E. (2019). Integrating the developmental niche: Relations among socialization goals, parental ethnotheories, daily activities, and parental responses to temperament. In *Toddlers, parents, and culture* (pp. 110–116). Routledge. <https://doi.org/10.4324/9781315203713-11>
- Montirosso, R., Cozzi, P., Putnam, S. P., Gartstein, M. A., & Borgatti, R. (2011). Studying cross-cultural differences in temperament in the first year of life: United States and Italy. *International Journal of Behavioral Development*, 35(1), 27–37. <https://doi.org/10.1177/0165025410368944>
- Oldehinkel, A. J., Hartman, C. A., De Winter, A. F., Veenstra, R., & Ormel, J. (2004). Temperament profiles associated with internalizing and externalizing problems in preadolescence. *Development and Psychopathology*, 16(2), 421–440. <https://doi.org/10.1017/s0954579404044591>
- Olino, T. M., McMakin, D. L., & Forbes, E. E. (2018). Toward an empirical multidimensional structure of anhedonia, reward sensitivity, and positive emotionality: An exploratory factor analytic study. *Assessment*, 25(6), 679–690. <https://doi.org/10.1177/1073191116680291>
- Ostlund, B. D., Measelle, J. R., Laurent, H. K., Conrath, E., & Ablow, J. C. (2017). Shaping emotion regulation: Attunement, symptomatology, and stress recovery within mother–infant dyads. *Developmental Psychobiology*, 59(1), 15–25. <https://doi.org/10.1002/dev.21448>
- Pham, C., Desmarais, E., Jones, V., French, B. F., Wang, Z., Putnam, S., Casalin, S., Martins Linhares, M. B., Lecannelier, F., Tuovinen, S., Heinonen, K., Raikkonen, K., Montirosso, R., Giusti, L., Park, S., Lee, E. G., Huitron, B., de Weerth, C., Beijers, R., ... Gartstein, M. A. (2022). Relations between bedtime parenting behaviors and temperament across 14 cultures. *Frontiers in Psychology*, 13, 7350.
- Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2012). Control networks and neuromodulators of early development. *Developmental Psychology*, 48(3), 827–835. <https://doi.org/10.1037/a0025530>
- Putnam, S. P., & Gartstein, M. A. (2017). Aggregate temperament scores from multiple countries: Associations with aggregate personality traits, cultural dimensions, and allelic frequency. *Journal of Research in Personality*, 67, 157–170.
- Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior and Development*, 29(3), 386–401. <https://doi.org/10.1016/j.infbeh.2006.01.004>
- Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and assessment of short and very short forms of the infant behavior questionnaire-revised. *Journal of Personality Assessment*, 96(4), 445–458. <https://doi.org/10.1080/00223891.2013.841171>
- Putnam, S. P., Jacobs, J., Gartstein, M. A., & Rothbart, M. K. (2010). Development and assessment of short and very short forms of the Early Childhood Behavior Questionnaire. *Poster presented at International Conference on Infant Studies, Baltimore, MD.*
- Raines, E. M., Viana, A. G., Trent, E. S., Woodward, E. C., Candelari, A. E., Zvolensky, M. J., & Storch, E. A. (2019). Effortful control, interpretation biases, and child anxiety symptom severity in a sample of children with anxiety disorders. *Journal of Anxiety Disorders*, 67(January), 102136. <https://doi.org/10.1016/j.janxdis.2019.102136>
- Reijntjes, A., Kamphuis, J. H., Prinzie, P., Boelen, P. A., Van Der Schoot, M., & Telch, M. J. (2011). Prospective linkages between peer victimization and externalizing problems in children: A meta-analysis. *Aggressive Behavior*, 37(3), 215–222. <https://doi.org/10.1002/ab.20374>
- Ren, H., Hart, C. H., Cheah, C. S. L., Porter, C. L., Nelson, D. A., Yavuz-Müren, H. M., Gao, W., Haron, F., Jiang, L., Kawashima, A., Shibazaki-Lau, A., Nakazawa, J., Nelson, L. J., Robinson, C. C., Selçuk, A. B., Evans-Stout, C., Tan, J.-P.,

- Yang, C., Quek, A.-H., & Zhou, N. (2023). Parenting measurement, normativeness, and associations with child outcomes: Comparing evidence from four non-Western cultures. *Developmental Science*, e13388.
- Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., Bird, H., Chen, W., Dobrea, A., Döpfner, M., Erol, N., Fombonne, E., Fonseca, A., Frigerio, A., Grietens, H., Hannesdottir, H., Kanbayashi, Y., Lambert, M., Larsson, B., ... Verhulst, F. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders*, 15(3), 130–142. <https://doi.org/10.1177/10634266070150030101>
- Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Harder, V. S., Otten, L., Bilenberg, N., Bjarnadottir, G., Capron, C., De Pauw, S. S. W., Dias, P., Dobrea, A., Döpfner, M., Duyme, M., Eapen, V., Erol, N., Esmaeili, E. M., Ezpeleta, L., Frigerio, A., Fung, D. S. S., ... Verhulst, F. C. (2011). International comparisons of behavioral and emotional problems in preschool children: Parents' reports from 24 societies. *Journal of Clinical Child and Adolescent Psychology*, 40(3), 456–467. <https://doi.org/10.1080/15374416.2011.563472>
- Riem, M. M. E., Van Ijzendoorn, M. H., Parsons, C. E., Young, K. S., De Carli, P., Kringelbach, M. L., & Bakermans-Kranenburg, M. J. (2017). Experimental manipulation of infant temperament affects amygdala functional connectivity. *Cognitive, Affective, & Behavioral Neuroscience*, 17(4), 858–868. <https://doi.org/10.3758/s13415-017-0518-8>
- Rothbart, M., & Derryberry, D. (1981). Development of individual differences in temperament. *Advances in Developmental Psychology*, 1, 37–86.
