Cultural Adaptation and Validity Testing of the Highly Sensitive Person Scale among the Working Population in China

ZHANG Yijun, PENG Dingjiang, YANG Mei, LIAO Jinjing

2026-08-26

ChinaRxiv

Abstract

Background Research on sensory processing sensitivity (SPS) is receiving increasing clinical attention. The Chinese version of the Highly Sensitive Person Scale (HSPS) currently used in China has mainly been tested among college students and urban populations, but this version has not yet been validated among the working-age labor population in China. Objective To obtain an SPS scale suitable for the Chinese working-age labor population by eliminating measurement bias related to educational level and place of birth. Methods Relevant items were revised according to the cognitive habits of the target population, and the initial scale was established through a pilot survey among people with low educational attainment. Stratified and purposive sampling was used to recruit participants from five labor-intensive enterprises in Guangdong, Chongqing, and Hunan, with a focus on frontline employees. A total of 2 015 valid questionnaires were collected (valid response rate: 80.6%) and randomly divided into an exploratory factor analysis group (EFA, n=1 007) and a confirmatory factor analysis group (CFA, n=1 008). After 6 months, 267 valid questionnaires were obtained for retesting. Python 3.13 and R were used to test reliability, validity, and measurement invariance, and to conduct between-group difference analysis and regression analysis in combination with demographic variables. Inverse probability weighting (IPW) was used to analyze the potential impact of gender imbalance. Results A 20-item Chinese version of the scale (HSPS-C20) was finalized. The three-dimensional bifactor model showed good fit: comparative fit index (CFI)=0.974, root mean square error of approximation (RMSEA)=0.038, and standardized root mean square residual (SRMR)=0.028. Cronbach' s α for the total scale was 0.923, and Cronbach' s α for each dimension ranged from 0.802 to 0.885. Test-retest reliability was acceptable [intraclass correlation coefficient (ICC) for the total scale=0.707, and ICCs for each dimension ranged from 0.442 to 0.620]. Bifactor dimensionality indices showed an explained common variance (ECV) of 0.742, a percent of uncontaminated correlations (PUC) of 0.679, and ω=0.924, supporting the total score as a unidimensional indicator. Strict measurement invariance was achieved across gender, educational level, and place-of-birth groups (ΔCFI<0.01). Gender and place of birth significantly predicted SPS scores (β=0.119-0.172, P<0.001), with a smaller effect size for place of birth (β=0.061-0.066, P<0.01). IPW sensitivity analysis indicated that gender imbalance did not substantially interfere with the conclusions of SPS assessment. Conclusion The revised HSPS-C20 has good reliability and validity and can serve as a reliable tool for assessing sensory processing sensitivity in the working-age labor population in China.

Full Text

Cultural Adaptation and Validity Testing of the Highly Sensitive Person Scale among the Working Population in China

ZHANG Yijun ^1*, PENG Dingjiang ^2, YANG Mei ^3, LIAO Jinjing ^1

1. 510010 Guangzhou, Guangdong Province, Department of Geriatrics, General Hospital of Southern Theater Command of the Chinese People' s Liberation Army

2. 510080 Guangzhou, Guangdong Province, Guangzhou University of Chinese Medicine

3. 401420 Chongqing, Community Health Service Center, Tonghui Street, Qijiang District

【Abstract】 **Background** Sensory processing sensitivity (SPS) has attracted increasing clinical attention. The Chinese version of the Highly Sensitive Person Scale (HSPS) currently used in China has mainly been tested among college students and urban populations, but this version has not yet been validated among the working-age labor population in China. **Objective** To obtain an SPS scale suitable for the Chinese working-age labor population by eliminating measurement bias related to education level and place of birth. **Methods** Relevant items were revised according to the cognitive habits of the target population, and an initial scale was established through a preliminary survey among groups with lower educational attainment. Stratified and purposive sampling was used to recruit participants from five labor-intensive enterprises in Guangdong, Chongqing, and Hunan, with frontline employees as the focus. A total of 2,015 valid questionnaires were collected (valid response rate: 80.6%) and randomly divided into two groups for exploratory factor analysis (EFA, n=1,007) and confirmatory factor analysis (CFA, n=1,008). Six months later, 267 valid retest questionnaires were obtained. Python 3.13 and R were used to test reliability and validity and measurement invariance, and demographic variables were incorporated for between-group difference analysis and regression analysis; inverse probability weighting (IPW) was used to analyze the potential impact of imbalance in sex proportions. **Results** A 20-item Chinese version of the scale was finally formed (HSPS-C20). The three-dimensional bifactor model showed good fit: comparative fit index (CFI) = 0.974, root mean square error of approximation ($RMSEA$) = 0.038, and standardized root mean square residual ($SRMR$) = 0.028. The total-scale Cronbach' s α was 0.923, and Cronbach' s α for each dimension ranged from 0.802 to 0.885; test–retest reliability [intraclass correlation coefficient (ICC) for the total scale = 0.707, with dimensional ICCs ranging from 0.442 to 0.620]. Bifactor dimensionality indices showed an explained common variance (ECV) of 0.742, percentage of uncontaminated correlations (PUC) of 0.679, and ω = 0.924, supporting use of the total score as a unidimensional indicator. Tests across sex, educational level, and place of birth all reached strict measurement invariance (ΔCFI<0.01). Sex and place of birth significantly predicted SPS scores (β=0.119–0.172, P<0.001), with the place-of-birth effect being smaller (β=0.061–0.066, P<0.01). IPW sensitivity analysis showed that sex imbalance did not substantially interfere with the SPS evaluation conclusions. **Conclusion** The revised HSPS-C20 has good reliability and validity and can serve as a reliable tool for assessing sensory processing sensitivity in China' s working-age labor population.

