



# Psychometric Properties of the Chinese Version of the Highly Sensitive Child Scale Across Age Groups, Gender, and Informants

Danni Liu<sup>1</sup> · Anouk van Dijk<sup>2</sup> · Shanyan Lin<sup>3</sup> · Zhenhong Wang<sup>4</sup> · Maja Deković<sup>5</sup> · Judith Semon Dubas<sup>1</sup>

Accepted: 6 April 2023 / Published online: 13 April 2023  
© The Author(s) 2023

## Abstract

Sensory Processing Sensitivity (SPS) is theorized to be a fundamental trait capturing children’s general sensitivity to the environment. Yet, scientific knowledge of SPS is mostly based on findings from Western cultures and few translated measures exist to assess children’s SPS outside of Western countries. Therefore, we developed the Chinese Highly Sensitive Child (HSC) scale. In Study 1, we investigated the scale’s psychometric properties for both self-reports ( $N=2925$ ,  $M_{\text{age}} = 11.74$  years, 43.3% girls) and caregiver reports ( $n=460$ ,  $M_{\text{child age}} = 9.02$  years, 44.0% girls). Findings replicated most psychometric properties found in international studies including: (a) a bifactor structure with one general sensitivity component and three specific components, (b) acceptable internal consistency of the total scale (although not for self-report of elementary school children, and not for the subscales), and (c) at least partial invariance across age groups, gender, and informants. In Study 2, we investigated convergent validity with related temperament and personality measures using self-reports from both elementary school children ( $n=845$ ,  $M_{\text{age}} = 9.71$  years, 41.9% girls) and middle school children ( $n=563$ ,  $M_{\text{age}} = 13.17$  years, 43.2% girls). Findings replicated bivariate associations found in Western studies: Ease of Excitation (EOE) was associated with more positive traits, whereas Aesthetic Sensitivity (AES) was associated with more negative traits, suggesting that EOE and AES may capture the “dark” and “bright side” of sensitivity, respectively. We hope that our studies help spur research on SPS across western and Chinese cultures.

**Keywords** Sensory processing sensitivity · Highly sensitive child scale · Environmental sensitivity · Chinese children · Bifactor model · Personality

# 1 Psychometric Properties of the Chinese Version of the Highly Sensitive Child Scale Across Age Groups, Gender, and Informants

Children vary in their reactions to the very same environment, which may be explained by individual differences in environmental sensitivity (Pluess et al., 2015). That is, some children may process environmental stimuli more deeply, leading them to respond more strongly to the environment. For instance, they may be more likely to develop aggressive behavior problems when growing up in negative parenting environments (Slagt et al., 2016), or they may profit significantly more from depression intervention programs (Pluess & Boniwell, 2015).

One prominent theory describing this phenomenon is Sensory Processing Sensitivity (SPS), which proposes the SPS trait describing and capturing a general trait of sensitivity in humans (Aron & Aron, 1997). Individuals high on this trait are characterized by greater behavioral inhibition in novel situations, greater awareness of sensory stimulation, deeper cognitive processing of environmental stimuli, and higher emotional and physiological reactivity (Aron et al., 2012). More recently, SPS theory was integrated into a broader meta-framework of Environmental Sensitivity (Pluess, 2015). In this framework, the SPS trait is advanced as a reliable psychological marker of individual differences in environmental sensitivity, given its core ability to capture heightened responsiveness to both positive and negative environments (Pluess et al., 2018; Greven et al., 2019).

SPS would be expected to be universal and global. After all, it was proposed as a “fundamental individual difference” (Aron & Aron, 1997, p. 347) and “a common, heritable and evolutionarily conserved trait” (Graven et al., 2019, p. 287). Research on SPS across cultures is therefore important to further support SPS theory and the Environmental Sensitivity meta-framework. Yet, few translated measures exist to assess children’s SPS outside of Western countries. Hence, in the current paper, we validated a Chinese translation of the Highly Sensitive Child (HSC) scale (Pluess et al., 2018) to help facilitate research on SPS across Western and Chinese cultures.

## 1.1 Existing Assessments of Sensory Processing Sensitivity

SPS in children can be assessed using self-report questionnaires (Pluess et al., 2018), caregiver reports (Slagt et al., 2018; Sperati et al., 2022), as well as observations (Davies et al., 2021; Lionetti et al., 2019). The most widely used measure is the self-report Highly Sensitive Child (HSC) scale (Pluess et al., 2018), which is developed from the self-report Highly Sensitive Person (HSP) scale (Aron & Aron, 1997). The HSC scale has been adapted into many languages, but mostly in Western cultures (e.g., Germany, Netherlands, Belgium, Italy, Spain; Costa-López et al., 2022; Sperati et al., 2022). Until recently, Chinese translations were lacking, even though China has the second-largest global child population (UNFPA, 2022).

Perhaps illustrating the relevance of a Chinese translation, while we worked on this study, two other studies on Chinese translations came out. Yet, their validation of the scale is limited to self-reports by middle school children (Dong et al., 2022) or caregiver reports for preschoolers (Zeng & Wang, 2022). Moreover, the self-report study (Dong et al., 2022) did not examine convergent validity with related tempera-

ment and personality measures. Yet, such information is crucial for understanding whether the self-report Chinese HSC scale relates to other established measures in similar patterns to Western findings (Pluess et al., 2018). Therefore, in the current paper, we aimed to more thoroughly validate our Chinese translation of the HSC in samples that cover both elementary school and middle school children. We investigated the psychometric properties of the self-report HSC scale and caregiver-report HSC scale (Study 1) as well as the convergent validity of the self-report HSC scale with related temperament and personality measures (Study 2).

## 1.2 Psychometric Qualities of the Highly Sensitive Child Scale

The original HSC scale contains 12 items making up three subscales: (a) Ease of Excitation (EOE), referring to being easily overwhelmed (e.g., finding it unpleasant to have a lot going on at once), (b) Aesthetic Sensitivity (AES), referring to having high aesthetic awareness (e.g., of smell and tastes), and (c) Low Sensory Threshold (LST), referring to being easily aroused by unpleasant sensory stimuli (e.g., loud noises; Pluess et al., 2018). Previous research has investigated the psychometric properties (i.e., factor structure, internal consistency, and measurement invariance) and convergent validity (i.e., bivariate associations with related temperament and personality measures) of the HSC scale internationally. We expect to replicate these findings in our Chinese sample.

Regarding psychometric properties of the HSC scale, research has revealed similar findings across Western and Chinese cultures, for both self-reports (Dong et al., 2022; Pluess et al., 2018; Weyn et al., 2021) and caregiver reports (Sperati et al., 2022; Zeng & Wang, 2022). Specifically, regarding factor structure, previous research has supported a bifactor structure of the HSC scale. This finding indicates that the HSC total score captures an overall sensitivity factor, while the three subscales capture additional variance representing different aspects of sensitivity. Regarding internal consistency, previous research generally found acceptable to good internal consistency for the HSC total scale and subscales. However, internal consistencies were generally higher for the self-report HSC total scale (versus subscales) and older children (versus younger children) (Pluess et al., 2018; Weyn et al., 2021). Regarding measurement invariance, most previous studies found full configural invariance and at least partial metric and partial scalar invariance across age groups (e.g., Sperati et al., 2022), gender (e.g., Weyn et al., 2021), and informants (e.g., Weyn et al., 2022). This indicates that the underlying bifactor structure of the HSC scale is conceptualized identically across these groups (i.e., full configural invariance) while the meaning attributed to the items (i.e., partial metric invariance) and reference point used when answering the items (i.e., partial scalar invariance) differed across these groups for invariant items (Weyn et al., 2022). Given the above-described culturally similar international findings on psychometric properties of the HSC scale, we sought to replicate them in Study 1, using child and caregiver reports, respectively. That is, for both child and caregiver reports, we expected a bifactor structure and acceptable to good internal consistency for the HSC total scale and subscales. Further, we expected full configural invariance and at least partial metric and partial scalar invariance across age groups and gender, and across child and caregiver reports.

Regarding convergent validity of the self-reported HSC scale with related temperament and personality measures, current evidence mostly stems from Western cultures (Iimura & Kibe, 2020; Pluess et al., 2018; Sperati et al., 2022; Weyn et al., 2021; Weyn et al., 2022). In general, convergent validity of HSC (sub)scales was well-supported by significant associations between HSC (sub)scales and related measures, including Behavioral Inhibition System (BIS), Behavioral Activation System (BAS), Negative Emotionality (NE), Positive Emotionality (PE), Neuroticism, Extraversion, and Openness. Specifically, research has generally found positive associations between HSC and all these measures except for Extraversion, which was negatively associated with HSC. Regarding subscales, EOE and LST were more strongly associated with measures reflecting sensitivity to *negative* environmental influences (i.e., BIS, NE, Negative Affect, and Neuroticism), whereas AES was more strongly associated with measures reflecting sensitivity to *positive* environmental influences (i.e., BAS, PE, Positive Affect, Extraversion, and Openness; Iimura, & Kibe, 2020; Pluess et al., 2018; Sperati et al., 2022; Weyn et al., 2021). These results suggest that the subscales may reflect different sensitivity components, with EOE and LST capturing sensitivity to negative aspects of the environment (i.e., the “dark side” of sensitivity), while AES capturing sensitivity to positive aspects of the environment (i.e., the “bright side” of sensitivity). This could explain why the HSC total score is able to capture heightened responsiveness to both positive and negative experiences (Pluess et al., 2018; Sperati et al., 2022; Weyn et al., 2021, 2022). To date, we know of only one Chinese study that has investigated convergent validity of the HSC scale (Zeng & Wang, 2022). This study has used the caregiver-report HSC scale among preschoolers. Contrary to Western findings (Pluess et al., 2018; Sperati et al., 2022; Weyn et al., 2021), it revealed significant positive associations of Extraversion with HSC and LST. Given this culturally inconsistent finding, a comprehensive examination of convergent validity of the HSC scale in the Chinese culture is warranted. Therefore, in Study 2, we sought to explore whether Western findings on convergent validity of the HSC scale can be replicated in self-reports of Chinese children.

