

# 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  
© The Author(s) 2012  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0022022112453316  
jccp.sagepub.com  


Helena R. Slobodskaya<sup>1</sup>, Maria A. Gartstein<sup>2</sup>,  
Atsuko Nakagawa<sup>3</sup>, and Samuel P. Putnam<sup>4</sup>

## Abstract

The present study addressed differences in infant and toddler temperament, utilizing translations of the Infant Behavior Questionnaire–Revised (IBQ-R) and the Early Childhood Behavior Questionnaire (ECBQ), for children growing up in the United States, Russia, and Japan. Results indicated a number of significant differences in higher-order dimensions and fine-grained components of early temperament between the three cultural groups. U.S. children scored higher for Surgency and related traits, compared to Japanese and Russian children; Negative Affectivity showed the opposite pattern of cross-cultural differences, wherein Japanese children received the highest scores from their caregivers. In addition, Japanese infants and toddlers scored lower for Effortful Control. Significant Culture  $\times$  Age interactions indicated that patterns of cross-cultural differences in different age groups varied across and within the three higher-order dimensions. Surgency, as well as positive affect to both low and high levels of intensity, showed a consistent pattern of decreasing cultural differences with age.

## Keywords

temperament, development, infancy, toddlerhood, cross-cultural differences

From the first days of life children differ in the ways they behave (Chess, Thomas, & Birch, 1965). These consistent patterns of behaviors (movements, facial expressions, vocalizations) along with the accompanying physiological processes across different situations (feeding, sleeping, playing, and other daily activities) have come to be known as temperament (Strelau, 1998). The psychobiological approach proposed by Rothbart and Derryberry (1981) has conceptualized temperament as constitutionally based individual differences in reactivity and self-regulation,

---

<sup>1</sup>FSBI Institute of Physiology SB RAMS, Novosibirsk State University, Novosibirsk, Russia

<sup>2</sup>Washington State University, Pullman, Washington, USA

<sup>3</sup>Nagoya City University, Nagoya, Japan

<sup>4</sup>Bowdoin College, Brunswick, Maine, USA

## Corresponding Author:

Helena R. Slobodskaya, FSBI Institute of Physiology SB RAMS, Timakova, 4, Novosibirsk, 630117, Russia.

Email: [hslob@physiol.ru](mailto:hslob@physiol.ru)

influenced by heredity, maturation, and experience. Reactivity refers to arousability of affect, motor activity, and attentional responses (i.e., orienting), assessed by threshold, latency, intensity, time to peak intensity, and recovery time of the reaction. Self-regulation refers to processes such as behavioral inhibition and self-soothing, serving to modulate reactivity (Rothbart & Bates, 1998).

Multiple assessment tools have been developed for the evaluation of temperament, with parent-report gaining most widespread use. Despite potential limitations of parent-report instruments (Kagan, 1994), they offer developmentally appropriate assessment with superior predictive validity relative to other temperament measures, including structured observations (Hart, Field, & Roitfarb, 1999; Pauli-Pott, Mertesacker, Bade, Haverkock, & Beckmann, 2003). Factor-analytic research provided evidence for a three-factor structure of temperament from infancy through early adolescence, including Surgency, Negative Affectivity, and Regulatory Capacity/Effortful Control (Rothbart, 2007). The content of the factors in the infant and toddler measures utilized in the current investigation are largely similar, with 10 of the 11 scales included on both measures loading on analogous factors. Scales found on one instrument but not the other largely reflect changes in the behavioral repertoire of the two ages. For instance, a vocal reactivity scale reflecting spontaneous babbling is found only on the infant questionnaire, whereas scales representing the ability to willfully control attention and behavior are included only on the early childhood instrument (see Putnam, Ellis, & Rothbart, 2001).

Cross-cultural research addressing temperament in the first few years of life has not been widespread, in contrast to the large-scale studies of personality that have demonstrated the robustness of the five-factor model from early childhood to adulthood across languages, cultures, and political systems (De Fruyt et al., 2009; Kohnstamm, Halverson, Mervielde, & Havill, 1998; McCrae et al., 2005). It should be noted that recent empirical work provided evidence for conceptual links between three higher-order childhood temperament traits of Surgency, Negative Affectivity, and Effortful Control and the Big Five personality traits of Extraversion, Neuroticism, and Conscientiousness, respectively (Caspi, Roberts, & Shiner, 2005; Rothbart, 2007). Therefore, cross-cultural temperament research could draw on the relevant findings in the area of personality. It has been shown that distribution of personality traits across cultures is organized geographically: European and American cultures differed from Asian and African cultures on two distinct dimensions linked to Extraversion and Neuroticism (Allik & McCrae, 2004). Studies of personality using peer and self-reports showed that for Extraversion, U.S. adults scored higher than Japanese and Russian participants, whereas for Neuroticism differences were in the opposite direction, and for Conscientiousness, results were not unequivocal (McCrae et al., 2005; Schmitt et al., 2007).

Cross-cultural temperament research has also revealed a number of mean-level differences in fine-grained traits between contrasting cultures. Using observational techniques, Chinese infants were found to be more fearful, timid, and more likely to cry intensely when separated from mothers than Caucasian infants (Kagan, Reznick, & Snidman, 1986), and Chinese toddlers were observed to have higher levels of behavioral inhibition than Canadian toddlers (Chen et al., 1998). Using parent ratings, Taiwanese infants were found to be less active, approachable, and adaptable, and more negative in mood than U.S. infants (Hsu, Soon, Stigler, Hong, & Liang, 1981); Japanese preschoolers were found to be more withdrawal-oriented, less flexible, and to express less positive affect than U.S. children (Windle, Iwawaki, & Lerner, 1988); and Russian infants were found to have lower levels of Smiling/Laughter, High and Low Intensity Pleasure, Perceptual Sensitivity and Vocal Reactivity, and higher levels of Distress to Limitations compared to their U.S. counterparts (Gartstein, Slobodskaya, & Kinsht, 2003).

Recently reported data showed that mean personality scores in samples of countries correlated significantly and substantially with Hofstede's dimensions of culture and a number of

socioeconomic indicators (Hofstede & McCrae, 2004). Among the three personality factors closely linked to child temperament, Extraversion was positively related to individualism (reflecting self-interest over group interest) and negatively related to power distance (reflecting acceptance of status differences). Neuroticism was associated with uncertainty avoidance (reflecting tolerance for ambiguity) and masculinity, and Conscientiousness was positively related to power distance and national poverty. There are two possible explanations for the observed cross-cultural differences in mean levels of personality traits. The biological explanation emphasizes heritability and the genetic transmission of temperament predispositions as the underlying mechanism and rests on the idea that the distributions of temperament genotypes differ among populations that vary in ethnicity, giving rise to cross-cultural differences in temperament and personality traits (Allik & McCrae, 2004). Environmental pathway models, on the other hand, stress the role of cultural values and corresponding patterns of childrearing that tend to be relatively stable, socializing children into phenotypical presentations of temperament that are desirable, appropriate, or at least tolerable within cultural norms (Kohnstamm, 1989).

Super and Harkness (1986) conceptualized the interface between the child and culture as the “developmental niche,” described as a function of (1) customs, especially those related to child rearing, (2) settings available to the child, and (3) caregiver psychosocial characteristics. All three factors are influenced by culture and cause children growing up in a particular country to acquire similar characteristics in the process of their development. Individualism-collectivism represents the most commonly applied construct in explaining cultural differences (Triandis & Suh, 2002). Socialization in collectivistic cultures has been described as focusing on emotional warmth/proximity that fosters acceptance of the group’s norms and values (Keller et al., 2004). Caregivers in collectivistic societies often respond to their infants’ needs in an anticipatory manner, blurring the self-other distinction, whereas caregivers in individualistic cultures tend to encourage the expression of positive emotions and focus on early self-regulation (Greenfield, Keller, Fuligni, & Maynard, 2003; Keller et al., 2004). Thus, culturally preferred phenotypes accepted by the caregivers may vary as a function of the individualistic versus collectivistic orientation of a given society, leading to differing parenting behavior, socialization practices, and subsequent cross-cultural differences in the development of individual differences in the experience and expression of both positive and negative emotions.

Evidence is accumulating that both biological and cultural explanations for cross-cultural differences have to be considered and development can be understood as the acquisition of cultural knowledge based in universal biological predispositions (Keller, 2008).

### *The Present Study*

The present study was designed to explore differences in higher order dimensions and fine-grained components of infant and toddler temperament in Japan, Russia, and the United States, using translations of the same parent-report instruments and addressing interactions between culture, gender, and age. A set of cultures known to vary in terms of important characteristics were selected. The decision to study infants and toddlers was based on a particular dearth of cross-cultural developmental research in early childhood, and because we anticipated that important cultural differences may emerge in the infancy-to-toddlerhood transition, negotiated differently depending on the cultural context, which shapes parental childrearing attitudes and behaviors as babies mature into toddlers.

The three countries included in this study are among the 10 most populous countries in the world with higher than 70% of urban population (Haub, 2010). Russia and the United States are the first and third largest countries in the world with areas 17,075 and 9,827 sq. km, respectively, whereas Japan has 378 sq. km. Russia is one of the most northern countries (61 degrees north of

the equator), whereas Japan and the United States are at much lower latitude of 36-37 degrees. Consequently, Russia is the coldest country in the world, with an average annual temperature below zero, whereas Japan and the United States are countries with mostly temperate climate and average annual temperature 14-15 °C (<http://www.climatetemp.info/>). At the time of data collection, gross domestic product (GDP) per capita was 33.63 for Japan, 45.59 for the United States, and 14.69 for Russia. Human Development Index (HDI), a composite measure of life expectancy at birth, literacy, educational attainment, and GDP per capita, was very high in Japan and the United States, with a ranking of 10 and 13, respectively; in Russia HDI was high, with a much lower ranking of 71 (Human Development Report, 2009). In terms of cultural dimensions, the United States has received the highest score for individualism on Hofstede's index, whereas Japan and Russia score in the middle range (Allik & Realo, 2004). Russia is among the highest ranking cultures on power distance (Hofstede, 2001). Japan has the highest ranking for masculinity, whereas the United States is moderate and Russia relatively low in masculinity. Both Japan and Russia have high rankings for uncertainty avoidance.

Drawing on prior research addressing cross-cultural differences in child and adult temperament and personality as well as associations between personality traits and dimensions of culture, geographic variables, and socioeconomic indicators (Allik & McCrae, 2004; Hofstede & McCrae, 2004; McCrae et al., 2005), we anticipated cultural effects for the three higher-order dimensions and the corresponding fine-grained components of infant and toddler temperament. Specifically, U.S. children were expected to score higher for Surgency and related traits, relative to Japanese and Russian children. For Negative Affectivity, the opposite pattern of cross-cultural differences was expected, with Japanese and Russian children scoring higher relative to U.S. children. No predictions were made about cultural effects for Effortful Control or mean-level differences between Japanese and Russian children, since the results of previous studies have been inconclusive. Although a priori predictions regarding possible interactions among culture, gender, and age could not be specified, they were examined to assess whether there is a gradual increase in cross-cultural differences from 3 months to 3 years of age, as predicted by the environmental model because of increased exposure, or whether the cross-cultural differences in temperament were most prominent early in the first year of life, more consistent with predictions based on the biological model. This study was expected to contribute to developmental cross-cultural psychology of individual differences since the two samples belong to contrasting cultural contexts, Western (European American) and East Asian (Japanese), whereas the third sample is from an Asian part of the most Eastern European country (Russia).