【Keywords】 sensory processing sensitivity; Highly Sensitive Person Scale; cross-cultural adaptation; working-age labor population; measurement invariance; reliability and validity testing

【Chinese Library Classification No.】 R 395 【Document Code】 A DOI: 10.12114/j.issn.1007-9572.2026.0095

Cultural adaptation and validity testing of the Highly Sensitive Person Scale among the working population in China

ZHANG Yijun ^1*, PENG Dingjiang ^2, YANG Mei ^3, LIAO Jinjing ^1

1. Department of Geriatrics, General Hospital of Southern Theater Command, Guangzhou 510010, China

2. Guangzhou University of Chinese Medicine, Guangzhou 510010, China

3. Community Health Service Center, Tonghui Street, Qijiang District, Chongqing 401420, China

*Corresponding author: ZHANG Yijun, Chief physician; E-mail: yijunzhang.2008@163.com

【Abstract】**Background** Sensory processing sensitivity (SPS) has attracted growing attention in clinical and psychological research. In China, the Chinese version of the 27-item Highly Sensitive Person Scale (HSPS) has been used mainly in studies of college students and urban residents. Its applicability to working-age adults, particularly frontline workers with diverse educational and regional backgrounds, has not been adequately examined. **Objective** This study sought to revise the Chinese HSPS-27 by addressing potential measurement bias related to educational attainment and place of birth and to develop a scale suitable for assessing SPS in the Chinese working-age population. **Methods** Several items were revised to better match

Citation: Zhang Yijun, Peng Dingjiang, Yang Mei, et al. Cultural adaptation and validity testing of the Highly Sensitive Person Scale among the working population in China [J]. Chinese General Practice, 2026. DOI: 10.12114/j.issn.1007-9572.2026.0095. [Epub ahead of print]. [www.chinagp.net]

Zhang Y J, Peng D J, Yang M, et al. Cultural adaptation and validity testing of the Highly Sensitive Person Scale among the working population in China[J]. Chinese General Practice, 2026. [Epub ahead of print].

the language habits and cognitive characteristics of the target population. A preliminary version of the scale was then developed after a pilot survey among participants with relatively low educational attainment. Participants were recruited through stratified and purposive sampling from five labor-intensive enterprises in Guangdong, Chongqing, and Hunan, with an emphasis on frontline employees. In total, 2 015 valid questionnaires were collected, yielding an effective response rate of 80.6%. The sample was randomly divided into two subsamples for exploratory factor analysis (EFA, n=1 007) and confirmatory factor analysis (CFA, n=1 008). After six months, 267 participants completed the retest. Python 3.13 and R were used to examine reliability, validity, measurement invariance, group differences, and regression models involving demographic variables. Inverse probability weighting (IPW) was also performed to assess whether the unequal gender distribution influenced the findings. **Results** The final scale consisted of 20 items and was named the HSPS-C20. A three-dimensional bifactor model showed good fit to the data, with a comparative fit index (*CFI*) of 0.974, a root mean square error of approximation (*RMSEA*) of 0.038, and a standardized root mean square residual (*SRMR*) of 0.028. The total scale showed high internal consistency, with a Cronbach' s α of 0.923. The values of Cronbach' s α for the three dimensions ranged from 0.802 to 0.885. The intraclass correlation coefficient (*ICC*) for the total scale was 0.707, while *ICC* for the dimensions ranged from 0.442 to 0.620. Bifactor indices further supported the use of the total score: the explained common variance (ECV) was 0.742, the percentage of uncontaminated correlations (PUC) was 0.679, and the omega coefficient was 0.924. Strict measurement invariance was supported across gender, educational attainment, and place of birth, with all ΔCFI values below 0.01. Gender and place of birth were significant predictors of SPS scores (β=0.119-0.172, P<0.001), although the effect associated with place of birth was small (β=0.061-0.066, P<0.01). The IPW sensitivity analysis suggested that the gender imbalance did not materially affect the main results. **Conclusion** The HSPS-C20 showed satisfactory reliability, validity, and measurement invariance in a large sample of Chinese working-age adults. The revised scale may serve as a practical and culturally appropriate instrument for assessing sensory processing sensitivity in this population.