### 1.3 Overview of the Present Research

The Environmental Sensitivity meta-framework and SPS theory may be reinforced by empirical evidence from different cultures and countries. This would first require the availability of a translated and validated measure. Therefore, we developed the child- and caregiver-reported Chinese HSC scales and addressed two complementing aims across two studies, covering both elementary and middle school children. In Study 1, we examined the psychometric properties (i.e., factor structure, internal consistency, and measurement invariance) of the child- and caregiver-reported HSC. For child reports, we involved  $N=2925$  Chinese elementary and middle school children (aged 6.92–16.75 years). For caregiver reports, we involved caregivers of a subsample of  $n=460$  elementary school children. We sought to replicate previous international findings and extend these to Chinese elementary school children in Study 1. Study 2 further extends previous research on the HSC in Chinese Children, examining convergent and discriminant validity by inspecting bivariate associations between self-reported HSC and related temperament and personality measures

(i.e., BIS and BAS, PE and NE, and the Big Five) in two subsamples of elementary ( $n=845$ ; aged 6.92–12.75 years) and middle school children ( $n=563$ ; aged 11–15.75 years). Data, analysis code, and the Chinese HSC scale are available through the Open Science Framework at <https://osf.io/rgan7/>.

## 2 Study 1

Study 1 investigated whether the psychometric properties of our child- and caregiver-reported HSC replicate prior international findings, that is, (1) a bifactor structure, (2) acceptable internal consistency for all (sub)scales, and (3) full configural invariance and at least partial metric and partial scalar invariance across age groups, gender and informants.

### 2.1 Method

#### 2.1.1 Participants

We used existing data from 4 child subsamples, yielding a total sample of 2925 children ( $M_{\text{age}} = 11.74$  years;  $SD=1.90$ ; range=6.92–16.75; 43.3% girls, 52.4% boys, 4.2% not reported). They were recruited from two public elementary schools (i.e., Grades 3–6) and two public middle schools (i.e., Grades 7–9) in two cities located in central and southeastern China. Data were collected from 13 more children but were excluded because these children were too distracted during questionnaire administration. For child-reported HSC, few data were missing (0.09%), with item-level missingness lower than 0.05% for all items.

Caregivers were recruited for one child subsample (response rate=88.63%). As a result, caregiver reports (75.7% mothers, 18.5% fathers, and 5.9% other relatives) were available for 460 elementary school children in Grades 3–4 ( $M_{\text{age}} = 9.02$  years;  $SD=0.64$ ; range=6.92 – 10.92; 44.0% girls, 56.0% boys). Data of 33 more caregivers were excluded because they spent less than an average of 2 s per item on the questionnaire ( $n=1$ ) or chose the wrong answer for a control item inserted in the questionnaire ( $n=32$ ; e.g., “Please select option 2 for this item”; DeSimone et al., 2015). For caregiver-reported HSC, there were no missing data. We used full information maximum likelihood (FIML; Enders & Bandalos 2001) in *Mplus* to handle missing data.

Combined reports from caregivers (for 460 children with caregiver reports) and children (for the remaining child sample) indicated that 78.2% of fathers and 74.9% of mothers had at least a middle school education, with 9.6% of fathers and 7.7% of mothers completing a university degree (4.6% missing). Sample size was determined based on recent validation studies on the HSC Scale, which generally used large sample sizes (e.g., Pluess et al., 2018; Wyne et al., 2021).

### 2.1.2 Procedure

Children completed questionnaires in their classrooms with one headteacher and researcher present. The researcher gave instructions and answered questions about the questionnaires when necessary. Caregivers completed an online questionnaire including the HSC scale and other measures not relevant to the current study through a link shared by headteachers.

At the end of the questionnaire, both children and caregivers were thanked, and children received a pen as a gift. We obtained informed consent from schools and caregivers and verbal assent from participating children. Data collection in the 4 subsamples was approved either by the ethics committee of the Faculty of Social and Behavioral Sciences of Utrecht University or the local Chinese Education Bureau.

### 2.1.3 Sensory Processing Sensitivity

SPS was assessed using the HSC scale, translated from the English version using translation and back-translation procedures (Brislin, 1970). We contacted the scale developer (i.e., M. Pluess) for 4 items that yielded differences in certain words or phrases between the back-translated version and the original version. We made the final decisions based on Pluess' feedback. For caregiver reports, we rephrased all items in the third person (i.e., "my child"; Slagt et al., 2018; Sperati et al., 2022).

The Ease of Excitation (EOE) subscale includes 5 items (e.g., "I find/my child finds it unpleasant to have a lot going on at once"), the Aesthetic Sensitivity (AES) subscale includes 4 items (e.g., "I love/my child loves nice smells"), and the Low Sensory Threshold (LST) subscale includes 3 items (e.g., "Loud noises make me/my child feel uncomfortable"). Each item was rated on a 7-point Likert scale with three anchors (1 = *not at all*, 4 = *moderately*, and 7 = *extremely*). We calculated the total and subscale scores as the mean across corresponding items.

### 2.1.4 Data Analysis

To examine the factor structure of the HSC scale, we conducted a series of Confirmatory Factor Analyses (CFAs) in *Mplus* 8.3 (Muthen & Muthen, 1998–2019) for both child and caregiver reports. Following existing research (Sperati et al., 2022; Weyn et al., 2021), we compared three competing models: (a) a one-factor model (i.e., HSC as a single factor), (b) a three-factor model (i.e., three subscales as uncorrelated factors), and (c) a bifactor model (i.e., one overarching general sensitivity factor and the three subscales as uncorrelated specific factors)<sup>1</sup>. We used the maximum likelihood robust estimator to deal with nonnormality (Satorra & Bentler, 1994). We first examined model fit for the three competing models separately. We considered model fit to be acceptable if indices were equal to or greater than 0.90 for the comparative fit index (CFI), equal to or smaller than 0.06 for the root mean square error of approximation (RMSEA), and equal to or smaller than 0.08 for the standardized root mean squared residual (SRMR; Kline 2005). We then compared the fit between the three models.

<sup>1</sup> For path diagrams of CFA models for all three models, see Sperati et al., 2022, p. 4.

We considered a model as having a better fit when  $\Delta CFI$  was at least 0.010,  $\Delta RMSEA$  0.015, and  $\Delta SRMR$  0.010 (Chen, 2007). We also used the Akaike information criterion (AIC) and Bayesian information criterion (BIC) as comparative fit indices, with smaller values indicating better fit (Raftery, 1995). In addition, if data favored the bifactor model, we examined the strength of the general factor relative to the specific factors by calculating the Explained Common Variance (ECV; i.e., the percentage of common variance explained by the general factor; Sijtsma 2009). The higher the ECV, the stronger the general factor. Lower ECV values indicate multidimensionality and support a bifactor structure (Rodriguez et al., 2016).

To examine internal consistency of the child- and caregiver-reported HSC (sub) scales, we used Cronbach's alpha ( $\alpha$ ) and McDonald's omega ( $\omega$ ), estimated using SPSS (version 26) and *Mplus*, respectively. Although Cronbach's  $\alpha$  is the most commonly used coefficient to examine internal consistency, it has strict assumptions that are hard to meet in reality (Dunn et al., 2014). Therefore, we also reported McDonald's  $\omega$ , a factor analytic model-based estimate of internal consistency providing a practical and appropriate alternative to Cronbach's  $\alpha$  (Dunn et al., 2014; Flora, 2020; McDonald, 1999). We considered values of both Cronbach's  $\alpha$  and McDonald's  $\omega$   $\leq .60$  as low, between 0.60 and 0.80 as acceptable, and  $\geq 0.80$  as good. In addition, if the data supported the bifactor model, we calculated bifactor-specific reliability indices: McDonald's Omega Hierarchical for the total scale ( $\omega_H$ ) and subscales ( $\omega_{HS}$ ; McDonald 1999). Specifically, McDonald's  $\omega_H$  estimates the proportion of variance of the HSC total scale scores explained by the general factor while controlling for the specific factors. Vice versa, McDonald's  $\omega_{HS}$  estimates the proportion variance of subscale scores explained by the specific factors while controlling for the general factor (Rodriguez et al., 2016). Higher values of McDonald's  $\omega_H$  indicate greater confidence in interpreting the total scale scores as due to the general factor, and higher values of  $\omega_{HS}$  greater confidence in interpreting the subscale scores as due to the specific factors (Green & Yang, 2015).