## Method

### Participants

*Japan.* A community sample of 284 mothers of infants (45% female) between 3 and 12 months of age were recruited in Nagoya, Japan's third largest industrial metropolis located near the centre of Japan. Parents, visiting public health centers for routine 3-month medical examinations, were initially recruited to take part in the temperament investigation and divided into three age groups. Questionnaires were handed to the parents of the 3- to 6-month-old group during the visits and were mailed to the homes at the appropriate time for the participants included in the 6- to 9- and 9- to 12-month age groups. In addition, questionnaires were disseminated through the local nursery to mothers whose babies were between 3 and 12 months of age. Mean age for the target children was 31.2 weeks ( $SD = 11.5$ ), ranging from 14 to 51 weeks of age.

The toddler sample included caregiver (97% mothers) reports for 318 children (49% female) from Aichi and Fukui Prefectures; Aichi Prefecture includes urban areas, whereas Fukui is more

rural. Data were collected in metropolitan Nagoya (Aichi prefecture), the commuter town of Konan in Aichi Prefecture, and the provincial town of Fukui in Fukui Prefecture. Questionnaires were distributed to mothers or other caregivers who attended medical check-ups for 18-month-old children at public health centers. Parents of older children received questionnaires by mail. Mean age for the target children was 22.7 months ( $SD = 5.6$ ), ranging from 18 to 36 months of age.

*Russia.* A community sample of 325 primary caregivers (94% mothers) of infants (47% female) between 3 and 12 months of age were recruited via public healthy child clinics. Questionnaires were distributed to parents during the visits for routine medical examinations in the first year of life. In addition, mothers whose babies were between 3 and 12 months of age were identified by research assistants and undergraduate students and approached in person. Mean age for the target children was 29.8 weeks ( $SD = 12.0$ ), ranging from 11 to 54 weeks of age.

The toddler sample included caregiver reports (93% mothers) for 229 children (52% female). Participants were approached via healthy child clinics, day care centers, and identified by research assistants and undergraduate students, who collected data as a class assignment. Parents whose children were between 18 and 36 months of age were recruited and provided with a questionnaire packet, which they completed and returned to the investigators. Mean age for the target children was 24.7 months ( $SD = 4.9$ ), ranging from 18 to 36 months of age. The Russian data were collected in Novosibirsk, the third biggest city in Russia and the business centre of Siberia, and in nearby regions.

*United States.* A community sample of 530 primary caregivers (90% mothers) of infants (50% female) between 3 and 12 months of age were recruited on the basis of birth announcements published in local newspapers, with information provided by birth centers, and through the parents' participation in local "Birth to 3" courses. Data were collected at two different locations: Eugene-Springfield, Oregon, and the San Francisco Bay Area, California. At Eugene-Springfield, primary caregivers of infants between 3 and 12 months of age were recruited; whereas at the San Francisco Bay Area, only parents of infants who were 3, 6, 9, or 12 months of age (plus or minus 2 weeks) were invited to participate. Mean age for the target children was 32.7 weeks ( $SD = 12.0$ ), ranging from 11 to 53 weeks of age.

The toddler sample included caregiver reports (98% mothers) for 319 children (52% female). Data were collected at Eugene-Springfield, Oregon, and the majority of participants were recruited between the ages of 3 and 12 months on the basis of birth announcements in the local newspaper and participation in local "Birth to 3" courses. The caregivers were contacted by telephone when their children were at least 18 months of age and asked to participate in the follow-up assessment. In addition, parents whose children were between 18 and 36 months of age were recruited through birth announcements. Mean age for the target children was 25.0 months ( $SD = 3.7$ ), ranging from 18 to 36 months of age. The U.S. participants were recruited by telephone and provided with a questionnaire packet by mail. Reminder calls and letters were used to contact parents who did not return questionnaires.

Although recruitment strategies had to be adapted to the cultural contexts of the three participating countries, these were expected to provide generally representative samples. In both Russia and Japan, as in the United States, all parents, regardless of their SES/affluence, are expected to receive regular wellness check-ups for their children.

## Materials and Procedure

*Infant Behavior Questionnaire-Revised (IBQ-R; Gartstein & Rothbart, 2003).* This 191-item parent-report instrument yields 14 scales that form three overarching factors: a Surgency factor containing Activity Level, Approach, High Intensity Pleasure, Perceptual Sensitivity, Smiling and Laughter, and Vocal Reactivity scales; a Negative Affectivity (NA) factor including Distress to

Limitations, Fear, Sadness, and negatively loading Falling Reactivity; and a Regulatory Capacity factor including Duration of Orienting, Soothability, Cuddliness/Affiliation, and Low Intensity Pleasure. Reliability and validity of the IBQ-R has been supported for samples from the United States (Gartstein & Rothbart, 2003), Japan (Nakagawa & Sukigara, 2005), and Russia (Gartstein et al., 2003), with mean alpha coefficients of .82 (range .77 to .89), .79 (range .61 to .92), and .86 (range .75 to .92), respectively. Comparisons of IBQ-R structure in Japanese, Russian, and U.S. samples have provided evidence for similarity of three overarching factors (Gartstein, Knyazev, & Slobodskaya, 2005; Nakagawa & Sukigara, 2005).

*Early Childhood Behavior Questionnaire (ECBQ; Putnam, Gartstein, & Rothbart, 2006).* This 201-item measure yields 18 scales that form three higher-order factors: Surgency, made up of Impulsivity, Activity Level, High-intensity Pleasure, Sociability, and Positive Anticipation; Negative Affectivity (NA), including Discomfort, Fear, Sadness, Frustration, Motor Activation, Perceptual Sensitivity, Shyness, and negatively loading Soothability; and Effortful Control (EC), containing Inhibitory Control, Attention Shifting, Low-Intensity Pleasure, Cuddliness, and Attention Focusing. Reliability and validity of the ECBQ have been supported for samples from the United States (Putnam et al., 2006), Russia (Kolmagorova, Slobodskaya, & Gartstein, 2008), and Japan (Nakagawa, Sukigara, & Mizuno, 2007), with mean alpha coefficients of .79 (range .61 to .89), .77 (range .63 to .89), and .81 (range .74 to .93), respectively. Comparison of ECBQ structure in Japanese, Russian, and U.S. samples provided evidence for similarity of three higher-order factors (Kolmagorova et al., 2008; Nakagawa et al., 2007).

*Procedure.* After the participating caregivers agreed to take part in this investigation and provided informed consent, infant caregiver samples were asked to complete the IBQ-R and toddler caregiver samples were administered the ECBQ. The questionnaires were administered in person, or provided to caregivers to be completed at home, at their convenience, to be returned to investigators at a later time, by mail or in-person. The completion of these paper-and-pencil measures generally takes about 1 hour. This research was approved by the respective Institutional Review Boards of the investigators' institutions.

## Results

### *Mean Differences Between Cultures*

To assess mean differences between the three cultures (Japan, Russia, and United States), as well as Culture  $\times$  Age and Culture  $\times$  Gender interactions, four three-way MANOVAs were performed with the IBQ-R and ECBQ higher-order temperament factors and fine-grained subscales. For IBQ-R, participants were divided into three age groups: 3 to 6 months ( $N = 371$ ), 6 to 9 months ( $N = 362$ ), and 9 to 12 months ( $N = 403$ ). For ECBQ, participants were divided into three age groups: 18 to 22 months ( $N = 366$ ), 22 to 27 months ( $N = 258$ ), and 27 to 36 months ( $N = 242$ ).<sup>1</sup> The composition of the cultural samples by gender and age groups is given in Table 1.

Significant main effects of culture were followed up by simple contrasts comparing the different cultural groups. The effect size was estimated using eta-squared ( $\eta^2$ ).

*IBQ-R.* A significant multivariate culture effect emerged for the three higher-order temperament scales (Wilks's  $\lambda = .82$ ,  $p < .001$ ,  $\eta^2 = .095$ ) and 14 subscales (Wilks's  $\lambda = .64$ ,  $p < .001$ ,  $\eta^2 = .197$ ). Table 2 presents the statistical results of all main effects and two-way and three-way interactions for each measure; descriptive statistics along with cross-cultural differences are shown in Table 3.

Largest effects of culture were found for Surgency and all subscales contributing to this factor, except Activity Level. U.S. infants' scores were the highest; Japanese infants exhibited the lowest levels of Approach, High Intensity Pleasure, Vocal Reactivity, and for overall Surgency,

**Table 1.** Composition of the Samples by Age Group and Gender

Age Group (in Months)	Sample					
	Japan		Russia		U.S.	
	M	F	M	F	M	F
3 to 6 <sup>a</sup>	58	39	54	56	82	83
6 to 9 <sup>a</sup>	40	39	51	52	93	87
9 to 12 <sup>a</sup>	57	51	67	44	88	97
18 to 22 <sup>b</sup>	90	95	41	54	40	46
22 to 27 <sup>b</sup>	29	25	39	37	65	63
27 to 36 <sup>b</sup>	42	37	29	29	49	56

Note. M = males; F = females.

a. Infant Behavior Questionnaire–Revised.

b. Early Childhood Behavior Questionnaire.

with Russian infants' scores mostly in the middle. The effects were moderated by gender for the Surgency factor and High Intensity Pleasure subscale such that girls' scores significantly differed across the three cultures, whereas Russian and Japanese boys were similar to one another and lower than U.S. male youngsters on these traits. These results also indicated that U.S. boys scored significantly higher on High Intensity Pleasure compared to girls, whereas gender differences in Japanese and Russian youngsters were insignificant (Figure 1A).

Significant Culture  $\times$  Age interactions also indicated that age moderated culture effects for Surgency and four of six scales contributing to this factor: Activity, Approach, High Intensity Pleasure, and Perceptual Sensitivity. As shown in Table 4, for Surgency, cross-cultural differences found in the total sample held in the youngest age group only; Japanese and Russian infants' scores in two older groups did not differ. For Activity, the effect of culture was significant only in the youngest age group, wherein U.S. infants' scores were higher than Japanese. For Approach and High Intensity Pleasure, the effect of culture was significant only in two younger groups (Figure 2A). For Perceptual Sensitivity, Japanese and U.S. infants' scores in the oldest group did not differ. Overall, the effect of culture for Surgency and associated IBQ-R scales was largest in the youngest age group.

Culture effects for Negative Affectivity (NA) showed the reverse pattern: U.S. infants' scores were the lowest for all subscales contributing to this factor, including reversed Falling Reactivity (Table 2). Japanese infants exhibited the highest levels of Distress to Limitations, and overall NA, and lowest Falling Reactivity. Russian infants' scores for NA and Distress to Limitations were in the middle, relative to the other two cultural groups. Russian infants did not significantly differ from Japanese youngsters on Fear and Sadness; compared to the U.S. infants, they were significantly higher for Sadness. The effect for Distress to Limitations was moderated by age, with the cross-cultural differences found in the total sample observed for the oldest age group only. Japanese and Russian infants in youngest group did not differ and were higher than the U.S. infants, whereas in the middle age group the effect of culture was not significant (Table 4).