【Key words】 Sensory processing sensitivity; Highly Sensitive Person Scale; Cultural adaptation; Working-age population; Measurement invariance; Reliability and validity

Sensory processing sensitivity (SPS) refers to an individual' s more acute perception of environmental stimuli and stronger emotional responses, and is an innate biological trait. Since Aron and his wife proposed this concept in 1997 [1], SPS has gradually become a research focus in psychology; it is estimated to occur in approximately 31% of the general population. Recent clinical observations have shown that SPS is closely associated with susceptibility to functional gastrointestinal disorders: highly sensitive individuals are prone to somatic discomfort such as abdominal pain and bloating, and often experience negative emotions such as anxiety or depression [2-3]. In view of this, the introduction and development of reliable SPS measurement tools is particularly important; however, there is still a lack in China of assessment methods for the general population that have undergone rigorous validation [4].

Previous researchers have introduced the Highly Sensitive Person Scale (HSPS) into China and completed its preliminary sinicization, verifying it as a multidimensional bifactor model [5]. However, prior study populations have focused on university students, urban residents, and online respondents. Yet online questionnaire completion is highly arbitrary, quality control is difficult to ensure, and retesting is also extremely difficult. In addition, existing Chinese versions of the scale have problems of insufficient cross-cultural adaptation and fail to consider China' s urban-rural population structure and educational conditions. Studies have shown that among China' s working-age population and non-urban employed groups, the proportions of those with an education below junior college reach 75.1% and 78.1%, respectively [6-7]. The original scale retains rigid literal translations such as "drowning," as well as contextual expressions with specific lifestyles such as "drinking coffee" and "art experience." Such differences in cultural habits and educational attainment can easily lead to deviations in understanding the scale, thereby affecting the reliability and validity of the scale in the general population.

Therefore, with authorization from the original author, the present study revised the scale for China' s working-age population. The study placed particular emphasis on the two variables of educational attainment and urban-rural differences. A pilot survey was first conducted among groups with an education below junior college, followed by large-scale systematic psychometric validation in the target population. The aim was to develop an SPS assessment tool for this population that conforms to local cultural characteristics.

1 Objects and Methods

1.1 Survey Participants

A stratified purposive sampling method was used, with place of birth (urban, non-urban) and educational attainment (below junior college, junior college and above) as stratification variables. "Place of birth" refers to the participant' s main place of residence before the age of 18 (based on household registration location), with the aim of reflecting the cultural background of the participant' s upbringing (urban or non-urban). The stratification targets were based on data from the National Bureau of Statistics' *China Labour Statistical Yearbook 2023* and the Ministry of Public Security' s 2023 household-registration population statistics: among the national employed population, those with an education below junior college account for approximately 77%, and those with non-urban household registration account for approximately 53% [7]. Accordingly, the present study preset the sampling targets for each stratum as approximately 77% of the sample having low educational attainment (below junior college) and approximately 53% having a non-urban place of birth, to ensure that the sample could represent China's working-age labor population in terms of cultural background and language comprehension ability for scale cultural adaptation and validation. To achieve the stratification targets, paper questionnaires were deliberately distributed to frontline employees in five labor-intensive enterprises (covering manufacturing, construction, and logistics) in Guangdong, Chongqing, and Hunan. This study was reviewed and approved by the Ethics Committee of the General Hospital of the Southern Theater Command (NO.ZLLK2023051), and all participants were informed of the purpose of the study and signed a paper informed consent form.

Inclusion criteria: (1) age 18 to <60 years, with the ability to complete the questionnaire independently; (2) questionnaire completion time of 8–20 min.

Exclusion criteria: (1) incomplete questionnaire or missing key information; (2) completion time <8 min or >20 min, with suspected perfunctory responding; (3) questionnaire not completed independently (questionnaires whose content was confirmed by two reviewers to be highly similar); (4) diagnosed mental or psychological illness, and current or previous history of taking psychotropic medications.