To examine measurement invariance of the HSC scale across age groups (elementary vs. middle school children), gender (boys vs. girls), and informants (child vs. caregiver reports), we used multigroup CFA analyses in *Mplus*. In the first step, we tested for configural invariance (i.e., if the same items load on the same factors across groups while factor loadings and intercepts are freely estimated across groups), which was established if the configural invariance model showed acceptable fit (i.e.,  $CFI \geq 0.90$ ,  $RMSEA \leq 0.06$ , and  $SRMR \leq 0.08$ ). In the second step, we tested for metric invariance (i.e., if factor loadings are the same across groups while intercepts are freely estimated across groups), which was established if the metric invariance model had a similar fit as the configural model (i.e., decreases in  $\Delta CFI < 0.010$ , increases in  $\Delta RMSEA < 0.015$ , and increases in  $\Delta SRMR < 0.030$ ; Chen 2007). If fit decreased, we tested for partial metric invariance by freeing the non-invariant factor loadings one by one across groups. Partial metric invariance was established if a model was found with a similar fit as the configural model and at least two invariant factor loadings per latent factor (Byrne et al., 1989). In the last step, we tested for scalar invariance (i.e., if factor loadings and intercepts are the same across groups), which was established if the scalar invariance model had a similar fit as the (partial) metric invariance model (i.e., decreases in  $\Delta CFI < 0.010$ , increases in  $\Delta RMSEA < 0.015$ , and increases



in  $\Delta SRMR < 0.010$ ; Chen 2007). If fit decreased, we tested for partial scalar invariance by freeing the non-invariant intercepts. In all steps, if fit indices disagreed, we relied on  $\Delta CFI$  because simulation studies suggest that  $\Delta CFI$  should be given more weight than  $\Delta RMSEA$  and  $\Delta SRMR$  in comparing nested models (Sellbom & Tellegen, 2019).

## 2.2 Results

### 2.2.1 Factor Structure

Descriptive statistics and intercorrelations between the HSC (sub)scales are presented in Table 1. For child-reported HSC, CFA results supported the bifactor model. The one-factor model had an unacceptable fit ( $CFI=0.630$ ;  $RMSEA=0.084$ ;  $SRMR=0.065$ ;  $AIC=140721.874$ ;  $BIC=140937.191$ ). Model fit was acceptable for the three-factor model ( $CFI=0.886$ ;  $RMSEA=0.047$ ;  $SRMR=0.037$ ;  $AIC=139791.70$ ;  $BIC=140024.96$ ) and the bifactor model ( $CFI=0.929$ ;  $RMSEA=0.041$ ;  $SRMR=0.031$ ;  $AIC=139632.65$ ;  $BIC=139919.74$ ). Yet, the data supported the bifactor model over the three-factor model:  $\Delta CFI=0.043$ ;  $\Delta RMSEA=-0.006$ ;  $\Delta SRMR=-0.006$ ,  $\Delta AIC=-159.053$ ;  $\Delta BIC=-105.223$ . The ECV of the general factor was 0.49, indicating that the general factor explained 49% of the common variance.

For caregiver-reported HSC, CFA results also supported the bifactor model. Again, the one-factor model had an unacceptable fit ( $CFI=0.651$ ;  $RMSEA=0.084$ ;  $SRMR=0.07$ ;  $AIC=21029.834$ ;  $BIC=21178.558$ ), whereas model fit was accept-

**Table 1** Study 1 Descriptive Statistics, Internal Consistency, and Intercorrelations of the HSC (Sub)scales

	<i>N</i>	<i>M</i>	<i>SD</i>	Range	$\alpha$ , 95% CI	$\omega$ , 95% CI	$\omega_{HHS}$ , 95% CI	HSC	EOE	LST
<i>Child Report</i>										
HSC	2925	4.95	0.85	1.00–7.00	0.65, [0.63, 0.67]	0.70, [0.68, 0.73]	0.55, [0.52, 0.58]	—		
EOE	2925	4.33	1.23	1.00–7.00	0.60, [0.58, 0.62]	0.63, [0.60, 0.67]	0.02, [-0.09, 0.12]	0.80***	—	
LST	2925	5.31	1.40	1.00–7.00	0.56, [0.53, 0.58]	0.59, [0.56, 0.62]	0.46, [0.39, 0.53]	0.65***	0.28***	—
AES	2925	5.46	1.06	1.00–7.00	0.47, [0.44, 0.50]	0.48, [0.44, 0.52]	0.41, [0.36, 0.46]	0.61***	0.20***	0.18***
<i>Caregiver Report</i>										
HSC	460	4.60	0.79	2.25–6.92	0.67, [0.62, 0.71]	0.73, [0.69, 0.78]	0.54, [0.47, 0.62]	—		
EOE	460	3.93	0.95	1.00–7.00	0.50, [0.43, 0.57]	0.54, [0.45, 0.63]	0.01, [-0.11, 0.14]	0.75***	—	
LST	460	4.60	1.45	1.00–7.00	0.64, [0.58, 0.69]	0.70, [0.65, 0.74]	0.44, [0.32, 0.57]	0.73***	0.34***	—
AES	460	5.45	1.04	1.50–7.00	0.56, [0.49, 0.62]	0.57, [0.50, 0.64]	0.47, [0.38, 0.57]	0.65***	0.21***	0.22***

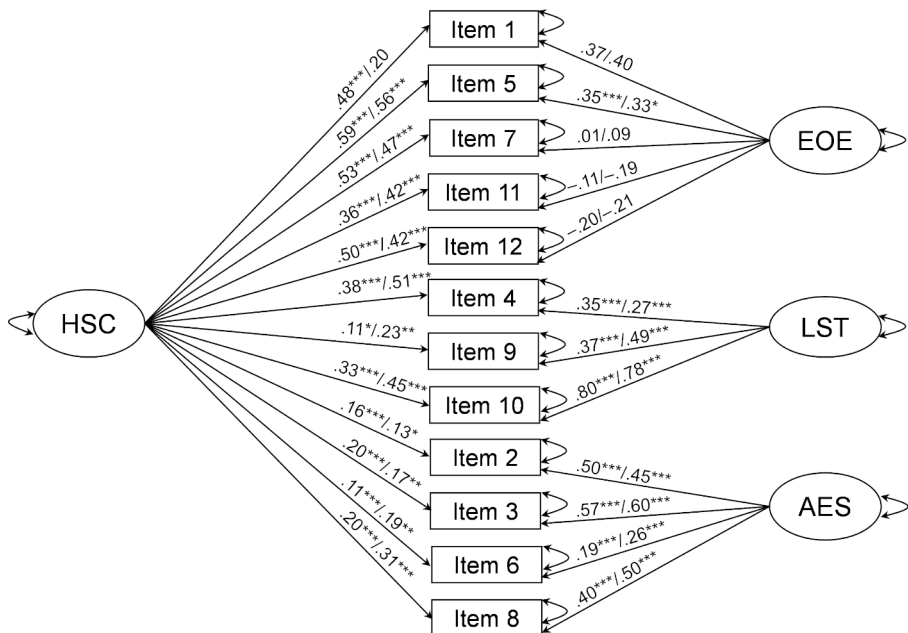
*Note.* HSC=Highly Sensitive Child total scale; EOE=Ease of Excitation; LST=Low Sensory Threshold; AES=Aesthetic Sensitivity;  $\omega_H$ =McDonald’s Omega Hierarchical for the total scale;  $\omega_{HS}$ =McDonald’s Omega Hierarchical for the subscales. \*\*\* $p < .001$



able for the three-factor model (CFI=0.897; RMSEA=0.047; SRMR=0.047; AIC=20886.46; BIC=21047.58), and the bifactor model<sup>2</sup> (CFI=0.974; RMSEA=0.026; SRMR=0.032; AIC=20846.75; BIC=21040.92). The data supported the bifactor model over the three-factor:  $\Delta\text{CFI}=0.077$ ;  $\Delta\text{RMSEA}=-0.021$ ;  $\Delta\text{SRMR}=-0.015$ ,  $\Delta\text{AIC}=-39.71$ ;  $\Delta\text{BIC}=-6.66$ . The ECV of the general factor was 0.43, indicating that the general factor explained 43% of the common variance. The standardized factor loadings for the final bifactor models for both informants are presented in Fig. 1.

## 2.2.2 Internal Consistency

Internal consistency values of the HSC (sub)scales for both informants are reported in Table 1. For child reports, internal consistency was acceptable for the HSC total scale and the EOE subscale, but low for the LST and AES subscales. For caregiver reports, internal consistency was acceptable for the HSC total scale and the LST subscale, but low for the EOE and AES subscales. We further examined bifactor-



**Fig. 1** Graphical Illustration and Standardized Factor Loadings of the Final Bifactor Model for Both Child-Report (Before the Slash) and Caregiver Report (Behind the Slash)

*Note.* HSC=Highly Sensitive Child; EOE=Ease of Excitation; LST=Low Sensory Threshold; AES=Aesthetic Sensitivity. The numbering of items we used was based on Table 2 of the original publication of the HSC scale (for the exact content of these items, see Pluess et al., 2018, p. 55).