Culture effects for Regulatory Capacity were much smaller and lacking a clear pattern of cross-cultural differences (Table 2). For Duration of Orienting, Russian infants' scores were the highest and Japanese infants' scores the lowest. Infants' scores for Soothability were similar across the samples. For Cuddliness and Low Intensity Pleasure, Japanese and Russian infants did not differ and were higher than the U.S. infants for the former and lower for the latter.

**Table 2.** Effects of Culture, Gender, Age, and Their Interactions on IBC-R Scales

Temperament Variable	Interactions													
	Culture		Gender		Age		C × G		C × A		G × A		C × G × A	
	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>
<b>Surgency</b>	<b>57.96***</b>	<b>.094</b>	<b>.002</b>	<b>.000</b>	<b>185.33***</b>	<b>.249</b>	<b>4.81**</b>	<b>.009</b>	<b>8.67***</b>	<b>.030</b>	<b>.66</b>	<b>.001</b>	<b>.74</b>	<b>.003</b>
Activity level	.88	.002	2.71	.002	84.41***	.132	.58	.001	4.33**	.015	.60	.001	1.21	.004
Approach	58.47***	.096	.05	.000	196.90***	.263	2.91	.005	16.24***	.055	.27	.000	.71	.003
High intensity pleasure	40.56***	.068	5.40*	.005	130.72***	.191	4.43*	.008	13.40***	.046	2.83	.005	.68	.002
Perceptual sensitivity	24.43***	.042	2.84	.003	79.96***	.126	2.21	.004	2.99*	.011	.38	.001	1.35	.005
Smiling & laughter	15.65***	.028	.35	.000	6.48**	.012	2.70	.005	2.21	.008	2.24	.004	.36	.001
Vocal reactivity	58.37***	.095	1.44	.001	59.91***	.098	.28	.001	1.57	.006	1.50	.003	1.07	.004
<b>Negative affectivity</b>	<b>24.96***</b>	<b>.043</b>	<b>1.28</b>	<b>.001</b>	<b>39.72***</b>	<b>.066</b>	<b>.54</b>	<b>.001</b>	<b>.51</b>	<b>.002</b>	<b>.22</b>	<b>.000</b>	<b>.55</b>	<b>.002</b>
Distress to limitations	15.41***	.027	6.37*	.006	80.76***	.127	.68	.001	4.93**	.018	.63	.001	.36	.001
Fear	19.05***	.033	23.31***	.021	55.47***	.091	1.83	.003	1.67	.006	.72	.001	.88	.003
Sadness	8.86***	.016	.04	.000	.88	.002	1.03	.002	1.42	.005	1.75	.003	.91	.003
Falling reactivity	15.70***	.028	.14	.000	5.16**	.009	1.21	.002	2.16	.008	.12	.000	.33	.001
<b>Regulatory capacity</b>	<b>4.53**</b>	<b>.008</b>	<b>1.44</b>	<b>.001</b>	<b>7.42**</b>	<b>.013</b>	<b>1.48</b>	<b>.003</b>	<b>5.70***</b>	<b>.020</b>	<b>.136</b>	<b>.000</b>	<b>.10</b>	<b>.000</b>
Duration of orienting	24.72***	.043	.64	.001	2.79	.005	2.88	.005	2.04	.007	.71	.001	.19	.001
Soothability	.01	.000	.07	.000	1.56	.003	.13	.000	1.39	.005	.96	.002	.42	.002
Cuddliness	7.40**	.013	.41	.000	46.36***	.077	.24	.000	2.97*	.011	1.81	.003	.74	.003
Low intensity pleasure	19.05***	.033	.97	.001	2.68	.005	.29	.001	7.01***	.025	.69	.001	.48	.002

Note. C = culture; G = gender; A = age. Error degrees of freedom for F statistics were 1,119 for the three higher-order temperament scales and 1,106 for 14 subscales. \*p < .05. \*\*p < .01. \*\*\*p < .001.

**Table 3.** Descriptive Statistics and Cross-Cultural Comparisons for the IBQ-R Scales

Temperament Variable	Japan (N = 284)		Russia (N = 325)		U.S. (N = 530)		Multiple Comparisons <sup>a</sup>
	M	SD	M	SD	M	SD	
<b>Surgency</b>	<b>4.27</b>	<b>.86</b>	<b>4.45</b>	<b>.81</b>	<b>4.77</b>	<b>.63</b>	<b>A&gt;J,R; R&gt;J</b>
Activity level	4.19	1.01	4.15	1.11	4.26	.86	—
Approach	4.15	1.49	4.84	1.36	4.98	1.07	A,R>J
High intensity pleasure	5.30	1.14	5.54	1.30	5.89	.70	A>J,R; R>J
Perceptual sensitivity	3.57	1.20	3.53	1.35	4.03	1.07	A>J,R
Smiling and laughter	4.42	1.00	4.37	1.14	4.74	.94	A>J,R
Vocal reactivity	3.99	1.01	4.31	1.10	4.71	.91	A>J,R; R>J
<b>Negative affectivity</b>	<b>3.52</b>	<b>.68</b>	<b>3.40</b>	<b>.81</b>	<b>3.17</b>	<b>.67</b>	<b>J&gt;A,R; R&gt;A</b>
Distress to limitations	3.99	.92	3.82	.92	3.62	.88	J>A,R; R>A
Fear	3.06	1.15	2.94	1.33	2.61	1.00	J,R>A
Sadness	3.62	.94	3.77	1.12	3.49	.90	R>A
Falling reactivity	4.60	.94	4.91	1.10	5.03	.99	A,R>J
<b>Regulatory capacity</b>	<b>4.66</b>	<b>.54</b>	<b>4.84</b>	<b>.76</b>	<b>4.77</b>	<b>.65</b>	<b>A,R&gt;J</b>
Duration of orienting	3.53	.99	4.13	1.18	3.75	1.03	R>A,J; A>J
Soothability	4.80	.82	4.83	1.12	4.81	1.03	—
Cuddliness	5.65	.59	5.64	.89	5.44	1.00	J,R>A
Low intensity pleasure	4.67	.95	4.73	1.22	5.07	.90	A>J,R

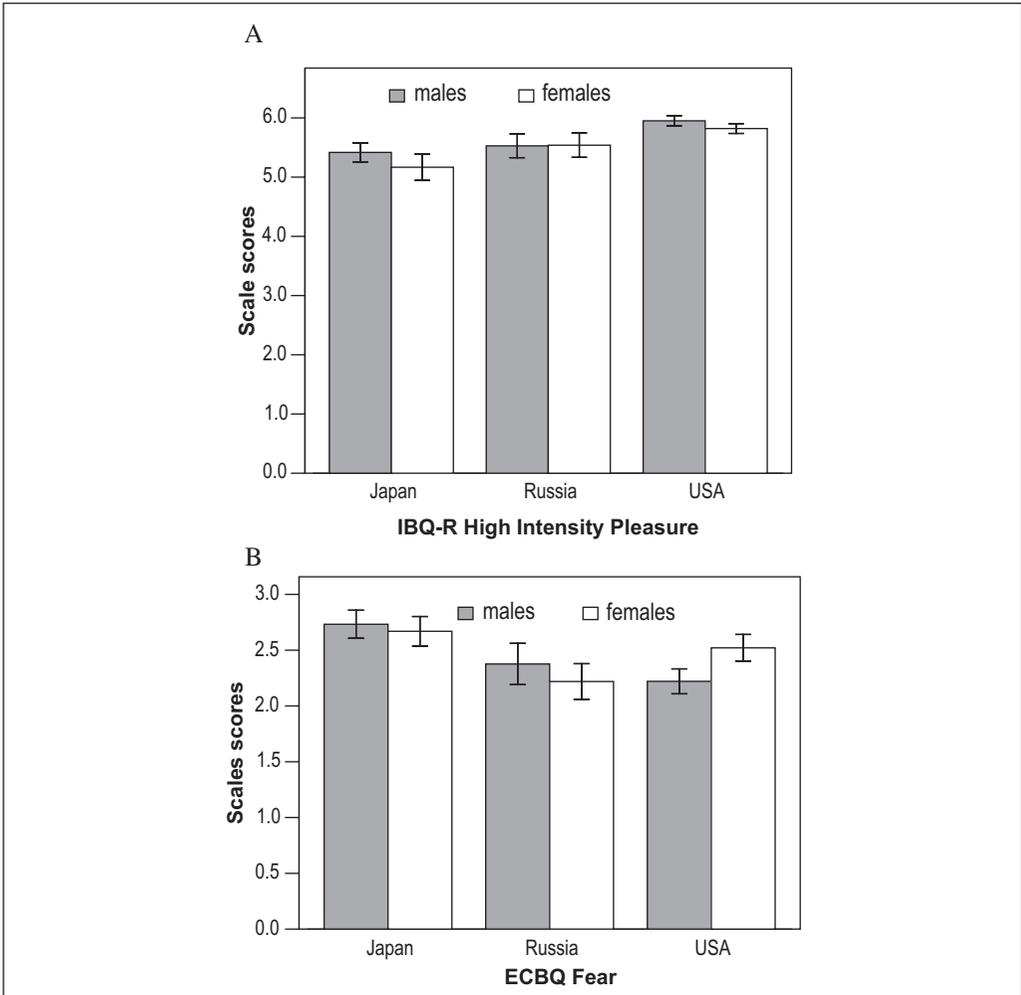
Note. a. Direction of culture effects when pairwise *t* test was significant at  $p < .05$ ; for example, J>A,R indicates that Japanese infants scored higher than U.S. and Russian infants.

Age moderated the culture effect for Regulatory Capacity and two subscales contributing to this factor, Cuddliness and Low Intensity Pleasure. Table 4 shows that for Regulatory Capacity and Low Intensity Pleasure, cross-cultural differences found in the total sample held in the youngest age group, whereas culture effect for Cuddliness was largest in the oldest group (Figure 2B).

Overall, the effects of culture on IBQ-R higher-order temperament scales accounted for 9% of the variance in Surgency, 4% of the variance in Negative Affectivity, and less than 1% of the variance in Regulatory Capacity. Among the subscales, the largest effects were found for Approach and Vocal Reactivity, accounting for up to 10% of the variance. Interactions between culture and age accounted for 3% of the variance in Surgency, 2% of the variance in Regulatory Capacity, and much less than 1% of the variance in Negative Affectivity. Among the subscales, the largest moderating effects of age were found for Approach and High Intensity Pleasure, accounting for around 5% of the variance. Interactions between culture and gender were significant for two scales, Surgency and High Intensity Pleasure, accounting for less than 1% of the variance.