1.2 Research Instruments and Methods

1.2.1 Cross-cultural Adaptation and Preliminary Investigation

This study used the original HSPS-27 scale developed by Aron and Aron in 1997[^1], comprising 27 items. The Cronbach' s α of the original total scale was 0.85–0.87. The initial theoretical assumption for the scale was a unidimensional trait. Smolewska et al.[^23] first identified three independent factors: ease of excitation (EOE), low sensory threshold (LST), and aesthetic sensitivity (AES). Lionetti et al.[^30] were the first to systematically validate a three-dimensional bifactor nested model; this model simultaneously includes a general sensitivity factor affecting all items and the above three specific dimensions. After localization validation in multiple countries, it has become the mainstream fitting model in this field and has been recognized by the original author team, providing the theoretical basis for the cultural adaptation and model validation in the present study. Because authorization for the Chinese translation of the scale used in this study had been obtained in advance from the original author Aron by email, the translation process followed Brislin' s classical back-translation procedure[^8]. After the first draft was completed, five experts in psychology, linguistics, and clinical medicine reviewed it round by round, examining item by item the degree to which each item, on the basis of conforming to the original scale, fit the cultural background of China' s age-appropriate labor population, and adapting words that could easily cause ambiguity. The final HSPS-27 Chinese version was obtained.

Thirty age-appropriate workers who had not grown up in towns and had no university education experience were selected for a face-to-face interview. They were asked item by item whether the items were "understandable, answerable, and ambiguous." The interview results showed that the respondents understood all items 100%, could answer them, and found no ambiguity. This indicated that the linguistic expression of the scale was feasible in the target population.

1.2.2 Data Collection and Processing

Data collection was conducted in two stages. Stage 1: from March to August 2023, a total of 2,500 paper questionnaires were distributed, and 2,015 valid questionnaires were recovered, with an effective recovery rate of 80.6%. Stage 2: in February 2024, six months later, 300 people were randomly selected from the qualified sample at that time for retesting, and 267 valid retest questionnaires were obtained, with an effective recovery rate of 89.0%.

Data processing was conducted with reference to the standard procedure of Zhao et al.[^9]. When questionnaires were collected on site, items with omissions were checked manually, and missing items were reminded and completed on the spot. Questionnaires with a completion time shorter than 8 min or longer than 20 min were uniformly excluded. If more than 20% of the content was unanswered, the questionnaire was directly discarded; for individual missing items, the overall mean of that item was used for imputation, thereby ensuring the completeness of the sample[^10].

1.2.3 Data Analysis Procedure

The total sample was divided into Sample 1 and Sample 2 using a random equal-allocation method. Baseline characteristics were first described, and then item analysis was conducted. The critical ratio method and item-total correlation analysis were used sequentially, and the corrected item-total correlation index (corrected item-total correlation, CITC) was calculated. Items were screened using P<0.05 and both the item-total correlation coefficient and CITC >0.3 as criteria[^11].

Reliability and validity verification: Bartlett' s test of sphericity was performed and the KMO value was calculated using the data matrix of Sample 1. When KMO $\geq$ 0.7 and Bartlett' s test of sphericity was significant at P<0.05, the data were considered suitable for factor analysis[^12]. Parallel analysis was used to determine common factors[^13], and exploratory factor analysis (EFA) was performed using maximum likelihood estimation and Promax oblique rotation. Redundant items were removed according to factor loadings and communality (h^2), and an intermediate version was formed. Six experts (including two each of psychology professors, scale-development specialists, and clinicians) were further invited to conduct content validity testing of the intermediate version. Using Lynn' s item content validity index (I-CVI) ≥ 0.83[^14] as the criterion, the final items were determined. Using Sample 2, five competing models were fitted (M1 one factor; M2 two factors; M3 three factors oblique; M4 three factors orthogonal; M5 three-dimensional bifactor), and comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), and maximum likelihood robust (MLR) estimation were used for model parameter estimation.

Using the Sample 2 data, average variance extracted (AVE) and composite reliability (CR) were used to test convergent validity. The square root of AVE was compared with the correlations between dimensions to determine discriminant validity; meanwhile, the heterotrait-monotrait ratio (HTMT) index was referenced[^15-16] to synthesize the evidence. Because self-report data are prone to common method bias (CMB), after the five models described above were compared and the optimal model was determined, a global method factor was introduced into that model to construct the corresponding nested model, and systematic error was detected by comparing differences in fit indices[^17]. Reliability was verified by evaluating internal consistency with Cronbach' s α coefficient, and the long-term stability of the scale was tested based on the two test-retest datasets[^18].