\* $p < .05$ ; \*\*\* $p < .001$

<sup>2</sup> We identified one small and nonsignificant negative residual variance for the LST item “My child doesn’t like loud noises.” Following Sperati and colleagues (2022), we fixed this residual variance to 20% of the observed item variance in all follow-up analyses.

specific indices. Omega hierarchical ( $\omega_H$ ) for the HSC total scale was 0.55 for child reports and 0.54 for caregiver reports. This indicates that 55% or 54% of the variance of the HSC total scale scores can be attributed to the general sensitivity factor after controlling for the variance due to the specific factors (EOE, LST, and AES). Next, we compared this  $\omega_H$  for the HSC total scale to  $\omega$  for the overall internal consistency (i.e., 0.70 and 0.73 for child and caregiver reports, respectively). This shows that the majority of reliable variance in the HSC total scale scores is attributable to the general factor. Turning to the unique explained variance of the subscales, we found that omega hierarchical ( $\omega_{HS}$ ) for the EOE subscale was 0.02 for child reports and 0.01 for caregiver reports, indicating that only 2% or 1% of the variance of the EOE subscale scores can be attributed to the EOE specific factor after controlling for the variance due to the general factor. Thus, the overwhelming majority of reliable variance in the EOE subscale scores is attributable to the general factor. The other two specific factors do seem to explain a majority of reliable variance of the subscales:  $\omega_{HS}$  for the LST and AES subscales were 0.46 and 0.41 for child reports, and 0.44, and 0.47 for caregiver reports. In sum, the general factor explained the majority of reliable variance for the HSC total scale and the EOE subscale, whereas the AES and LST specific factors explained the majority of reliable variance for the AES and LST subscales, respectively.

### 2.2.3 Measurement Invariance

We tested measurement invariance of the HSC scale within child reports (i.e., across age groups and gender) and between child and caregiver reports (i.e., across informants) using multigroup bifactor CFA analyses. If we found non-invariant items, they were freed to vary across groups for all following models. Results of all invariance models as well as non-invariant items are shown in Table 2.

Across age groups (elementary vs. middle school children), we identified one small and nonsignificant negative residual variance in the elementary school group for the configural invariance model, which we resolved by imposing constraints following Sperati and colleagues (2022)<sup>3</sup>. We found evidence for full configural, full metric, and partial scalar invariance. One item (i.e., “I notice it when small things have changed in my environment”) was variant, with a higher intercept in elementary school children. Across gender, we found evidence for full configural, full metric, and partial scalar invariance. One item (i.e., “I don’t like watching TV programs that have a lot of violence in them”) was variant, with a higher intercept for girls. Finally, across child and caregiver reports, we found evidence for configural invariance, partial metric invariance (three items were variant with mixed directions), and partial scalar invariance across child and caregiver reports (two items were variant with mixed directions; see Table 2).

<sup>3</sup> For all steps of the measurement invariance analyses, we fixed the residual variance of the EOE item “I find it unpleasant to have a lot going on at once” to 20% of the observed item variance in the elementary school children group.

**Table 2** Study 1 Measurement Invariance Analyses Results Within Child Reports and Between Child and Caregiver Reports

Invariance Model	CFI	RMSEA	SRMR	AIC	BIC	ΔCFI	ΔRMSEA	ΔSRMR	Non-invariant items
<b>Within Child Reports</b>									
<i>Elementary versus Middle School Children</i>									
Configural	0.926	0.041	0.033	138717.982	139286.182				
Metric	0.920	0.039	0.037	138729.216	139177.794	0.006	0.002	0.004	
Scalar	0.905	0.041	0.040	138764.471	139165.201	0.015	0.002	0.003	
Partial scalar	0.911	0.040	0.039	138748.015	139154.726	0.006	0.001	0.001	HSC6 <sup>+</sup>
<i>Boys versus Girls</i>									
Configural	0.929	0.041	0.033	133285.877	133855.900				
Metric	0.934	0.036	0.036	133278.825	133730.092	0.005	0.005	0.003	
Scalar	0.904	0.041	0.039	133357.708	133761.474	0.030	0.005	0.003	
Partial scalar	0.932	0.035	0.036	133276.784	133686.488	0.002	0.001	0.000	HSC9 <sup>+</sup>
<b>Between Child and Caregiver Reports</b>									
Configural	0.978	0.021	0.030	43641.499	44099.815				
Metric	0.931	0.033	0.051	43667.783	44029.611	0.047	0.012	0.021	
Partial metric	0.969	0.023	0.039	43635.643	44011.944	0.009	0.002	0.009	HSC6*, HSC8*, HSC1*
Scalar	0.893	0.040	0.044	43676.709	44014.416	0.076	0.017	0.005	
Partial scalar	0.966	0.023	0.039	43634.567	43981.922	0.003	0.000	0.000	HSC11 <sup>+</sup> , HSC2 <sup>+</sup>

Note. HSC=Highly Sensitive Child total scale; CFI=comparative fit index; RMSEA=root mean square error of approximation; SRMR=standardized root mean squared residual; AIC=Akaike information criterion; BIC=Bayesian information criterion; Δ=absolute change value between consecutive invariance models. \* Items with invariant factor loadings. + Items with invariant intercepts. The numbering of items we used was based on Table 2 of the original publication of the HSC scale (for the exact content of these items, see Pluess et al., 2018, p. 55)

## 2.3 Discussion

Most Study 1 findings support the expected psychometric properties of the Chinese HSC scale. We found a bifactor structure and acceptable total scale internal consistency for both child- and caregiver-reported HSC, and partial invariance across age groups, gender, and informants. However, unexpectedly, we also found that the internal consistency of the HSC subscales was generally low for both child and caregiver reports. Overall, most Study 1 findings replicate findings from international samples and extend these to Chinese elementary school children. In Study 2, we further extend current work on our HSC scale by examining its convergent validity using self-reports from both elementary and middle school children.

## 3 Study 2

Study 2 investigated, for the first time, whether the bivariate associations between the self-reported Chinese HSC (sub)scales and related temperament and personality measures replicate prior findings found in Western samples. Based on theory and Western findings, we expected the HSC total scale to be positively associated with Behavioral Inhibition System (BIS), Behavioral Activation System (BAS), Negative Emotionality (NE), Positive Emotionality (PE), Neuroticism, and Openness, but negatively with Extraversion. For the EOE and LST subscales, we expected larger associations with measures that reflect sensitivity to more negative experiences (i.e., BIS, NE, Neuroticism) than with measures that reflect sensitivity to more positive experiences (i.e., BAS, PE, Extraversion, Openness). For the AES subscale, we expected the opposite pattern to that of EOE and LST. We did not have expectations on how the HSC (sub)scales would associate with Agreeableness and Conscientiousness because of unclear predictions from theory and mixed findings (i.e., Pluess et al., 2018; Weyn et al., 2021). We examined the expected associations in two subsamples of Chinese elementary school and middle school children.

### 3.1 Method

#### 3.1.1 Participants

Study 2 included two convenience subsamples from Study 1: one elementary school subsample and one middle school subsample. Children in each subsample completed the HSC scale. They completed a different set of additional temperament and personality measures, generally in line with previous research conducted in these two age groups (i.e., temperament measures for younger children and the Big-Five personality traits for older children; Pluess et al., 2018).

The elementary school subsample included 845 children ( $M_{\text{age}} = 9.71$  years;  $SD = 1.10$ ; range = 6.92–12.75; 41.9% girls, 57.5% boys, 0.6% not reported), recruited from one school in a small city located in central China. Additional measures included BIS, BAS, PE, and NE. Missingness was low (0.8%). We therefore used pairwise deletion in SPSS to handle missing data.

The middle school subsample included 563 children ( $M_{\text{age}} = 13.17$  years;  $SD = 1.10$ ; range = 11–15.75; 43.2% girls, 56.7% boys, 0.2% not reported), recruited from two schools in two small cities located in central and southeastern China. Measures included BIS, BAS, NE, PE, and—additional to the elementary school sample—all Big-Five personality traits. About one-third of the subsample (i.e., 36.2%) were not administered the scales for NE, PE, Agreeableness, and Conscientiousness because administrators of one school requested shorter questionnaires. Missing data<sup>4</sup> were therefore imputed using the multiple imputation feature of *Mplus*, using Bayesian estimation to predict the plausible missing values from all observed data (Asparouhov & Muthén, 2010). Key covariates<sup>5</sup> and all study variables were included in the imputation model. Fifty imputed datasets were generated and used for all subsequent analyses, using *Mplus* and SPSS to obtain the pooled results.

### 3.1.2 Measures

**Sensory Processing Sensitivity.** The HSC scale to assess SPS is described in Study (1) We here examine its internal consistency for the two subsamples used in Study (2) Internal consistency for the elementary school subsample was lower than expected:  $\alpha = 0.41$  and  $\omega = 0.53$  for HSC,  $\alpha = 0.42$  and  $\omega = 0.50$  for EOE,  $\alpha = 0.48$  and  $\omega = 0.50$  for LST, and  $\alpha = 0.39$  and  $\omega = 0.41$  for AES. Internal consistency for the middle school subsample was also somewhat lower than in Study 1, but still had acceptable  $\omega$  for the total scale:  $\alpha = 0.53$  and  $\omega = 0.61$  for HSC,  $\alpha = 0.48$  and  $\omega = 0.55$  for EOE,  $\alpha = 0.51$  and  $\omega = 0.54$  for LST, and  $\alpha = 0.41$  and  $\omega = 0.43$  for AES.