**ECBQ.** Significant multivariate culture effect emerged for the three higher-order temperament scales (Wilks's  $\lambda = .95$ ,  $p < .001$ ,  $\eta^2 = .027$ ) and 18 subscales (Wilks's  $\lambda = .58$ ,  $p < .001$ ,  $\eta^2 = .241$ ). Table 5 presents the statistical results of all main effects and two-way and three-way interactions for each measure; descriptive statistics along with cross-cultural differences are shown in Table 6.

For the Surgency factor and High Intensity Pleasure and Sociability subscales, U.S. toddlers' scores were the highest; Japanese toddlers exhibited the lowest levels of High Intensity Pleasure, whereas Russian toddlers exhibited the lowest levels of Activity. For Sociability



**Figure 1.** Effects of Gender and Culture on Temperament Traits: A – IBQ-R High Intensity Pleasure, B – ECBQ Fear in Boys and Girls of Different Cultures

Note. Bars are 95% confidence intervals.

and the overarching Surgency dimension, Japanese and Russian toddlers' scores did not differ significantly. For the Negative Affectivity factor and Fear, Shyness, and Sadness subscales, Japanese toddlers' scores were the highest, while Russian toddlers exhibited the highest levels of Discomfort and the lowest levels of Motor Activation, Perceptual Sensitivity, and Shyness. U.S. toddlers' scores were the lowest for Soothability. Japanese toddlers' scores were the lowest, whereas Russian and U.S. toddlers' scores did not differ for the Effortful Control factor and Low-Intensity Pleasure, Cuddliness, and Attention Focusing subscales. For Attention Shifting, cross-cultural differences were in the opposite direction: Japanese toddlers scored higher than Russian and U.S. toddlers. For Inhibitory Control, U.S. toddlers' scores were the highest, while Japanese and Russian toddlers did not differ.

The effects of culture were moderated by gender for the Negative Affectivity factor and Fear subscale such that Japanese boys were significantly higher than U.S. and Russian boys on these traits, whereas Japanese girls were higher than Russian but did not differ from U.S. girls; for

**Table 4.** Culture Effects on IBQ-R Scales in Different Age Groups<sup>a</sup>

Temperament Variable	3-6 Months		6-9 Months		9-12 Months	
	F <sup>b</sup>	η <sup>2</sup>	F <sup>c</sup>	η <sup>2</sup>	F <sup>d</sup>	η <sup>2</sup>
<b>Surgency</b>	<b>36.11***(A&gt;J,R; R&gt;J)</b>	<b>.164</b>	<b>13.65*** (A&gt;J,R)</b>	<b>.071</b>	<b>9.85*** (A&gt;J,R)</b>	<b>.047</b>
Activity level	5.38** (A>J)	.029	.48	.003	1.77	.009
Approach	52.65*** (A,R>J)	.226	19.63*** (A,R>J)	.100	.76	.004
High intensity pleasure	30.24*** (A>J,R; R>J)	.143	5.97** (A>J,R)	.033	2.26 (A>J)	.011
Perceptual sensitivity	16.34*** (A>J,R)	.083	5.69** (A>J,R)	.031	7.15** (A>R)	.035
Smiling and laughter	6.21** (A>J,R)	.033	3.91* (A,J>R)	.022	10.00*** (A>J,R)	.048
Vocal reactivity	14.80*** (A,R>J)	.076	23.70*** (A>J,R)	.118	24.77*** (A> R,J;R>J)	.110
<b>Negative affectivity</b>	<b>8.81*** (J,R&gt;A)</b>	<b>.046</b>	<b>7.53** (J,R&gt;A)</b>	<b>.040</b>	<b>10.52*** (J&gt;A,R;R&gt;A)</b>	<b>.050</b>
Distress to limitations	4.22* (J,R>A)	.023	1.42	.008	22.75*** (J>A,R;R>A)	.102
Fear	2.59 (R>A)	.014	9.31*** (J,R>A)	.050	9.80*** (J>A,R)	.047
Sadness	2.95 (J>A)	.016	4.40* (R>A)	.024	3.85* (R>A,J)	.019
Falling reactivity	14.29*** (A>R,J; R>J)	.073	3.86* (A>J)	.021	2.52 (R>J)	.012
<b>Regulatory capacity</b>	<b>10.31*** (A,R&gt;J)</b>	<b>.053</b>	<b>.17</b>	<b>.001</b>	<b>6.12** (R&gt;A,J)</b>	<b>.030</b>
Duration of orienting	8.41*** (A,R>J)	.044	3.43* (R>J)	.019	20.02*** (R>A,J)	.091
Soothability	1.47	.008	.05	.000	1.61	.008
Cuddliness	1.11	.006	2.31 (J>A)	.013	8.67*** (J,R>A)	.042
Low intensity pleasure	22.91*** (A>J,R)	.113	7.49*** (A,J>R)	.040	1.05	.005

Note. a. Direction of culture effects when pairwise t test was significant at  $p < .05$ ; for example, J>A,R indicates that Japanese infants scored higher than U.S. and Russian infants.

b. Degrees of freedom were 2,369 for the three higher-order temperament scales and 2,361 for the 14 subscales.

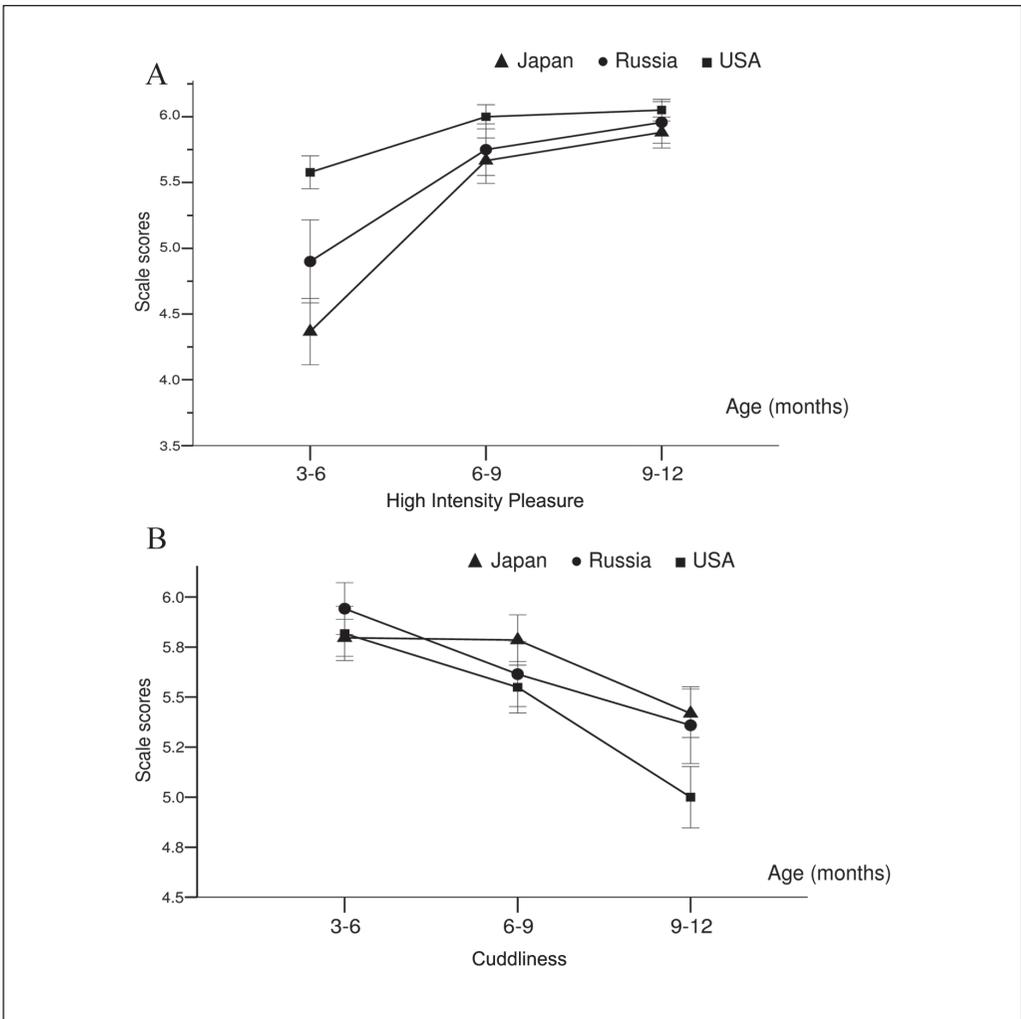
c. Degrees of freedom were 2,359 for the three higher-order temperament scales and 2,355 for the 14 subscales.

d. Degrees of freedom were 2,401 for the three higher-order temperament scales and 2,400 for the 14 subscales.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$  (all two-tailed tests).

Fear, U.S. girls were also higher than Russian girls. Follow-up analyses of significant Culture × Gender interactions also indicated that U.S. boys scored significantly lower on ECBQ Negative Affectivity and Fear compared to girls, whereas gender differences in Japanese and Russian youngsters were insignificant (Figure 1B).

Significant Culture × Age interactions also indicated that age moderated culture effects for Impulsivity, Sociability, and Shyness subscales. As shown in Table 7, for Impulsivity, although the effect of culture was not significant in the total sample and in two older groups, there were significant cross-cultural differences in the youngest age group: Russian toddlers' scores were the lowest, while Japanese and U.S. toddlers' scores did not differ. For Sociability, cross-cultural differences found in the total sample held in the youngest age group only (Figure 3A), and for Shyness, the effect of culture was significant only in two younger groups (Figure 3B).



**Figure 2.** Effects of Age and Culture on IBQ-R Temperament Subscales: A – High Intensity Pleasure, B – Cuddliness in Japanese, Russian, and U.S. Infants of Different Age Groups  
 Note. Bars are 95% confidence intervals.

Overall, the effects of culture on ECBQ higher-order temperament scales accounted for 1.2% of the variance in Surgency and slightly more than 2% of the variance in Negative Affectivity and Effortful Control. Among the subscales, the largest effects were found for Discomfort (5.6% of the variance), Fear (4.4% of the variance), and Soothability (3.9% of the variance). Interactions between culture and age accounted for less than 2% of the variance, with the largest effects for Impulsivity (1.5% of the variance) and Shyness (1.3% of the variance). Interactions between culture and gender were significant for the Negative Affectivity factor and Fear subscale, accounting for less than 2% of variance.