The stability of the scale across groups was tested by grouping according to sex, educational level, and place of birth, and measurement invariance testing was completed (with ΔCFI ≤ 0.01 and ΔRMSEA ≤ 0.015 as criteria)[^19]. If measurement invariance was established, demographic variables such as sex, educational level, and place of birth were directly used for one-way analysis of variance or t tests, and effect sizes (η^2 or Hedges' g)[^20] were calculated to compare differences between groups. Finally, regression analysis was performed to determine the predictive effects of each variable on the HSPS total score and each dimensional score.

1.3 Statistical Methods

Python and R were used for statistical analysis. Normally distributed quantitative data were expressed as $\bar{x} \pm s$, and independent-samples t tests were used for comparisons between two groups; quantitative data not conforming to a normal distribution were expressed as M (P_{25}, P_{75}), and the Mann-Whitney U test was used. Categorical data were expressed as percentages (%), and intergroup comparisons were performed using the χ^2 test. The intraclass correlation coefficient (ICC) was used to evaluate test-retest reliability of the scale[^18]. $P<0.05$ was considered statistically significant. Considering the large sample size in this study, in addition to parametric tests, the corresponding nonparametric tests were also reported.

the validation results for robustness verification. The effect size for between-group comparisons was uniformly expressed using Hedges' g.

2 Results

2.1 Descriptive analysis

2.1.1 Baseline characteristics Through pretesting, a culturally adapted scale was obtained for testing. A total of 2 015 valid samples were collected and randomly divided into Sample 1, assigned to the exploratory factor analysis (EFA) group ($n = 1\ 007$), and Sample 2, assigned to the confirmatory factor analysis (CFA) group ($n = 1\ 008$). The survey participants were predominantly male (90.5%), aged 18–22 years (53.6%), had an education level of junior college or below (75.3%), and were mainly from non-urban areas (71.1%). Among them, female participants accounted for only 9.5%. Because of the imbalance in the sex ratio, inverse probability weighting (IPW) was ultimately used for sensitivity analysis. Comparisons of sex, age, education level, and place of origin between the two groups showed no statistically significant differences ($P > 0.05$), as shown in Table 1.

Table 1 Baseline characteristics of participants

Variable	EFA group ($n = 1\ 007$)	CFA group ($n = 1\ 008$)	χ^2 (Z) value	P value
Sex			0.004	0.949
Male	912 (90.6)	912 (90.5)		
Female	95 (9.4)	96 (9.5)		
Age			0.600	0.548
18–22 years	537 (53.3)	543 (53.9)		
23–27 years	283 (28.1)	276 (27.4)		
28–32 years	99 (9.8)	90 (8.9)		
33–37 years	59 (5.9)	69 (6.9)		
( )38 years	29 (2.9)	30 (3.0)		
Education level			0.555	0.456
Junior college or below	765 (76.0)	752 (74.6)		
Junior college or above	242 (24.0)	256 (25.4)		

Variable	EFA group ($n = 1\ 007$)	CFA group ($n = 1\ 008$)	χ^2 (Z) value	P value
Place of origin			1.577	0.209
Non-urban	703 (69.8)	729 (72.3)		
Urban	304 (30.2)	279 (27.7)		

Note: EFA = exploratory factor analysis; CFA = confirmatory factor analysis; indicates Z value.

2.1.2 Item analysis First, a normality test was performed on the full sample (n = 2 015). The mean values of the items were 2.47–3.71, and the standard deviations were 1.44–1.75. The Shapiro–Wilk normality test results showed that all items had skewed distributions. Because the absolute values of skewness for all items were < 0.84 and the absolute values of kurtosis were < 1.18, meeting the criteria of $|\text{skewness}| < 2$ and $|\text{kurtosis}| < 7$, the data for all items were considered to approximately follow a normal distribution[^21], and subsequent analyses could be carried out. The CITC of the items was 0.352–0.700, all higher than the retention criterion of 0.30[^22]. The extreme-group method was used to compare the extreme groups (the upper 27% and lower 27%) by independent-samples t test. The results showed that the between-group differences for all items were statistically significant ($P < 0.001$), indicating that each item had good discrimination, as shown in Table 2.

2.2 Structural validity

2.2.1 Test of suitability for factor analysis Using Sample 1 (n = 1 007), the calculated KMO value was 0.962, and Bartlett' s test of sphericity was significant (χ^2 = 11 064.82, $P < 0.001$). Therefore, the data were suitable for further factor analysis, as shown in Table 2.