**Behavioral Inhibition and Activation.** We used the 20-item Behavioral Inhibition and Behavioral Activation Scales (BIS-BAS; Carver & White 1994; Muris et al., 2005). The BIS includes 7 items (e.g., “I worry about making mistakes”). The BAS includes 13 items (e.g., “I crave for excitement and new sensations”). Items were rated on a 4-point Likert scale ranging from 1 = *not at all true* to 4 = *very true*. Items were averaged to create total scores for BIS and BAS. Descriptive statistics and Cronbach’s  $\alpha$ s of all personality and temperament measures are provided in Table 3.

**Negative and Positive Emotionality.** We used the Early Adolescent Temperament Questionnaire-Revised (EATQ-R; Capaldi & Rothbart 1992). Negative Emotionality was computed as the average of items from the Fear, Frustration, and Shyness subscales (e.g., “I worry about getting into trouble”). Positive Emotionality was computed as the average of items from the Surgency, Pleasure Sensitivity, Perceptual Sensitivity, and Affiliation subscales (e.g., “I wouldn’t be afraid to try something like mountain climbing”). Items were rated on a 5-point Likert scale, ranging from 1 = *almost always untrue of you*, to 5 = *almost always true of you*.

**The Big-Five Personality Traits.** We assessed Agreeableness, Extraversion, Neuroticism, Openness, and Conscientiousness using the 60-item Chinese Big Five Personality Inventory (Zhou et al., 2017). Items (e.g., “I like to play with my class-

<sup>4</sup> Missing data were nearly exclusively due to certain measures not administered in one school, with in total only 8 responses missing due to participants across all study variables.

<sup>5</sup> Covariates in the imputation model included school location, father education level, mother education level, child gender, and child age.

**Table 3** Study 2 Descriptive Statistics and Cronbach’s Alphas of Personality and Temperament Measures, and Bivariate Zero-Order and Partial Correlations Between Child-Reported HSC (Sub)scales and Personality and Temperament Measures

	<i>M</i> ( <i>SD</i> )	<i>α</i>	HSC		EOE		LST		AES	
			Zero-order	Zero-order	Partial	Zero-order	Partial	Zero-order	Partial	
<i>Elementary School</i>										
BIS	2.72 (0.54)	0.55	0.27***	0.27***	0.25***	0.15***	0.11**	0.04	0.06	
BAS	2.77 (0.45)	0.67	0.11**	0.04	0.04	0.01	0.01	0.16***	0.16***	
NE	3.02 (0.58)	0.71	0.28***	0.32***	0.31***	0.14***	0.09*	0.01	0.03	
PE	3.23 (0.49)	0.60	0.15***	−0.03	−0.04	0.05	0.07*	0.29***	0.29***	
<i>Middle School</i>										
BIS	2.78 (0.46)	0.51	0.39***	0.29***	0.25***	0.17***	0.11**	0.30***	0.28***	
BAS	2.73 (0.47)	0.76	0.21***	0.19***	0.19***	−0.11**	−0.17***	0.30***	0.30***	
NE	3.11 (0.59)	0.77	0.33***	0.32***	0.29***	0.14	0.08	0.13*	0.10	
PE	3.30 (0.50)	0.68	0.10	−0.04	−0.08***	0.01	0.02	0.26***	0.27***	
Extra- ver- sion	3.26 (0.81)	0.86	−0.10*	−0.16***	−0.15***	−0.19***	−0.17***	0.19***	0.22***	
Agree- able- ness	3.66 (0.68)	0.76	0.08	−0.05	−0.09	0.15	0.16	0.08	0.09	
Con- scien- tious- ness	3.16 (0.61)	0.84	−0.02	−0.13*	−0.15**	0.04	0.07	0.10	0.12*	
Open- ness	3.27 (0.67)	0.82	0.04	−0.10*	−0.12**	−0.02	−0.01	0.24***	0.25***	
Neu- roti- cism	3.06 (0.83)	0.86	0.29***	0.33***	0.32***	0.05	−0.02	0.13**	0.10*	

*Note.* HSC=Highly Sensitive Child total scale; EOE=Ease of Excitation; LST=Low Sensory Threshold; AES=Aesthetic Sensitivity; BIS=Behavioral Inhibition System; BAS=Behavioral Activation System; NE=Negative Emotionality; PE=Positive Emotionality. Missing data were handled with multiple imputation in *Mplus* for the middle school subsample. Coefficients and significance are based on pooled results of 50 imputed data sets. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

mates,” “I always worry that something bad will happen.”) were rated on a 5-point Likert scale ranging from 1=*not at all true* to 5=*very true*. Items were averaged to create total scores for each trait.

### 3.1.3 Data Analysis

All analyses were run separately for the elementary school and middle school subsamples. To examine convergent validity of the HSC (sub)scales with related measures, we examined bivariate zero-order correlations between the HSC (sub)scales and related measures (i.e., BIS, BAS, PE, NE, and the Big Five Personality traits). For subscales, we additionally reported their partial correlations with related measures controlling for the contribution of the other two HSC subscales. We considered a correlation coefficient  $r$  from 0.10 as small, 0.30 as medium, and 0.50 as large (Cohen, 1992). If correlations were equal to or higher than  $r = .50$ , this would indicate an issue with divergent validity (Sperati et al., 2022). Last, to estimate how much of the variance of the HSC and its subscales was accounted for by related temperament and personality measures, we inspected the explained variance of the HSC (sub)scales by running a series of multiple regression models with all related measures simultaneously included as predictors of the HSC (sub)scales.

## 3.2 Results

### 3.2.1 Bivariate Correlations

Table 3 presents bivariate zero-order and partial correlations between child-reported HSC (sub)scales and related temperament and personality measures. Convergent validity of the total scale was generally supported by findings from both the elementary school subsample (BIS, BAS, NE, and PE) and middle school subsample (same measures plus the Big Five traits). HSC was positively associated with BIS, BAS, NE, PE, and neuroticism, and negatively with Extraversion. However, unexpectedly, Openness was not related to HSC. No associations were above 0.50, supporting divergent validity.

Convergent validity was also clearly supported for most subscales. EOE had stronger associations with more negative traits (i.e., BIS, NE, and Neuroticism) than with more positive traits (i.e., BAS, PE, Extraversion, and Openness). Moreover, EOE's associations with more negative traits remained of similar magnitude and significant after partialling out the other two subscales. As expected, AES showed the inverse pattern of associations to that of EOE (except that it was also moderately associated with BIS in the middle school sample). Moreover, AES's associations with more positive traits also remained of similar magnitude and significant after partialling out the other two subscales. Findings for LST were somewhat mixed. In the elementary school sample, they followed the expected pattern of stronger associations with more negative versus positive traits. Associations of LST with more negative traits remained significant but decreased after partialling out the other two subscales. However, in the middle school sample, associations were stronger for positive versus negative traits, and associations with both negative and positive traits remained significant (although decreased for negative traits) after partialling out the other two subscales.



### 3.2.2 Multivariate Regression

We inspected how much variance of the HSC (sub)scales would be explained by related temperament and personality measures. In the elementary school subsample, BIS, BAS, NE, and PE combined explained a significant amount of variance for the HSC (sub)scales: 12%, 12%, 3%, and 9% for HSC, EOE, LST, and AES respectively. In the middle school sample, explained variance was also significant: BIS, BAS, NE, PE, and the Big Five traits combined explained 24%, 21%, 15%, and 20% of the variance of HSC, EOE, LST, and AES, respectively (see Table 4 for parameter estimates). Across two subsamples, explained variance was lower for LST than for other (sub)scales.

### 3.3 Discussion

Study 2 findings clearly supported convergent validity for the total scale and the EOE and AES subscales and partially supported convergent validity for LST. Study 2 also suggests that the HSC (sub)scales may still have good validity despite their low internal consistencies in the Study 2 subsamples.

## 4 General Discussion

Sensory Processing Sensitivity (SPS) and the term “highly sensitive” have become increasingly popular both within and outside academia (Hellwig & Roth, 2021). However, scientific knowledge of SPS is mostly based on findings from Western cultures (Greven et al., 2019). It is important to promote examination of the cultural generalizability of SPS, that is, whether the SPS trait captures heightened responsiveness to both positive and negative environments across cultures. We therefore conducted two studies to examine our Chinese translation of the HSC scale. In Study 1, we examined the psychometric properties (i.e., factor structure, internal consistency, and measurement invariance) of the child- and caregiver-reported HSC. In Study 2, we examined convergent validity of the self-reported HSC with related temperament and personality measures. We extend previous research by also including Chinese elementary school children, examining child as well as caregiver reports, and investigating convergent validity of the self-reported HSC in Chinese children.

### 4.1 Psychometric Properties of the Chinese Child- and Caregiver-Reported HSC Scales

Study 1 replicated most psychometric properties of the child- and caregiver-reported HSC found in previous international studies. That is, results supported a bifactor structure of the HSC scale, acceptable internal consistency for the total scale (but not the subscales), and partial measurement invariance across age groups, gender, and informants. Overall, these findings suggest that our HSC scale—or at least, the total scale—can be used to examine SPS in Chinese elementary and middle school children.