### *Age Differences in Cultural Effects*

To examine whether cultural effects are stable over age, we used eta values that reflect the correlations between the grouping variable and the outcomes (Huberty, 2002). Procedures described

**Table 5.** Effects of Culture, Gender, Age, and Their Interactions on ECBQ Scales

Temperament Variable	Interactions													
	Culture		Gender		Age		C × G		C × A		G × A		C × G × A	
	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>	F	η <sup>2</sup>
<b>Surgency</b>	<b>5.06**</b>	<b>.012</b>	<b>3.23</b>	<b>.004</b>	<b>2.34</b>	<b>.005</b>	<b>1.54</b>	<b>.004</b>	<b>1.10</b>	<b>.005</b>	<b>.41</b>	<b>.001</b>	<b>2.08</b>	<b>.010</b>
Impulsivity	.31	.001	.16	.000	.46	.001	.52	.001	3.15*	.015	1.08	.003	.32	.001
Activity level	9.70***	.023	8.97**	.011	2.88	.007	.72	.002	.22	.001	1.04	.002	1.76	.008
High intensity pleasure	12.93***	.030	15.63***	.018	.58	.001	.76	.002	1.35	.006	.88	.002	2.19	.010
Sociability	4.49**	.011	.11	.000	7.49**	.017	.09	.000	2.39*	.011	.95	.002	.40	.002
Positive anticipation	.22	.001	.02	.000	17.05***	.039	2.43	.006	2.29	.011	.31	.001	.22	.007
<b>Negative affectivity</b>	<b>10.77***</b>	<b>.025</b>	<b>.65</b>	<b>.001</b>	<b>5.08**</b>	<b>.012</b>	<b>5.10**</b>	<b>.012</b>	<b>.28</b>	<b>.001</b>	<b>.48</b>	<b>.001</b>	<b>1.51</b>	<b>.007</b>
Discomfort	24.76***	.056	.13	.000	5.43**	.013	2.14	.005	1.15	.005	1.19	.003	1.13	.005
Fear	19.27***	.044	.08	.000	2.59	.006	7.32**	.017	.89	.004	.18	.000	3.36**	.016
Sadness	12.33***	.028	1.15	.001	.77	.002	.43	.001	.42	.002	.74	.002	1.29	.006
Frustration	2.54	.006	6.12*	.007	.41	.001	1.34	.003	.17	.001	.24	.001	.13	.001
Motor activation	8.80***	.020	6.14*	.007	.56	.001	2.59	.006	.77	.004	.22	.001	2.52*	.012
Perceptual sensitivity	7.96**	.019	.20	.000	8.66***	.020	1.59	.004	1.09	.005	.94	.002	3.43**	.016
Shyness	7.52**	.018	5.85*	.007	.74	.002	.81	.002	2.70*	.013	.93	.002	1.09	.005
Soothability	16.97***	.039	.07	.000	5.92	.014	1.99	.005	.86	.004	.05	.000	.33	.002
<b>Effortful control</b>	<b>9.56***</b>	<b>.022</b>	<b>13.75***</b>	<b>.016</b>	<b>20.38***</b>	<b>.046</b>	<b>.39</b>	<b>.001</b>	<b>1.04</b>	<b>.005</b>	<b>1.52</b>	<b>.004</b>	<b>.39</b>	<b>.002</b>
Inhibitory control	20.81***	.047	14.15***	.017	8.77***	.020	1.58	.004	1.82	.009	1.27	.003	.27	.001
Attention shifting	5.54**	.013	2.31	.003	6.11**	.014	.17	.000	1.37	.006	.47	.001	1.77	.008
Low intensity pleasure	12.20***	.028	13.11***	.015	1.31	.003	.30	.001	.44	.002	1.25	.003	.34	.002
Cuddliness	14.76***	.034	6.89**	.008	8.55***	.020	.05	.000	.49	.002	.99	.002	.54	.003
Attention focusing	3.63*	.009	.24	.000	26.27***	.059	.22	.001	1.80	.008	2.12	.005	.67	.003

Note. C = culture; G = gender; A = age. Error degrees of freedom for F statistics were 848 for the three higher-order temperament scales and 842 for 18 subscales.

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

in Lipsey and Wilson (2001) were followed, with Fisher’s z-transformed eta value weighted by the inverse of the variance. Mean effect sizes and confidence intervals were computed using SPSS macros (Wilson, 2006) and converted back to correlations after all analyses were performed. Table 8 shows effect sizes of culture for eight subscales that appear in both the IBQ-R and ECBQ: Activity, High Intensity Pleasure, Perceptual Sensitivity, Fear, Sadness, Soothability, Cuddliness, and Low Intensity Pleasure, along with three higher-order factor values and mean effect sizes for temperament subscales. The heterogeneity tests (*Q* statistic) indicated significant differences among the effect sizes for High Intensity Pleasure, Low Intensity Pleasure, and Soothability subscales, as well as for Surgency and Effortful Control higher-order dimensions.

We then tested age of child and temperament questionnaire as moderators in weighted multiple regression analyses using SPSS macros created by Wilson (2006). For High Intensity Pleasure, the fixed effects model was significant,  $Q_B(2) = 14.21, p < .001$ , and the residual  $Q_E(3) = 5.08, p = .17$ , indicated that the remaining variability across effect sizes was homogeneous. Both moderators accounted for significant variability in effect size: cultural differences decreased with age ( $B = -.10, p < .001$ ), although for ECBQ they were higher than for IBQ-R ( $B = .24, p < .01$ ). For Surgency, the fixed effects model was also significant,  $Q_B(2) = 25.55, p < .001$ , and the residual  $Q_E(3) = 1.378, p = .71$ , indicated that the remaining variability across effect sizes was homogeneous. Age was a significant predictor of effect size ( $B = -.08, p < .01$ ), indicating that cultural differences decreased with age, whereas the temperament questionnaire moderator did not account for a significant amount of variance. Figure 4 graphically demonstrates this decrease in the effect of culture at older ages.

For Low Intensity Pleasure, the fixed effects model was significant,  $Q_B(2) = 8.62, p < .05$ , but the residual variability was not homogeneous,  $Q_E(3) = 8.07, p < .05$ ; thus, we used the random

**Table 6.** Descriptive Statistics and Cross-Cultural Comparisons for the ECBQ Scales

Temperament Variable	Japan (N = 318)		Russia (N = 229)		U.S. (N = 319)		Multiple Comparisons <sup>a</sup>
	M	SD	M	SD	M	SD	
<b>Surgency</b>	<b>4.83</b>	<b>.60</b>	<b>4.84</b>	<b>.79</b>	<b>5.01</b>	<b>.54</b>	<b>A&gt;J,R</b>
Impulsivity	5.09	.75	4.99	.96	5.04	.63	—
Activity level	4.76	.85	4.50	.96	4.83	.76	A,J>R
High intensity pleasure	4.42	.94	4.68	1.17	4.84	.97	A>J,R; R>J
Sociability	5.08	1.21	5.20	1.17	5.44	.91	A>J,R
Positive anticipation	4.81	1.04	4.82	1.10	4.89	.85	—
<b>Negative affectivity</b>	<b>3.03</b>	<b>.54</b>	<b>2.87</b>	<b>.63</b>	<b>2.86</b>	<b>.51</b>	<b>J&gt;A,R</b>
Discomfort	2.50	.68	2.89	.98	2.39	.85	R>A,J
Fear	2.70	.82	2.29	.93	2.38	.76	J>A,R
Sadness	3.10	.88	2.75	.97	2.71	.88	J>A,R
Frustration	3.48	.91	3.37	1.18	3.57	.89	—
Motor activation	2.04	.68	1.74	.98	1.93	.65	A,J>R
Perceptual sensitivity	4.09	1.04	3.81	1.19	4.03	1.05	A,J>R
Shyness	3.48	1.00	3.14	.96	3.30	1.05	A,J>R; J>A
Soothability	5.13	1.01	5.05	1.05	5.43	.71	A>J,R
<b>Effortful control</b>	<b>4.36</b>	<b>.55</b>	<b>4.55</b>	<b>.67</b>	<b>4.64</b>	<b>.56</b>	<b>A,R&gt;J</b>
Inhibitory control	3.51	.99	3.61	1.26	4.12	.94	A>J,R
Attention shifting	4.68	.66	4.56	.86	4.52	.62	J>A,R
Low intensity pleasure	4.65	.74	4.91	.92	4.96	.71	A,R>J
Cuddliness	4.84	.72	5.24	1.01	5.19	.82	A,R>J
Attention focusing	4.12	1.08	4.45	.98	4.40	.86	A,R>J

Note. a. Direction of culture effects when pairwise *t* test was significant at  $p < .05$ ; for example, J>A,R indicates that Japanese toddlers scored higher than US and Russian toddlers.

effects model with maximum likelihood estimation. The results showed that the model was significant,  $Q_B(2) = 6.31, p < .05$ , with insignificant residual variance,  $Q_E(3) = 6.45, p = .09$ . Age was a significant predictor of effect size ( $B = -.07, p < .05$ ), indicating that cultural differences decreased with age, whereas the temperament questionnaire moderator was not significant. For Soothability, the fixed effects model was significant,  $Q_B(2) = 10.69, p < .01$ , and the residual,  $Q_E(3) = 3.09, p = .38$ , indicated that the remaining variability across effect sizes was homogeneous. However, neither age nor questionnaire accounted for a significant amount of variance ( $B = .02$  and  $.08$ , respectively,  $p > .05$ ). When we compared the mean effect size for IBQ-R and ECBQ for Soothability, the between-groups homogeneity statistic was significant,  $Q_B(1) = 10.09, p = .002$ , indicating that variability in cultural effects was explained by significant differences between the IBQ-R (mean ES =  $.06$ ; CI =  $.00$  to  $.14$ ) and ECBQ (mean ES =  $.20$ ; CI =  $.12$  to  $.27$ ). For Effortful Control higher-order dimension, neither the fixed effects model,  $Q_B(2) = 1.72$  nor the random effects model,  $Q_B(2) = .93$ , were significant,  $p > .05$ , indicating that age of child and temperament questionnaire were not significant predictors of effect size (Figure 4B).

## Discussion

This study provides evidence of cross-cultural differences between infants and toddlers in Japan, Russia, and the United States in higher-order dimensions and fine-grained components of

**Table 7.** Culture Effects on ECBQ Scales in Different Age Groups<sup>a</sup>

Temperament Variable	18-22 Months		22-27 Months		27-36 Months	
	F <sup>b</sup>	η <sup>2</sup>	F <sup>c</sup>	η <sup>2</sup>	F <sup>d</sup>	η <sup>2</sup>
<b>Surgency</b>	<b>4.94** (A&gt;J,R)</b>	<b>.027</b>	<b>2.00</b>	<b>.015</b>	<b>.19</b>	<b>.002</b>
Impulsivity	5.82** (A,J>R)	.032	.54	.004	.72	.006
Activity level	4.98** (A,J>R)	.027	3.17* (A>R)	.024	3.06* (A>R)	.025
High intensity pleasure	7.87*** (A>J)	.042	5.04** (A,R>J)	.038	2.00	.016
Sociability	8.33*** (A>J,R)	.044	1.36	.011	.13	.001
Positive anticipation	1.26	.007	.51	.004	3.92* (J>A)	.032
<b>Negative affectivity</b>	<b>6.98** (J&gt;A,R)</b>	<b>.037</b>	<b>2.96 (J&gt;A,R)</b>	<b>.023</b>	<b>2.76 (J&gt;A)</b>	<b>.023</b>
Discomfort	20.71*** (J,R>A;R>J)	.104	5.35** (R>A)	.040	4.33** (R>A)	.035
Fear	11.43*** (J>A,R)	.060	6.08** (J>A,R)	.046	5.91** (J>A,R)	.047
Sadness	9.60*** (J>A,R)	.051	2.40 (J>R)	.019	4.11* (J>A)	.033
Frustration	1.06	.006	.37	.003	1.45	.012
Motor activation	6.32** (A,J>R)	.034	5.88** (A,J>R)	.044	.66	.005
Perceptual sensitivity	2.12 (J>R)	.012	5.32** (A,J>R)	.040	2.22	.018
Shyness	7.12** (J>A,R)	.038	5.32** (A,J>R)	.040	1.11	.009
Soothability	4.30* (A>J,R)	.023	7.84*** (A>R)	.058	7.08** (A>J,R)	.056
<b>Effortful control</b>	<b>10.58*** (A,R&gt;J)</b>	<b>.055</b>	<b>.84</b>	<b>.007</b>	<b>3.90* (A,R&gt;J)</b>	<b>.032</b>
Inhibitory control	11.23*** (A,R>J)	.059	10.44*** (A>J,R)	.076	5.83** (A>J,R)	.046
Attention shifting	1.30	.007	4.37* (J>A,R)	.033	2.52	.021
Low intensity pleasure	7.88*** (A,R>J)	.042	1.75	.014	5.83** (A,R>J)	.047
Cuddliness	6.07** (A,R>J)	.033	3.15* (A,R>J)	.024	7.80** (A,R>J)	.061
Attention focusing	6.96** (A,R>J)	.037	2.16 (R>A)	.017	.97	.008

Note. a. Direction of culture effects when pairwise *t* test was significant at  $p < .05$ ; for example, J>A,R indicates that Japanese toddlers scored higher than US and Russian toddlers.

b. Degrees of freedom were 2,363 for the three higher-order temperament scales and 2,358 for the 14 subscales.