Table 2 Summary of data quality tests

Test item	Test result	Judgment criterion	Conclusion
Valid sample size	2 015	>500	Sufficient
Normality test			
Skewness range	0.08–0.84	<2	Approximately normal
Kurtosis range	0.04–1.18	<7	Approximately normal
Item discrimination			

Figure 1 Scree plot of HSPS-27 based on parallel analysis

Figure 1: Figure 1 Scree plot of HSPS-27 based on parallel analysis

Test item	Test result	Judgment criterion	Conclusion
Critical ratio method	$P < 0.001$	Significant	Good
CITC range	0.352–0.700	>0.30	Acceptable
Prerequisites for factor analysis			
KMO	0.962	>0.90	Highly suitable
Bartlett' s test of sphericity	$\chi^2 = 11\ 064.82$, $df = 351$, $P < 0.001$	Significant	Suitable
Overall internal consistency of the scale	Cronbach' s $\alpha = 0.923$	>0.70	Excellent

Note: CITC = corrected item-total correlation; KMO and Bartlett tests were based on Sample 1 ($n = 1\ 007$).

2.2.2 Parallel analysis The actual eigenvalues of the first 3 factors (11.648, 1.821, 1.409) all exceeded the 95% eigenvalues of the random model (Figure 1), while the 4th factor (0.987) was lower than the random critical value (1.230). Combined with the elbow shape of the scree plot, this was judged to be a three-factor structure. Referring to the theoretical framework of the original authors[^1,^23], the 3 factors were still named EOE, LST, and AES.

Note: EFA = exploratory factor analysis; EOE = ease of excitation; LST = low sensory threshold; AES = aesthetic sensitivity.

Figure 1 Scree plot of HSPS-27 based on parallel analysis

2.2.3 Exploratory Factor Analysis

Promax oblique rotation was used to conduct EFA on sample 1. With a factor loading <0.40 or excessively high cross-loading as the deletion criterion, items (8, 9, 11, 18, and 13) were deleted in sequence (Figure 2). After review of content validity, the item content validity index (item content validity index, *I-CVI*) values of items 5 and 12 were both lower than 0.83, and their agreement was also relatively low, so they were deleted [24]. Among them, the agreement for item 15 was 0.140, lower than 0.20, but considering that its *I-CVI* met the standard, if this item were deleted, only 4 items would remain under the

AES factor; therefore, it was retained based on content coverage. Finally, 20 items were retained. The *I-CVI* values of the retained items were all 0.833-1.000, and the overall scale average content validity index (Scale content validity index/Average, *S-CVI/Ave*) was 0.975, forming the final Chinese version HSPS-C20. The second round of EFA still supported a three-factor structure, with a cumulative explained variance of 59.20%. However, the loadings of some items (such as item14, 19, 20, 21, etc.) were lower than 0.40, with loadings between 0.35 and 0.40, slightly below the preset threshold. Nevertheless, their loadings were still mainly concentrated on the corresponding AES factor, and no large-scale cross-factor drift occurred, indicating that the three-factor structure was stable and also suggesting room for future optimization of item wording (Figure 3).

*Note: * indicates items deleted from EFA.*

Figure 2 Heatmap of factor loadings of HSPS-27 based on exploratory factor analysis (EFA)

2.2.4 Confirmatory Factor Analysis

Using sample 2 (n=1 008), five competing models were fitted and compared; the three-dimensional bifactor model M5 was optimal [25]. *CFI*=0.974, *TLI*=0.967, *RMSEA*=0.038, *SRMR*=0.028. According to Akaike information criterion/Bayesian information criterion (AIC/BIC), the AIC difference Δ AIC was 0 (Table 3). The standardized factor loadings of the M5 model were as follows: general factor G loadings were 0.45-0.76 (P<0.001); among the special factor loadings on EOE, only item 26 was relatively high (0.614), and the rest were below 0.23; LST

Figure 3 Factor loading matrix of HSPS–C20 after the second round of EFA

loadings 0.36–0.50; AES loadings 0.32–0.57. Bifactor dimensionality indices: the explained common variance (ECV) was 0.742, the percentage of uncontaminated correlations (PUC) was 0.679, and the ω hierarchical coefficient was 0.924. Because ECV > 0.70 indicates that the total score of HSPS-C20 can be used as a unidimensional index, and PUC close to 0.70 suggests that multidimensionality has little bias in the estimation of the G factor, retaining the three specific factors was supported; see Figure 4.