**Table 4** Study 2 Multivariate Regression Analyses Predicting the HSC Total Scale and Subscales for Each Subsample

	HSC			EOE			LST			AES		
	B	B SE	$\beta$	B	B SE	$\beta$	B	B SE	$\beta$	B	B SE	$\beta$
<i>Elementary School</i>												
BIS	0.20	0.05	0.15***	0.28	0.08	0.14***	0.27	0.11	0.10*	0.03	0.08	0.01
BAS	-0.01	0.06	-0.01	-0.08	0.08	-0.04	-0.16	0.12	-0.05	0.19	0.09	0.08*
NE	0.23	0.05	0.20***	0.46	0.07	0.26***	0.23	0.10	0.10*	-0.05	0.07	-0.03
PE	0.18	0.05	0.13***	-0.10	0.07	-0.05	0.14	0.10	0.05	0.57	0.08	0.27***
R <sup>2</sup>	0.12***			0.12***			0.03***			0.09***		
<i>Middle School</i>												
BIS	0.40	0.08	0.28***	0.31	0.11	0.14**	0.29	0.17	0.11	0.59	0.10	0.30***
BAS	0.20	0.07	0.14**	0.36	0.10	0.17***	-0.29	0.13	-0.11*	0.36	0.09	0.19***
NE	0.14	0.08	0.15	0.27	0.10	0.18**	0.21	0.22	0.12	-0.06	0.08	-0.05
PE	0.03	0.10	0.02	-0.12	0.11	-0.07	0.09	0.24	0.04	0.16	0.11	0.10
Extraversion	-0.11	0.04	-0.13*	-0.14	0.06	-0.11*	-0.30	0.09	-0.20**	0.06	0.06	0.05
Agreeableness	0.10	0.07	0.11	0.06	0.08	0.05	0.30	0.19	0.18	-0.01	0.07	-0.01
Conscientiousness	-0.12	0.07	-0.12	-0.19	0.10	-0.13*	-0.13	0.18	-0.08	-0.03	0.08	-0.02
Openness	0.06	0.06	0.06	-0.05	0.08	-0.04	0.13	0.13	0.07	0.15	0.07	0.12*
Neuroticism	0.02	0.04	0.02	0.13	0.06	0.11*	-0.10	0.10	-0.07	-0.03	0.06	-0.03
R <sup>2</sup>	0.24***			0.21***			0.15***			0.20***		

Note. HSC=Highly Sensitive Child total scale; EOE=Ease of Excitation; LST=Low Sensory Threshold; AES=Aesthetic Sensitivity; BIS=Behavioral Inhibition System; BAS=Behavioral Activation System; NE=Negative Emotionality; PE=Positive Emotionality. Missing data were handled with multiple imputation in Mplus for the Middle school subsample. Coefficients and significance are based on pooled results of 50 imputed data sets. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

We found clear support for the bifactor structure of the data. Thus, the HSC total score captures an overall trait of sensitivity, whereas the three subscales explain additional variance in specific sensitivity aspects. This suggests that Chinese children may differ in their general sensitivity, but also in the extent to which they are easily aroused by external stimuli (Low Sensory Threshold; LST), overwhelmed by external and internal demands (Ease of Excitation; EOE), and stimulated by aesthetic stimuli (Aesthetic Sensitivity; AES) (Pluess et al., 2018). Bifactor-specific reliability indices support the use of both a general and specific sensitivity factors. For both informants, the general sensitivity factor explained an overwhelming proportion of the reliable variance of the HSC total score, supporting the use of raw total HSC scores as a measure of general sensitivity. Bifactor-specific reliability indices for the subscales showed that LST and AES explained additional reliable variance. However, the EOE specific factor barely explained any reliable variance. Such results were consistent with one previous study (Weyn et al., 2021) and may be because the EOE subscale contains several items with nonsignificant loadings for the EOE specific factor. Overall, the bifactor structure supports the use of the HSC scale to assess both general and specific aspects of sensitivity in Chinese children, although the EOE subscale may need some improvement (for a recent example, see: Weyn et al., 2022).

Internal consistency for the HSC total scale was supported in most samples. In the overall sample of Study 1, the HSC total scale had acceptable internal consistency for both child and caregiver reports. However, Study 2 revealed acceptable total scale internal consistency (i.e., McDonald's omega) only for the middle school subsample, and not for the elementary school subsample. Although this result is in line with other studies finding relatively lower internal consistencies for younger versus older children (Pluess et al., 2018; Weyn et al., 2021), it still questions the use of our HSC total scale for elementary school children. Possibly, the 7-point Likert scale with three anchors (i.e., 1 = *not at all*, 4 = *moderately*, and 7 = *extremely*) made it difficult for younger children to understand and answer the items. This might be especially true for the elementary school children from our subsample, who came from a small city in central China, had no previous experience with filling out questionnaires, and were quite young. Researchers studying elementary school children may thus consider using our caregiver-reported HSC total scale, which did show acceptable internal consistency in that sample.

Internal consistency for the HSC subscales was generally low across studies and informants. This, albeit occasionally found in previous research (e.g., Pluess et al., 2018, Study 3; Weyn et al., 2021; Yano et al., 2021), did not replicate most previous research (Dong et al., 2022; Pluess et al., 2018; Sperati et al., 2022). Low internal consistency for the HSC subscales may be due to (1) the small numbers of items, and (2) the extremeness (e.g., “I love nice smells/tastes” instead of “I like nice smells/tastes”) and negative wordings (e.g., I *don't* like it when things change in my life) of some items that might have caused high mean scores and low variation in responses (see Weyn et al., 2022 for details). Supporting this reasoning, a recent study has found improved internal consistency of the HSC subscales after solving the aforementioned limitations (e.g., by adding more items and omitting items with extremeness and negative wordings; Weyn et al.,

2022). Taken together, our results suggest that the self-reported HSC total scale for middle school children and caregiver-reported HSC total scale for elementary school children are quick, convenient, and reliable measures of general sensitivity. However, regarding our subscales, an improved version may be needed before one can confidently use them (see Weyn et al., 2022).

Our measurement invariance results suggest that our HSC can be used across age groups, gender and informants. We found full configural invariance, suggesting that the concepts of general sensitivity, as well as three sensitivity components, are shared across age groups, gender, and informants. Moreover, we found full metric invariance across age groups and gender, suggesting that similar meaning is attributed to the HSC items across elementary and middle school children, and across boys and girls. Yet, scalar invariance was only partial, suggesting different reference points were used for the non-invariant item across elementary and middle school children, and across boys and girls. Specifically, across age group, the item *I notice it when small things have changed in my environment* had a higher intercept among elementary than middle school children. This indicates that given the same level of general sensitivity, elementary school children had a higher score than middle school children on this item (Chen et al., 2019). Possibly, young children are more curious about their surroundings and thus notice more subtleties. Across gender, the item *I don't like watching TV programs that have a lot of violence in them* had a higher intercept among girls than boys, indicating that given the same level of general sensitivity, girls had a higher score than boys on this item. This might be due to gender differences in violent media usage, with boys watching more violent TV than girls and having stronger attraction to TV violence (Rosenkoetter et al., 2004). Finally, across child and caregiver reports, only partial metric and partial scalar invariance were achieved, with 3 and 2 items exhibiting metric and scalar non-variance, respectively. Thus, caregivers and children differed in the meaning they attributed to the 3 invariant items as well as the reference point they used for the 2 invariant items (Weyn et al., 2022). Taken together, given that we did not achieve full metric invariance across any group, researchers using the HSC should refrain from comparing *observed* means (i.e., sum/mean scores) across age groups, gender, and informants (Steinmetz, 2013). Instead, researchers could compare *latent* means to capture children's SPS in samples including different age groups, genders, and informants (Schmitt et al., 2011; Steinmetz, 2013).

#### 4.2 Associations of the Self-Reported Chinese HSC Scale with Related Temperament and Personality Measures

Study 2 replicated most patterns of convergent correlations found in previous Western studies between the self-reported HSC (sub)scales and related temperament and personality measures: Behavioral Inhibition System (BIS), Behavioral Activation System (BAS), Negative Emotionality (NE), Positive Emotionality (PE), Neuroticism, Openness, and Extraversion. Convergent validity was clearly supported for the HSC total scale, given that 6 out of 7 associations were in the expected directions. This suggests that the associations between the HSC and

other related temperament and personality measures are largely alike in Chinese and Western children. We also found clear support for convergent validity of the EOE and AES subscales: EOE was more strongly associated with negative traits (i.e., BIS, NE, Neuroticism), whereas AES was more strongly associated with positive traits (i.e., BAS, PE, Openness, and Extraversion). This suggests that Chinese and Western children not only share the same sensitivity components of EOE and AES — as evidenced by the CFA results — but may also share the implications of EOE and AES. For the LST subscale, we found partial support for convergent validity. We expected to find larger associations of LST with negative versus positive traits, which we found in the elementary school subsample, but not the middle school subsample. Here, LST exhibited slightly larger associations with positive than negative traits — a pattern of results that was also reported in previous research in Western samples (Sperati et al., 2022; Weyn et al., 2021). Last, Study 2 demonstrated that the HSC (sub)scales, despite relating meaningfully to other established measures, were not fully captured by these measures. Overall, findings suggest that our self-report HSC (sub)scales—or at least, the HSC total scale and the EOE and AES subscales—have good convergent and divergent validity in Chinese children.