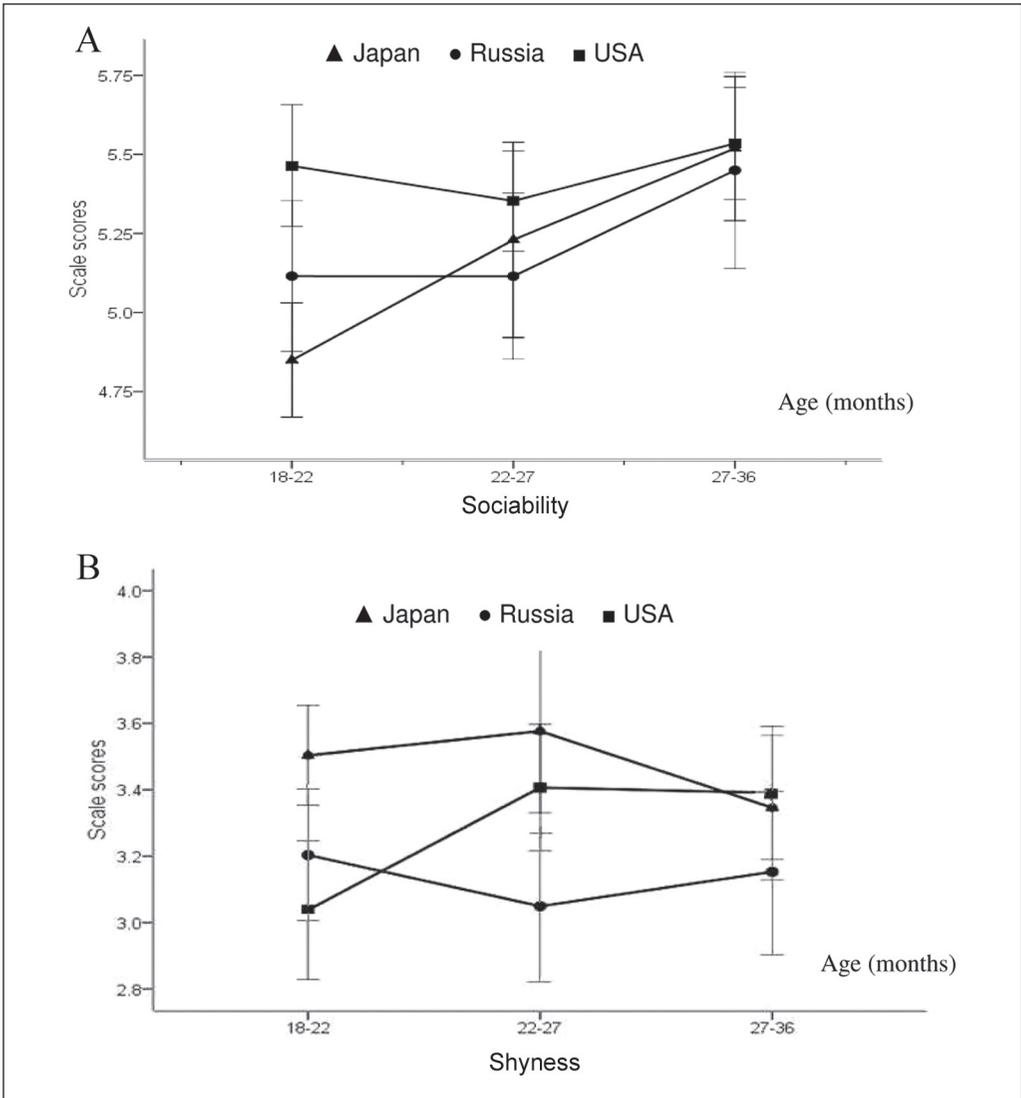
c. Degrees of freedom were 2,255 for the three higher-order temperament scales and 2,254 for the 14 subscales.

d. Degrees of freedom were 2,239 for the three higher-order temperament scales and 2,239 for the 14 subscales.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$  (all two-tailed tests).

temperament, as well as the moderating role of gender and age with respect to the effects of culture. Our findings supported the prediction that U.S. children would score higher for the overarching dimension of Surgency and related traits, relative to Japanese and Russian children. In infancy, these differences were significant for five of the six fine-grained traits contributing to this factor; in toddlerhood, for three of the five. In addition, Russian infants exhibited higher levels of Surgency and three of the six related traits than Japanese infants, whereas in toddlerhood, Japanese and Russian children's scores were mostly similar. Consistent with our predictions, the Negative Affectivity factor produced an opposite pattern, wherein U.S. infants' scores for this higher-order dimension and all related traits were the lowest, and Japanese infants scores indicated greater Negative Affectivity at the dimension level and on two of the four fine-grained traits. In toddlerhood, Japanese children had the highest scores for this overarching temperament dimension and three of its eight fine-grained components.

These observed differences are consistent with previous findings demonstrating that U.S. children expressed more positive affect (smiling, laughter, pleasure, approach) and less negative affect (crying, fear, shyness) relative to East Asian (Chen et al., 1998; Hsu et al., 1981; Kagan et al., 1986; Windle et al., 1988) and Russian children (Gartstein et al., 2003). Results from prior studies were obtained from both parent reports and behavioral observations, so they do not appear to be simply artifacts of potential evaluative biases. Our findings are also consistent with



**Figure 3.** Effects of Age and Culture on ECBQ Temperament Subscales: A – Sociability, B – Shyness in Japanese, Russian, and U.S. Toddlers of Different Age Groups  
*Note.* Bars are 95% confidence intervals.

the results of adult and adolescent personality studies showing that Americans are higher in Extraversion and lower in Neuroticism, relative to participants from Russia and Japan (McCrae et al., 2005, 2010). For the third higher-order dimension of temperament, Effortful Control, both in infancy and in toddlerhood, Japanese children had the lowest scores, paralleling the surprising findings that Japanese score very low on the Conscientiousness scale (McCrae, 2002).

To our knowledge, no prior study compared temperament of Japanese and Russian children, but relevant evidence from cross-cultural personality research should be considered in this context. For example, Allik and McCrae's (2004) analyses of self-reports from 36 cultures around the world showed that Japanese and Russians had very similar personality profiles, although Japanese were slightly lower in Extraversion and slightly higher in Neuroticism than Russians (McCrae, 2002). By contrast, in observer ratings from 51 cultures, Japanese were higher in

**Table 8.** Estimates of Culture Effects for Temperament Traits Across Age Groups

Temperament Variable	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Q
	ES (95% CI)	ES (95% CI)	ES (95% CI)	ES (95% CI)	ES (95% CI)	ES (95% CI)	
Activity level	.17 (.07 to .27)	.06 (-.05 to .16)	.10 (-.00 to .19)	.16 (.06 to .26)	.16 (.03 to .27)	.16 (.03 to .28)	4.01
High intensity pleasure	.38 (.29 to .46)	.18 (.08 to .28)	.10 (.01 to .20)	.20 (.10 to .30)	.20 (.07 to .31)	.13 (-.00 to .25)	19.29**
Perceptual sensitivity	.29 (.19 to .38)	.17 (.07 to .27)	.19 (.09 to .28)	.11 (.01 to .21)	.20 (.08 to .31)	.13 (.01 to .26)	7.30
Fear	.12 (.02 to .22)	.22 (.12 to .32)	.22 (.12 to .31)	.24 (.15 to .34)	.21 (.09 to .33)	.22 (.09 to .33)	3.82
Sadness	.13 (.02 to .22)	.16 (.05 to .25)	.14 (.04 to .23)	.23 (.13 to .32)	.14 (.02 to .26)	.18 (.06 to .30)	2.59
Soothability	.09 (-.01 to .19)	.00 (-.10 to .10)	.09 (-.01 to .19)	.15 (.05 to .25)	.24 (.12 to .35)	.24 (.11 to .35)	13.78*
Cuddliness	.08 (-.02 to .18)	.11 (.01 to .21)	.20 (.11 to .30)	.18 (.08 to .28)	.16 (.03 to .27)	.25 (.12 to .36)	6.39
Low intensity pleasure	.34 (.24 to .42)	.20 (.10 to .30)	.07 (-.03 to .17)	.20 (.10 to .30)	.12 (-.00 to .24)	.22 (.09 to .33)	16.69**
Surgency	.40 (.32 to .49)	.27 (.17 to .36)	.22 (.12 to .31)	.16 (.06 to .26)	.12 (-.00 to .24)	.04 (-.08 to .17)	27.92***
Negative affectivity	.21 (.11 to .31)	.20 (.10 to .30)	.22 (.13 to .31)	.19 (.09 to .29)	.15 (.03 to .27)	.15 (.03 to .28)	1.51
Effortful control	.23 (.13 to .32)	.03 (-.07 to .13)	.17 (.08 to .27)	.24 (.14 to .33)	.08 (-.04 to .20)	.18 (.05 to .30)	11.89*
Total <sup>a</sup>	.23 (.13 to .32)	.16 (.06 to .26)	.18 (.08 to .27)	.18 (.08 to .28)	.16 (.04 to .28)	.15 (.02 to .27)	1.33