Table 3 Comparison of confirmatory factor analysis (CFA) model fits

Model	*CFI*	*TLI*	*RMSEA*	*SRMR*	AIC	BIC	ΔAIC
M1 Single factor	0.867	0.850	0.080	0.060	68634.3	68831	789.8

Model	*CFI*	*TLI*	*RMSEA*	*SRMR*	AIC	BIC	ΔAIC
M2 Two factors	0.901	0.888	0.070	0.053	68332.4	68534	487.9
M3 Three oblique factors	0.962	0.956	0.044	0.035	67904.4	68115.8	59.9
M4 Three orthogonal factors	0.819	0.798	0.094	0.285	68957.1	69153.7	1112.6
M5 Bifactor with three dimensions	0.974	0.967	0.038	0.028	67844.5	68139.5	0

Note: *CFI* = comparative fit index; *TLI* = Tucker-Lewis index; *RMSEA* = root mean square error of approximation; *SRMR* = standardized root mean square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; ΔAIC = difference in Akaike information criterion. $CFI/TLI > 0.90$, $RMSEA < 0.08$, and $SRMR < 0.08$ were considered criteria for good fit.

2.2.5 Convergent Validity and Discriminant Validity Analysis

Validation was conducted with Sample 2 (n = 1,008). The *AVE* values for the dimensions of HSPS-C20 were 0.449–0.466, not reaching the traditional criterion of 0.5; however, the *CR* values were 0.802–0.887, all exceeding 0.70, indicating good composite reliability. In the *HTMT* test, the ratios among the dimensions of HSPS-C20 (EOE-LST = 0.796, EOE-AES = 0.742, LST-AES = 0.660) were all below the threshold criterion of 0.85. The results indicated that the three dimensions had sufficient discriminant validity [26]; see Table 4.

Figure 4. Optimal Fitting Model of HSPS-C20 Scale

Figure 2: Figure 4. Optimal Fitting Model of HSPS-C20 Scale

2.2.6 Measurement Invariance Test

Using the main sample (n = 2,015), multigroup CFA was conducted by sex, education, age, and birthplace. Configural invariance, weak (loading) invariance, strong (intercept) invariance, and strict (residual) invariance were tested sequentially. Each

Figure 4 Optimal Fitting Model of HSPS-C20 Scale

Table 4 Internal consistency reliability, convergent and discriminant validity

Dimension	Cronbach's α	McDonald ω	AVE	CR	Discriminant validity
EOE	0.885	0.887	0.466	0.887	EOE-AES=0.742
AES	0.802	0.802	0.449	0.802	EOE-LST=0.796
LST	0.836	0.837	0.462	0.837	AES-LST=0.660
HSPS-C20	0.923	0.925			

Note: EOE = ease of excitation; AES = aesthetic sensitivity; LST = low sensory threshold; AVE = average variance extracted; CR = composite reliability; HTMT = heterotrait-monotrait ratio.

The CFI values of all groups were above 0.950, and the $RMSEA$ values were all below 0.060. After stepwise constraints, the decreases in ΔCFI were all less than 0.01, and changes in $\Delta RMSEA$ did not exceed 0.015, indicating that the HSPS-C20 achieved strict measurement invariance across sex, educational level, age, and place of birth groups. Therefore, cross-group comparisons were sufficiently reasonable, as shown in Table 5.

2.2.7 Common method bias analysis

After adding one global method factor to the three-dimensional two-factor baseline model, the changes in model fit indices were small, with $\Delta CFI = 0.008$ and $\Delta RMSEA = 0.005$. The results indicated that CMB was very low and could be ignored, and did not interfere with the main results of this study[^27].

2.3 Reliability analysis

2.3.1 Internal consistency test

The Cronbach' s α coefficients of each HSPS-C20 dimension were 0.802-0.885, and the Cronbach' s α coefficient of the HSPS-C20 total scale was 0.923; all exceeded 0.70, indicating that the scale had good internal consistency, as shown in Table 4.

2.3.2 Test-retest reliability analysis

The 6-month retest results showed that the ICC of the HSPS-C20 total scale was 0.707 (95%CI=0.64-0.76) . According to the judgment criteria of Koo et al.[^28], the retest reliability of the total scale reached a moderate level: EOE=0.620 (moderate) , LST=0.529 (low) , and AES=0.442 (low) , as shown in Table 6.