A recent trend in SPS research is the effort to distinguish between the negative versus positive component of sensitivity (De Gucht et al., 2022; Weyn et al., 2022), or the so-called “dark” versus “bright side” of sensitivity (Sperati et al., 2022). Specifically, EOE and LST may represent the “dark side” of sensitivity capturing greater sensitivity to *negative* environments, whereas AES may represent the “bright side” of sensitivity capturing greater sensitivity to *positive* environments (Pluess et al., 2018; Sperati et al., 2022; Weyn et al., 2021). In line with this notion, a recent study using classic twin design methodology showed that the genetic influences underlying EOE and LST are relatively distinct from AES (Assary et al., 2021). Further evidence comes from a validation study showing that collapsing EOE and LST into one factor provided a better fit to the data (Weyn et al., 2022). Turning to our results, we found that positive associations of EOE with negative traits (i.e., BIS, NE, Neuroticism) and positive associations of AES with positive traits (i.e., BAS, PE, Extraversion, Conscientiousness, Agreeableness, and Openness) remained of similar magnitude and significant after controlling for the other two subscales. Our findings thus clearly support EOE and AES as the “dark” and “bright side” of sensitivity, respectively.

As for LST, converging evidence from our and others’ research seems to suggest that it may not represent the “dark” or “bright side” of sensitivity per se. Our and others’ findings suggest that associations between LST and negative traits may be mostly accounted for by EOE (Sperati et al., 2022; Weyn et al., 2021). However, LST does differ from EOE in that it had positive associations with more positive traits (e.g., Conscientiousness, Effortful Control), whereas EOE mostly had negative associations with positive traits. Moreover, multiple regression analyses revealed much lower amounts of explained variance for LST than for the other HSC (sub) scales (Pluess et al., 2018; Sperati et al., 2022), suggesting that “LST may capture aspects that are more specific to sensitivity and not otherwise reflected in existing temperament questionnaires” (Sperati et al., 2022, p. 8). Similarly, a study in adults also found that LST was the only SPS subscale that seems “not fully explained by

established personality traits” (Hellwig & Roth, 2021, p. 10). Collectively, it would thus be important for future research to further clarify the role of LST and which types of environments it may enhance children’s sensitivity to.

### 4.3 Strengths and Limitations

This research has several strengths. It is the first study investigating the psychometric properties of the Chinese version of the HSC scale that included (a) Chinese elementary school children, (b) both self- as well as caregiver reports, and (c) a wide range of personality and temperament measures to investigate convergent validity. Moreover, we were able to examine measurement invariance of our Chinese HSC scale across age groups, gender, and informants.

This research also has several limitations. First, all measures were based on self- and caregiver reports. To reduce shared method variance and possible social desirability bias, it would be useful for future research to examine the association of our Chinese HSC scale with observer-rated environmental sensitivity measures (e.g., Lionetti et al., 2019), cognitive attention tasks (e.g., change detection task; Jagiellowicz et al., 2011), genetic markers (e.g., dopamine-related genes; Chen et al., 2011), or physiological markers (e.g., heart rate variability; Miller et al., 2021). Second, our convergent validity measures were reported by children only. It would be informative to have caregivers also report on other temperament and personality measures, which would have enabled us to also investigate the convergent validity of caregiver-reported HSC. Third, although the HSC scale showed acceptable total scale internal consistency in Study 1, we found low internal consistency for the HSC total scale in Study 2 and for the HSC subscales in both Studies 1 and 2. Despite these lower reliabilities, however, we generally still found well-supported convergent validity of the self-reported HSC (sub)scales. Indeed, researchers have suggested that “When effects are significant despite low reliability, this implies either that effects are very large in the population or that the reliability is actually higher than the alpha indicates” (p. 1270, Keijzer et al., 2022). Finally, we did not examine whether our HSC scale moderates the association between negative and positive environments and outcomes (i.e., criterion validity), which will need to be corroborated by future research.

### 4.4 Implications

The availability of a validated Chinese SPS questionnaire may be relevant for both research and practice. First, it may facilitate cross-cultural research on SPS. For example, researchers could examine measurement invariance of SPS across China and other countries or examine differences in SPS and its associations with other measures between China and other countries. Such research may reveal potential SPS-related cultural differences (e.g., different connotations of items) which may facilitate adjusting the SPS questionnaire to better fit within the Chinese context. Second, it may spur research examining the cultural generalizability of SPS: does this trait moderate positive and negative environmental influences in Chinese children as it does in western children? Such research may facilitate

intervention programs to use the SPS questionnaire as a tool to early identify children who are more at risk but also more likely to exceptionally benefit from enriched environments in the Chinese context.

## 4.5 Conclusion

Overall, our Chinese translation of the HSC replicated most psychometric properties found in international studies: (a) a bifactor structure with one general sensitivity component and three specific components, (b) acceptable internal consistency of the total scale (although not for self-report of elementary school children, and not for the subscales), and (c) at least partial invariance across age groups, gender, and informants. Our findings also supported convergent validity of the Chinese HSC as found in Western studies, suggesting that EOE may capture the “dark side” of sensitivity, whereas AES may capture the “bright side” of sensitivity. Taken together, we recommend that researchers using our Chinese HSC could: (a) use the HSC total scale to assess general sensitivity, but be careful to use the less reliable subscales to investigate the different sensitivity components; (b) use child reports for middle school children and caregiver reports for elementary school children; and (c) use latent scores when comparing means across age groups, gender, and informants. Given that examining cultural differences in SPS is an important and yet understudied topic (Greven et al., 2019), we hope that our studies may help spur more future research on SPS in cross-cultural contexts.

**Funding** This research was supported by the China Scholarship Council Grant (201906870051) awarded to Danni Liu and the National Natural Science Foundation of China (31971004) awarded to Zhenhong Wang.

**Data Availability** All data that support the findings of this study are available through the Open Science Framework at <https://osf.io/rgan7/>.

**Code Availability** All syntax of analyses run for this study are available through the Open Science Framework (see link above).

## Compliance with Ethical Standards

**Conflict of interest** We have no known conflict of interest to disclose.

**Research Involving Human Participants and/or Animals** All procedures performed in this study were in line with the principles of the Declaration of Helsinki. This study was approved by the Faculty Ethics Committee of the Social and Behavioural Sciences Faculty of Utrecht University.

**Informed Consent** Informed consent was obtained from all individual participants of the study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission



directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aron, E. N., & Aron, A. (1997). Sensory-processing sensitivity and its relation to introversion and emotionality. *Journal of Personality and Social Psychology*, *73*, 345–368. <https://doi.org/10.1037/0022-3514.73.2.345>.
- Aron, E. N., Aron, A., & Jagiellowicz, J. (2012). Sensory Processing sensitivity: A review in the light of the evolution of Biological Responsivity. *Personality and Social Psychology Review*, *16*(3), 262–282. <https://doi.org/10.1177/1088868311434213>.
- Asparouhov, T., & Muthén, B. (2010). Multiple imputation with Mplus. *MPlus Web Notes*, *29*, 238–246.
- Assary, E., Zavos, H. M. S., Krapohl, E., Keers, R., & Pluess, M. (2021). Genetic architecture of environmental sensitivity reflects multiple heritable components: A twin study with adolescents. *Molecular Psychiatry*, *26*(9), 4896–4904. <https://doi.org/10.1038/s41380-020-0783-8>.
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456–466.
- Capaldi, D. M., & Rothbart, M. K. (1992). Development and validation of an early adolescent temperament measure. *The Journal of Early Adolescence*, *12*(2), 153–173. <https://doi.org/10.1177/0272431692012002002>.
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319–333. <https://doi.org/10.1037/0022-3514.67.2.319>.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 464–504.
- Chen, C., Chen, C., Moyzis, R., Stern, H., He, Q., Li, H., Li, J., Zhu, B., & Dong, Q. (2011). Contributions of dopamine-related genes and environmental factors to highly sensitive personality: A multi-step neuronal System-Level Approach. *PLOS ONE*, *6*(7), e21636. <https://doi.org/10.1371/journal.pone.0021636>.
- Chen, M., Zee, M., Koomen, H. M. Y., & Roorda, D. L. (2019). Understanding cross-cultural differences in affective teacher-student relationships: A comparison between dutch and chinese primary school teachers and students. *Journal of School Psychology*, *76*, 89–106. <https://doi.org/10.1016/j.jsp.2019.07.011>.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>.
- Costa-López, B., Ruiz-Robledillo, N., Albaladejo-Blázquez, N., Baryła-Matejczuk, M., & Ferrer-Cascales, R. (2022). Psychometric Properties of the Spanish Version of the highly sensitive child scale: The parent version. *International Journal of Environmental Research and Public Health*, *19*(5), <https://doi.org/10.3390/ijerph19053101>.
- Davies, P. T., Hentges, R. F., Coe, J. L., Parry, L. Q., & Sturge-Apple, M. L. (2021). Children’s dove temperament as a differential susceptibility factor in child rearing contexts. *Developmental Psychology*, *57*, 1274–1290. <https://doi.org/10.1037/dev0001215>.
- De Gucht, V., Woestenburg, D. H. A., & Wilderjans, T. F. (2022). The different faces of (high) sensitivity, toward a more comprehensive measurement instrument. Development and validation of the sensory Processing Sensitivity Questionnaire (SPSQ). *Journal of Personality Assessment*, *104*(6), 784–799. <https://doi.org/10.1080/00223891.2022.2032101>.
- DeSimone, J. A., Harms, P. D., & DeSimone, A. J. (2015). Best practice recommendations for data screening. *Journal of Organizational Behavior*, *36*, 171–181.
- Dong, Q., Zhou, L., Wang, W., Wei, X., Pluess, M., & Ma, X. (2022). Development and validation of the Chinese Version of the highly sensitive child scale: Understanding environmental sensitivity and depressive symptoms in adolescents. *Frontiers in Psychiatry*, *13*, 999150. <https://doi.org/10.3389/fpsy.2022.999150>.