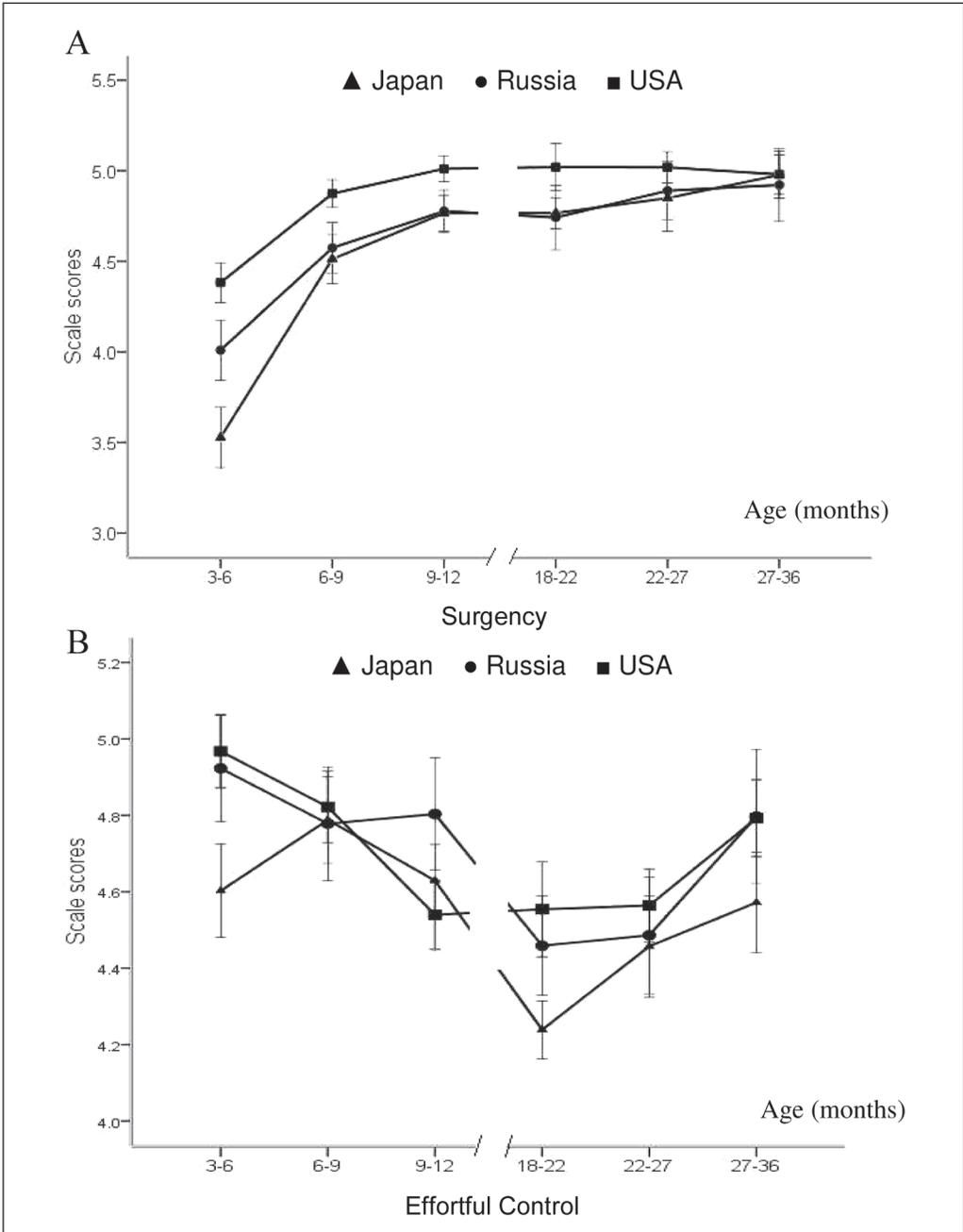
Note. Group 1: 3 to 6 months of age ( $N = 371$ ); Group 2: 6 to 9 months ( $N = 362$ ); Group 3: 9 to 12 months ( $N = 403$ ); Group 4: 18 to 22 months ( $N = 366$ ), Group 5: 22 to 27 months ( $N = 258$ ); and Group 6: 27 to 36 months ( $N = 242$ ). ES = effect size. CI = 95% confidence interval of estimated correlation. Q = homogeneity statistic.

a. Mean ES estimate for subscales.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$  (all two-tailed tests).

Extraversion and lower in Neuroticism than Russians (McCrae et al., 2005). For observer-ratings of 12- to 17-year-old adolescents from 24 countries (McCrae et al., 2010), Russian adolescents appeared to be more extraverted, more emotionally stable, and less neurotic than older Russians and more similar to Americans than to Japanese. Our findings show that more than 20 years “after the fall,” Russian infants’ temperament ratings were mostly in the middle range relative to Japanese and U.S. infants’ scores, but somewhat closer to the Japanese ratings. In toddlerhood, however, Russian children’s scores tended to be closer to the U.S. ratings, particularly in regards to aspects of negative emotionality.

One possible explanation for the heterogeneous findings in adults and children of different ages is that cross-cultural differences in temperament/personality could reflect the combination of relatively stable national differences due to genetic variation and characteristic adaptations that change over time as a consequence of changes in cultural values and practices (Allik & McCrae, 2004; Greenfield et al., 2003; Hofstede & McCrae, 2004; Keller, 2008; Super & Harkness, 1986). It may be that infants’ temperament ratings reflect more biologically based dispositions that make Japanese and Russian adult personality profiles very similar to each other



**Figure 4.** Temperament factors in Japanese, Russian and US children of different age groups: A - Surgency, B - Effortful Control  
 Note. Bars are 95% confidence intervals.

and different from those of Americans, whereas ratings from toddlerhood through adolescence may be more influenced by current parenting and socialization that make Japanese and Russians more Westernized. Perhaps these transitional developmental periods (i.e., toddlerhood and adolescence) elicit such variable socialization practices because of the particular challenges they

pose. That is, during the toddler period, parents prepare their children for culturally determined educational systems, and in adolescence, preparation for adulthood likely dictates additional culturally prescribed parenting practices.

At the cultural level, Extraversion is positively associated with Western geographical location and Euro-American culture, Gross Domestic Product (GDP) per capita, Human Development Index, and individualism, and negatively associated with Hofstede's (2001) cultural dimension of power distance, whereas Neuroticism is closely related to uncertainty avoidance and masculinity. Conscientiousness is positively associated with mean temperature and cultural dimension of power distance and negatively related to latitude and GDP per capita (Allik & McCrae, 2004; Hofstede & McCrae, 2004; McCrae et al., 2005). Cross-cultural differences in Surgency and Negative Affectivity found in the present study were in accord with these geographic and cultural correlates.

To our knowledge, this is the first study to show that Japanese infants and toddlers had lower scores on Effortful Control higher-order dimension, similar to the low scores on Conscientiousness noted for adult Japanese participants (McCrae, 2002; Schmitt et al., 2007). One possible explanation for these counterintuitive findings is that people make ratings in relation to culture-specific standards; and although no substantial culture-related differences in standards for Conscientiousness have been found (Möttus et al., 2012), research evidence suggests that Asian parents strongly value behavioral and emotional control in their children and emphasize parental control in childrearing (Chen, 2009; Greenfield et al., 2003). Considering that Japanese infants and toddlers had also higher scores for Fear, with cultural differences emerging toward the end of the first year of life, it might be possible that the Japanese eco-cultural context fosters development of a more reactive fear-based control system that does not overlap completely with more voluntarily engaged effortful control systems based in executive attention related mechanisms.

Overall, the effects of culture were in the small to medium range, accounting, on average, for 3.7% of the variance in IBQ-R subscales and 2.4% of the variance in ECBQ subscales. Gender accounted for .3% and .5% of the variance in IBQ-R and ECBQ subscales, respectively, while age accounted for 8.2% and 1.3%, respectively. Similar estimates for culture, gender, and age were recently described in a study of 12- to 17-year-old adolescents from 24 countries (McCrae et al., 2010), wherein these were identified as 3.6%, 2.8%, and .2%, respectively. Thus, cross-cultural differences in early temperament found in the present study were similar in magnitude to personality differences in adolescence, whereas the impact of age decreases and gender becomes a more substantial influence at older ages. Results concerning the interaction effects are perhaps the most important findings of the present study, because they qualify the main effects of culture and could help to better understand development of temperament in a cultural context.

Gender moderated a number of cultural differences in fine-grained traits and higher-order dimensions. For instance, cross-cultural differences on Negative Affectivity and Fear in toddlerhood were considerably more prominent in males than females. Viewing this interaction between culture and gender from another perspective, the gender difference was not consistent across the three cultures, with girls from the United States appearing to be substantially higher than boys on Fear, whereas gender differences in Japanese and Russian youngsters were nonsignificant. These findings have to be characterized as preliminary; however, this pattern of results is similar to recently reported findings, wherein U.S. boys exhibited lower levels of fearfulness later in childhood (3 to 7 years of age) relative to girls, coupled with nonsignificant gender differences for Finnish children (Gaias et al., 2012). It should be also noted that, overall, Culture  $\times$  Gender interactions effects were small, accounting for less than 1% of the variance in infants' trait scores and 1% to 2% of the variance in toddlers' trait scores. Nonetheless, these results have important implications for the meaning of gender differences in temperament, inviting the interpretation that socialization pressures in some countries accentuate gender differences for certain domains of temperament (i.e., High Intensity Pleasure and Fear). Future research is required to further

explore potential mechanisms responsible for the observed pattern of results; however, existing studies suggest that boys and girls in the United States receive differential treatment from parents very early in life (Leaper, Anderson, & Sanders, 1998; Lindsey & Mize, 2001). Furthermore, Hofstede (2001) found that Americans are often socialized to accept more distinct gender roles because of their parents' differentiated duties within the home, and this modeling leads to emphasizing assertiveness in males and pleasantness in females.

The present study also demonstrated that cultural effects on early temperament differ depending on child age. Whether this variation results from differing cultural influences on the different traits or from differing patterns of maturation due to genetic makeup of ethnic populations remains unclear, but either alternative in its pure form seems implausible. First, in recent decades many countries have experienced rapid economic and societal changes and seem to move toward a less differentiated cultural pattern (Gouveia & Ros, 2000). Both in Japan and Russia, cultural values and practices have become more Westernized (Ispa, 2002; Jandt, 2007), which may lead to increasing Westernization of parental childrearing attitudes (Chen et al., 1998). Second, both Russia and the United States have a high proportion of international migrants (Human Development Report, 2009) and thus are likely to be genetically heterogeneous populations. Both biological predispositions and cultural practice may be relevant for the development of temperament, and the relationship of biology and culture should be systematic (Keller, 2008). It will be up to future studies to outline the gene-environment interplay taking place in the context of temperament development in different regions of the world in a more precise manner.

A consistent age trend in cross-cultural differences represents a particularly notable finding. Specifically, effects of culture for the Surgency factor, as well as High and Low Intensity Pleasure fine-grained scales, decreased from 3 months to 3 years of age. Although nature and nurture are both likely to contribute to the emergence of cross-cultural differences in temperament, this pattern of results is more consistent with the implications of the biological effects model, wherein more prominent discrepancies are noted early in life, and not amplified via greater exposure/experience. It would be useful to examine the possibility that caregivers in Russia and Japan have shifted their beliefs/attitudes and parenting behaviors regarding the expression of positive emotions and other elements of Surgency in directions that are more consistent with Western approaches, which could help account for the apparently equalizing effects of caregiving suggested by these results. To our knowledge, this is the first study to demonstrate age-related differences on cultural differences in temperamental traits. Thus, our findings need to be replicated with other cultures using different measures and different informants.