- Rothbart, M. K., & Hwang, J. (2005). Temperament and the development of competence and motivation. In *Handbook of competence and motivation* (pp. 167–184).
- Rubin, K. H., Hemphill, S. A., Chen, X., Hastings, P., Sanson, A., Coco, A. L., Zappulla, C., Chung, O.-B., Park, S.-Y., Doh, H. S., Chen, H., Sun, L., Yoon, C.-H., & Cui, L. (2006). A crosscultural study of behavioral inhibition in toddlers: East–West–North–South. *International Journal of Behavioral Development*, 30(3), 219–226.
- Sánchez-Pérez, N., Fuentes, L. J., Eisenberg, N., & González-Salinas, C. (2018). Effortful control is associated with children's school functioning via learning-related behaviors. *Learning and Individual Differences*, 63(February), 78–88. <https://doi.org/10.1016/j.lindif.2018.02.009>
- Sánchez-Pérez, N., Putnam, S. P., Gartstein, M. A., & González-Salinas, C. (2020). ADHD and ODD symptoms in toddlers: Common and specific associations with temperament dimensions. *Child Psychiatry and Human Development*, 51(2), 310–320. <https://doi.org/10.1007/s10578-019-00931-3>
- Schwartz, C. E., Snidman, N., & Kagan, J. (1999). Adolescent social anxiety as an outcome of inhibited temperament in childhood. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(8), 1008–1015. <https://doi.org/10.1097/00004583-199908000-00017>
- Schwartz, C. E., Wright, C. I., Shin, L. M., Kagan, J., & Rauch, S. L. (2003). Inhibited and uninhibited infants “grown up”: Adult amygdalar response to novelty. *Science*, 300(5627), 1952–1953. <https://doi.org/10.1126/science.1083703>
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Slobodskaya, H., Kozlova, E., Han, S.-Y., Gartstein, M. A., & Putnam, S. P. (2019). Cross-cultural differences in temperament. In *Toddlers, parents, and culture* (pp. 29–37). Routledge. <https://doi.org/10.4324/9781315203713-3>
- Slobodskaya, H. R., Gartstein, M. A., Nakagawa, A., & Putnam, S. P. (2013). Early temperament in Japan, the United States, and Russia: Do cross-cultural differences decrease with age? *Journal of Cross-Cultural Psychology*, 44(3), 438–460. <https://doi.org/10.1177/0022022112453316>
- Snijders, T., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage Publishers.
- Stevens, G., & Featherman, D. L. (1981). A revised socioeconomic index of occupational status. *Social Science Research*, 10(4), 364–395. [https://doi.org/10.1016/0049-089X\(81\)90011-9](https://doi.org/10.1016/0049-089X(81)90011-9)
- Stifter, C. A., Putnam, S., & Jahromi, L. (2008). Exuberant and inhibited toddlers: Stability of temperament and prediction to behavior problems. *Development and Psychopathology*, 20, 401–421.
- Super, C. M., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development*, 9(4), 545–569.
- Van de Vijver, F., & Hambleton, R. K. (1996). Translating tests. *European Psychologist*, 1(2), 89–99. <https://doi.org/10.1027/1016-9040.1.2.89>
- Verhulst, F. C., & Achenbach, T. M. (1995). Empirically based assessment and taxonomy of psychopathology: Cross-cultural applications. A review. *European Child & Adolescent Psychiatry*, 4(2), 61–76. <https://doi.org/10.1007/BF01977734>
- Watson, D., Stasik, S. M., Ellickson-Larew, S., & Stanton, K. (2015). Extraversion and psychopathology: A facet-level analysis. *Journal of Abnormal Psychology*, 124(2), 432–446. <https://doi.org/10.1037/abn0000051>
- Weeks, M., Ploubidis, G. B., Cairney, J., Wild, T. C., Naicker, K., & Colman, I. (2016). Developmental pathways linking childhood and adolescent internalizing, externalizing, academic competence, and adolescent depression. *Journal of Adolescence*, 51, 30–40. <https://doi.org/10.1016/j.adolescence.2016.05.009>

White, L. K., McDermott, J. M., Degnan, K. A., Henderson, H. A., & Fox, N. A. (2011). Behavioral inhibition and anxiety: The moderating roles of inhibitory control and attention shifting. *Journal of Abnormal Child Psychology*, 39(5), 735–747. <https://doi.org/10.1007/s10802-011-9490-x>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Campagna, A. X., Desmarais, E. D., French, B., Underwood, J. J., Majdandžić, M., Beijers, R., de Weerth, C., Lee, E. G., Huitron, B., Ahmetoglu, E., Benga, O., Raikkonen, K., Heinonen, K., Gonzalez-Salinas, C., Slobodskaya, H., Kozlova, E., Linhares, M. B. M., Lecannelier, F., Casalin, S., ... Gartstein, M. A. (2023). Temperament and behaviour problems in children: A multilevel analysis of cross-cultural differences. *Infant and Child Development*, 32(5), e2443. <https://doi.org/10.1002/icd.2443>