- Dunn, T. J., Baguley, T., & Brunsten, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. <https://doi.org/10.1111/bjop.12046>.
- Enders, C., & Bandalos, D. (2001). The relative performance of full information Maximum Likelihood Estimation for Missing Data in Structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 430–457. [https://doi.org/10.1207/S15328007SEM0803\\_5](https://doi.org/10.1207/S15328007SEM0803_5).
- Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which Coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, *3*(4), 484–501. <https://doi.org/10.1177/2515245920951747>.
- Green, S. B., & Yang, Y. (2015). Evaluation of dimensionality in the assessment of internal consistency reliability: Coefficient alpha and omega coefficients. *Educational Measurement: Issues and Practice*, *34*, 14–20.
- Greven, C. U., Lionetti, F., Booth, C., Aron, E. N., Fox, E., Schendan, H. E., Pluess, M., Bruining, H., Acevedo, B., Bijtbeier, P., & Homberg, J. (2019). Sensory Processing sensitivity in the context of environmental sensitivity: A critical review and development of research agenda. *Neuroscience & Biobehavioral Reviews*, *98*, 287–305. <https://doi.org/10.1016/j.neubiorev.2019.01.009>.
- Hellwig, S., & Roth, M. (2021). Conceptual ambiguities and measurement issues in sensory processing sensitivity. *Journal of Research in Personality*, *93*, 104130. <https://doi.org/10.1016/j.jrp.2021.104130>.
- Jagiellowicz, J., Xu, X., Aron, A., Aron, E., Cao, G., Feng, T., & Weng, X. (2011). The trait of sensory processing sensitivity and neural responses to changes in visual scenes. *Social Cognitive and Affective Neuroscience*, *6*(1), 38–47. <https://doi.org/10.1093/scan/nsq001>.
- Keijzer, R., van Schooten, E., van der Rijst, R., & Admiraal, W. (2022). Individual characteristics of students in vocational education moderating the relationship between school engagement and vocational identity. *European Journal of Psychology of Education*, *37*(4), 1255–1283. <https://doi.org/10.1007/s10212-021-00580-y>.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford Press.
- Lionetti, F., Aron, E. N., Aron, A., Klein, D. N., & Pluess, M. (2019). Observer-rated environmental sensitivity moderates children's response to parenting quality in early childhood. *Developmental Psychology*, *55*(11), 2389–2402. <https://doi.org/10.1037/dev0000795>.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum.
- Miller, J. G., Chahal, R., Kirshenbaum, J. S., Ho, T. C., Gifuni, A. J., & Gotlib, I. H. (2021). Heart rate variability moderates the effects of COVID-19-related stress and family adversity on emotional problems in adolescents: Testing models of differential susceptibility and diathesis stress. *Development and Psychopathology*, 1–12. <https://doi.org/10.1017/S095457942100033X>.
- Muris, P., Meesters, C., de Kanter, E., & Timmerman, P. E. (2005). Behavioural inhibition and behavioural activation system scales for children: Relationships with Eysenck's personality traits and psychopathological symptoms. *Personality and Individual Differences*, *38*(4), 831–841. <https://doi.org/10.1016/j.paid.2004.06.007>.
- Pluess, M. (2015). Individual differences in environmental sensitivity. *Child Development Perspectives*, *9*(3), 138–143. <https://doi.org/10.1111/cdep.12120>.
- Pluess, M., & Boniwell, I. (2015). Sensory-Processing Sensitivity predicts treatment response to a school-based depression prevention program: Evidence of Vantage Sensitivity. *Personality and Individual Differences*, *82*, 40–45. <https://doi.org/10.1016/j.paid.2015.03.011>.
- Pluess, M., Assary, E., Lionetti, F., Lester, K. J., Krapohl, E., Aron, E. N., & Aron, A. (2018). Environmental sensitivity in children: Development of the highly sensitive child scale and identification of sensitivity groups. *Developmental Psychology*, *54*(1), 51–70. <https://doi.org/10.1037/dev0000406>.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. <https://doi.org/10.1037/met0000045>.
- Rosenkoetter, L. I., Rosenkoetter, S. E., Ozretich, R. A., & Acock, A. C. (2004). Mitigating the harmful effects of violent television. *Journal of Applied Developmental Psychology*, *25*(1), 25–47. <https://doi.org/10.1016/j.appdev.2003.11.005>.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. V. Eye, & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Sage.
- Schmitt, N., Golubovich, J., & Leong, F. T. L. (2011). Impact of Measurement Invariance on Construct Correlations, Mean Differences, and Relations with External Correlates: An illustrative example using big five and RIASEC measures. *Assessment*, *18*(4), 412–427. <https://doi.org/10.1177/1073191110373223>.

- Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment*, 31(12), 1428–1441. <https://doi.org/10.1037/pas0000623>.
- Sijtsma, K. (2009). On the Use, the Misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1), 107–120. <https://doi.org/10.1007/s11336-008-9101-0>.
- Slagt, M., Dubas, J. S., Deković, M., & van Aken, M. A. G. (2016). Differences in sensitivity to parenting depending on child temperament: A meta-analysis. *Psychological Bulletin*, 142, 1068–1110. <https://doi.org/10.1037/bul0000061>.
- Slagt, M., Dubas, J. S., van Aken, M. A. G., Ellis, B. J., & Deković, M. (2018). Sensory processing sensitivity as a marker of differential susceptibility to parenting. *Developmental Psychology*, 54(3), 543–558. <https://doi.org/10.1037/dev0000431>.
- Sperati, A., Spinelli, M., Fasolo, M., Pastore, M., Pluess, M., & Lionetti, F. (2022). Investigating sensitivity through the lens of parents: Validation of the parent-report version of the highly sensitive child scale. *Development and Psychopathology*, 1–14. <https://doi.org/10.1017/S0954579422001298>.
- Steinmetz, H. (2013). Analyzing observed Composite differences across groups: Is partial measurement Invariance Enough? *Methodology*, 9(1), 1–12. <https://doi.org/10.1027/1614-2241/a000049>.
- Weyn, S., Van Leeuwen, K., Pluess, M., Lionetti, F., Greven, C. U., Goossens, L., Colpin, H., Van Den Noortgate, W., Verschuere, K., Bastin, M., Van Hoof, E., De Fruyt, F., & Bijttebier, P. (2021). Psychometric properties of the highly sensitive child scale across developmental stage, gender, and country. *Current Psychology*, 40(7), 3309–3325. <https://doi.org/10.1007/s12144-019-00254-5>.
- Weyn, S., Van Leeuwen, K., Pluess, M., Lionetti, F., Goossens, L., Bosmans, G., Van Den Noortgate, W., Debeer, D., Bröhl, A. S., & Bijttebier, P. (2022). Improving the measurement of environmental sensitivity in children and adolescents: The highly sensitive child Scale–21 item version. *Assessment*, 29(4), 607–629. <https://doi.org/10.1177/1073191120983894>.
- Yano, K., Kase, T., & Oishi, K. (2021). The Associations between sensory Processing sensitivity and the big five personality traits in a Japanese sample. *Journal of Individual Differences*, 42(2), 84–90. <https://doi.org/10.1027/1614-0001/a000332>.
- Zeng, S., & Wang, Z. (2022). Revision of the Chinese Version of the highly sensitive child scale-parent report. *Studies of Psychology and Behavior*, 20(1), 101. <https://doi.org/10.12139/j.1672-0628.2022.01.015>.
- Zhou, Y., Li, D., Li, X., Wang, Y., & Zhao, L. (2017). Big five personality and adolescent internet addiction: The mediating role of coping style. *Addictive Behaviors*, 64, 42–48. <https://doi.org/10.1016/j.addbeh.2016.08.009>.
- UNPFA (2022). *State of World Population Report 2022*. <https://www.familyplanning.org.nz/news/2022/unpfa-state-of-world-population-report-2022>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Danni Liu<sup>1</sup> · Anouk van Dijk<sup>2</sup> · Shanyan Lin<sup>3</sup> · Zhenhong Wang<sup>4</sup> · Maja Deković<sup>5</sup> · Judith Semon Dubas<sup>1</sup>

✉ Danni Liu  
d.liu@uu.nl

<sup>1</sup> Department of Developmental Psychology, Utrecht University, Utrecht, The Netherlands

<sup>2</sup> Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands

- 
- <sup>3</sup> Department of Psychology, University of Turin, Torino, Italy
  - <sup>4</sup> School of Psychology, Shaanxi Provincial Key Research Center of Child Mental and Behavioral Health, Shaanxi Normal University, Xi'an, China
  - <sup>5</sup> Department of Clinical Child and Family Studies, Utrecht University, Utrecht, The Netherlands