It should also be noted that, overall, Culture  $\times$  Age interaction effects were small, accounting for four to nine times smaller amounts of variance than main effects of culture. In addition, although the cultural effect decreased with age for several traits, the effect of age was less clear for other traits, with cultural impact demonstrated only at later or intermediate time points. It is also worth reiterating that our study showed several cross-cultural differences that were consistent across age groups and that the pattern of these differences strongly resembles the picture that has consistently emerged from cross-cultural studies of adult and adolescent personality.

The present study has a number of limitations that should be addressed in future research. First, our sample only consisted of participants from three (although highly diverse) countries; research on cross-cultural differences would be improved by a broader range of cultures and larger and more representative groups of participants. Second, although evidence has provided support for the validity of parental report of child temperament, these findings should be confirmed with data from observational measures and information from other caregivers when possible. Furthermore, because child temperament development appears to be influenced by culture, future cross-cultural studies should include measures of parent-child interactions, parenting values and practices, and caregiving arrangements and family dynamics, to further articulate potential

environmental pathways. Finally, the cross-sectional design did not allow for an exploration of the developmental trends/growth curves. Although the existing evidence indicates that the results from longitudinal studies are generally consistent with cross-sectional findings, future research should utilize longitudinal data.

### Acknowledgments

The authors are grateful to the participants of the study; colleagues in Russia, Japan, and the United States; and anonymous reviewers for their helpful comments.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partly supported by a grant of the Russian Humanitarian Foundation (No. 11-06-00517a) to the first author, by the United States of America National Institute of Mental Health Behavioral Science Track Award for Rapid Transition (R3MH61831A) to the second author, and by a Grant-in-aid (No. 17330143) for scientific research from the Ministry of Education, Culture, Sports, Science and Technology of Japan to the third author.

### Note

1. To determine whether using actual age rather than age category affected the results, all analyses reported in the article were also run with age as covariate. Although this approach slightly altered the *F* values, significant multivariate and univariate culture effects for all higher-order temperament scales and subscales were the same.

### References

- Allik, J., & McCrae, R. R. (2004). Toward a geography of personality traits: Patterns of profiles across 36 cultures. *Journal of Cross-Cultural Psychology, 35*, 13-28.
- Allik, J., & Realo, A. (2004). Individualism-collectivism and social capital. *Journal of Cross-Cultural Psychology, 35*, 29-49.
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56*, 453-484.
- Chen, X. (2009). Culture and early socio-emotional development. In R. E. Tremblay, R. G. Barr, R. De V. Peters, & M. Boivin (Eds.), *Encyclopedia on early childhood development* (pp. 1-6). Montreal, Quebec, Canada: Centre of Excellence for Early Childhood Development. Retrieved from May 23, 2010, from <http://www.child-encyclopedia.com/documents/ChenANGxp.pdf>
- Chen, X., Hastings, P. D., Rubin, K. H., Chen, H., Cen, G., & Stewart, S. L. (1998). Child-rearing attitudes and behavioral inhibition in Chinese and Canadian toddlers: A cross-cultural study. *Developmental Psychology, 34*, 677-686.
- Chess, S., Thomas, A., & Birch, H. G. (1965). *Your child is a person: A psychological approach to parenthood without guilt*. New York: Viking Press, Inc.
- De Fruyt, F., De Bolle, M., McCrae, R. R., Terracciano, A., Costa, P. T., Jr., & 39 Collaborators from the Adolescent Personality Profiles of Cultures Project (APPOC). (2009). The universal structure of personality in adolescence: NEO-PI-R and NEO-PI-3 findings from 24 cultures. *Assessment, 16*, 301-311.
- Gaias, L. M., Räikkönen, K., Komsu, N., Gartstein, M. A., Fisher, P. A., & Putnam, S. P. (2012). Cross-cultural temperamental differences in infants, children, and adults in the United States of America and Finland. *Scandinavian Journal of Psychology, 53*, 119-128.

- Gartstein, M. A., Knyazev, G. G., & Slobodskaya, H. R. (2005). Cross-cultural differences in the structure of infant temperament: United States of America (U.S.) and Russia. *Infant Behavior and Development, 28*, 54-61.
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via a revision of the infant behavior questionnaire. *Journal of Infant Behavior and Development, 26*, 64-86.
- Gartstein, M. A., Slobodskaya, H. R., & Kinsht, I. A. (2003). Cross-cultural differences in the first year of life: United States of America (U.S.) and Russia. *International Journal of Behavioral Development, 27*, 316-328.
- Gouveia, V. V., & Ros, M. (2000). The Hofstede and Schwartz models for classifying individualism at the cultural level: Their relation to macro-social and macro-economic variables. *Psicothema, 12*, 25-33.
- Greenfield, P. M., Keller, H., Fuligni, A., & Maynard, A. (2003). Cultural pathways through universal development. *Annual Review of Psychology, 54*, 461-491.
- Hart, S., Field, T., & Roitfarb, M. (1999). Depressed mothers' assessments of their neonates' behaviors. *Infant Mental Health Journal, 20*, 200-210.
- Haub, C. (2010). *2010 world population data sheet*. Washington, DC: Population Reference Bureau.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). Beverly Hills, CA: Sage.
- Hofstede, G., & McCrae, R. R. (2004). Personality and culture revisited: Linking traits and dimensions of culture. *Cross-Cultural Research, 38*, 52-88.
- Hsu, C., Soon, W., Stigler, J. W., Hong, C., & Liang, C. (1981). The temperamental characteristics of Chinese babies. *Child Development, 52*, 1337-1340.
- Huberty, C. J. (2002). A history of effect size indices. *Educational and Psychological Measurement, 62*, 227-240.
- Human Development Report. (2009). *Overcoming barriers: Human mobility and development*. Basingstoke, Hampshire, UK: Palgrave Macmillan.
- Ispa, J. M. (2002). Russian child care goals and values: From Perestroika to 2001. *Early Childhood Research Quarterly, 17*, 393-413.
- Jandt, F. E. (2007). *An introduction to intercultural communication: Identities in a global community* (5th ed.). Thousand Oaks, CA: Sage.
- Kagan, J. (1994). *Galen's prophecy: Temperament in human nature*. New York: Basic Books.
- Kagan, J., Reznick, J. S., & Snidman, N. (1986). Temperamental inhibition in early childhood. In R. Plomin & J. Dunn (Eds.), *The study of temperament: Changes, continuities, and challenges* (pp. 53-79). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Keller, H. (2008). Culture and biology: The foundation of pathways of development. *Social and Personality Psychology Compass, 2*, 668-681.
- Keller, H., Yovsi, R., Borke, J., Kartner, J. H., Jensen, H., & Papaligoura, Z. (2004). Developmental consequences of early parenting experiences: Self-recognition and self-regulation in three cultural communities. *Child Development, 75*, 1745-1760.
- Kohnstamm, G. A. (1989). Temperament in childhood: Cross-cultural and sex differences. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 483-508). New York: Wiley.
- Kohnstamm, G. A., Halverson, C. F., Jr., Mervielde, I., & Havill, V. L. (1998). *Parental descriptions of child personality: Developmental antecedents of the big five?* Mahwah, NJ: Lawrence Erlbaum Associates.
- Kolmagorova, A. V., Slobodskaya, H. R., & Gartstein, M. A. (2008). A Russian adaptation of the Early Childhood Behavior Questionnaire (ECBQ) for measuring temperament. *Psikhologicheskii Zhurnal, 29*, 121-136.
- Leaper, C., Anderson, K. J., & Sanders, P. (1998). Moderators of gender effects on parents' talk to their children: A meta-analysis. *Developmental Psychology, 34*(1), 3-27.
- Lindsey, E. W., & Mize, J. (2001). Contextual differences in parent-child play: Implications for children's gender role development. *Sex Roles, 44*(3-4), 155-176.

- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- McCrae, R. R. (2002). NEO-PI-R data from 36 cultures: Further intercultural comparisons. In R. R. McCrae & J. Allik (Eds.), *The Five-Factor Model of personality across cultures* (pp. 105-125). New York: Kluwer Academic/Plenum Publishers.
- McCrae, R. R., Terracciano, A., & 79 Members of the Personality Profiles of Cultures Project. (2005). Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology, 89*, 407-425.
- McCrae, R. R., Terracciano, A., De Fruyt, F., De Bolle, M., Gelfand, M. J., Costa, P. T. & 42 Collaborators of the Adolescent Personality Profiles of Cultures Project. (2010). The validity and structure of culture-level personality scores: Data from ratings of young adolescents. *Journal of Personality, 78*(3), 815-838.
- Möttus, R., Allik, J., Realo, A., Pullmann, H., Rossier, J., Zecca, G., et al. (2012). Comparability of self-reported Conscientiousness across 21 Countries. *European Journal of Personality, 26*, 303-317.
- Nakagawa, A., & Sukigara, M. (2005). How are cultural differences in the interpretation of infant behavior reflected in the Japanese Revised Infant Behavior Questionnaire? *Japanese Journal of Educational Psychology, 53*, 491-503.
- Nakagawa, A., Sukigara, M., & Mizuno, R. (2007, April). *Cultural effects reflected in the Early Childhood Behavior Questionnaire for Japanese toddlers: Psychometrics and factor structure*. Paper presented at Society for Research in Child Development Biennial Meeting, Boston.
- Pauli-Pott, U., Mertesacker, B., Bade, U., Haverkock, A., & Beckmann, D. (2003). Parental perceptions of infant temperamental development. *Infant Behavior and Development, 26*, 27-48.
- Putnam, S. P., Ellis, L. K., & Rothbart, M. K. (2001). The structure of temperament from infancy through adolescence. In A. Eliasch & A. Angleitner (Eds.), *Advances/proceedings in research on temperament* (pp. 165-182). Germany: Pabst Scientist Publisher.
- 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*, 386-401.
- Rothbart, M. K. (2007). Temperament, development and personality. *Current Directions in Psychological Science, 16*, 207-212.
- Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: (Vol. 3), Social, emotional, and personality development* (5th ed., pp. 105-176). New York: Wiley.
- Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (Vol. 1, pp. 37-86). Hillsdale, NJ: Erlbaum.
- Schmitt, D. P., Allik, J., McCrae, R. R., Benet-Martinez, V., Alcalay, L., Ault, L., et al. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology, 38*, 173-212.
- Strelau, J. (1998). *Temperament: A psychological perspective*. New York: Plenum Press.
- Super, C., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development, 9*, 545-569.
- Triandis, H. C., & Suh, E. M. (2002). Cultural influences on personality. *Annual Review of Psychology, 53*, 133-160.
- Wilson, D. B. (2006). *Meta-analysis macros for SAS, SPSS, and Stata*. Retrieved February 15, 2011, from <http://mason.gmu.edu/~dwilsonb/ma.html>
- Windle, W., Iwawaki, S., & Lerner, R. M. (1988). Cross-cultural comparability of temperament among Japanese and American preschool children. *International Journal of Psychology, 23*, 547-567.