2.4 Difference analysis and correlation analysis of demographic characteristics in HSPS-C20 total scores and dimensional scores

2.4.1 Age

According to the descriptive statistical results by age group (Table 1), the sample sizes of the age groups differed considerably and did not satisfy homogeneity of variance; therefore, the Kruskal-Wallis H test was used to compare differences among different age groups in HSPS-C20 total scores and dimensional scores. The results showed that differences among age groups in AES dimensional scores were statistically significant ($H = 12.445$, $P = 0.014$), whereas intergroup differences in the HSPS-C20 total score ($H = 9.296$, $P = 0.054$), EOE dimension ($H = 8.418$, $P = 0.077$), and LST dimension ($H = 8.443$, $P = 0.077$) were not significant. Thus, age had a relatively limited effect on HSPS-C20 scores.

2.4.2 Sex

Welch' s t test was used to analyze the effect of sex differences on SPS. The results showed that women scored higher than men on the HSPS-C20 total score and on all dimensional scores ($P < 0.001$), with the overall effect size reaching a moderate level (Table 7). Combined with the measurement invariance results, the higher SPS scores among women can be regarded as genuine group differences.

2.4.3 Educational level

Welch' s t test was used to analyze differences in HSPS-C20 total scores and dimensional scores across educational levels. The results showed that differences in all dimensions and in the total score were not significant ($P > 0.05$) (Table 7). Combined with the measurement invariance results, no significant differences were found among groups with different educational levels in HSPS-C20 total scores or dimensional scores, suggesting that educational—

Table 5 Results of Measurement Invariance Tests

Group	Model	CFI	$RMSEA$	ΔCFI	$\Delta RMSEA$	Judgment
Sex	Configural invariance	0.947	0.039			Baseline
Sex	Metric invariance	0.947	0.038	−0.000	−0.001	Passed
Sex	Scalar invariance	0.941	0.039	−0.005	+0.001	Passed
Sex	Strict invariance	0.938	0.039	−0.003	−0.000	Passed
Educational level	Configural invariance	0.950	0.039			Baseline
Educational level	Metric invariance	0.950	0.038	−0.000	−0.001	Passed
Educational level	Scalar invariance	0.948	0.037	−0.001	−0.001	Passed
Educational level	Strict invariance	0.946	0.037	−0.002	−0.000	Passed
Place of birth	Configural invariance	0.951	0.038			Baseline
Place of birth	Metric invariance	0.950	0.038	−0.001	−0.001	Passed
Place of birth	Scalar invariance	0.949	0.037	−0.001	−0.001	Passed
Place of birth	Strict invariance	0.947	0.036	−0.001	−0.000	Passed
Age	Configural invariance	0.949	0.039			Baseline

Group	Model	CFI	$RMSEA$	ΔCFI	$\Delta RMSEA$	Judgment
Age	Metric invariance	0.947	0.039	−0.002	−0.000	Passed
Age	Scalar invariance	0.944	0.039	−0.002	−0.000	Passed
Age	Strict invariance	0.939	0.039	−0.006	+0.001	Passed

Note: $\Delta CFI \leq 0.010$ and $\Delta RMSEA \leq 0.015$ were used as the criteria for invariance testing.

overall level of education.

2.4.4 Place of Birth

Welch' s t test was used to compare differences in the HSPS-C20 total score and dimensional scores by place of birth. The results showed that differences in all dimensions and in the total score were significant, and the scores of the urban group were higher than those of the non-urban group: total score $t = -3.543$, $P < 0.001$, $g = 0.173$; EOE $t = -3.380$, $P < 0.001$, $g = 0.168$; LST $t = -2.293$, $P = 0.022$, $g = 0.111$; AES $t = -3.444$, $P < 0.001$, $g = 0.166$ (Table 7). However, because of the relatively large difference in sample size, the effect sizes reported used the corrected Hedges' g, and the effect sizes for all dimensions were small ($g = 0.111$–0.173).

2.4.5 Multiple Linear Regression Analysis

Using demographic characteristics as independent variables, standardized multiple linear regression analyses were conducted separately with the HSPS-C20 total score and the scores of the three dimensions as dependent variables. The results showed that sex had the greatest impact on the HSPS-C20 total score and each dimensional score ($\beta = 0.119$–0.172, $P < 0.001$), indicating that among demographic variables, sex was the strongest independent predictor; place of birth significantly predicted the HSPS-C20 total score and EOE and AES scores ($P = 0.003$, 0.003, 0.006), but its predictive effect on the LST score was not statistically significant ($P = 0.056$). Educational attainment predicted only the AES score ($\beta = 0.052$, $P = 0.028$), and its predictive effects on the HSPS-C20 total score, EOE and LST scores were not statistically significant. Age did not significantly predict the HSPS-C20 total score or any dimensional scores ($P >$ 0.05). Moreover, R^2 was 0.019–0.036, indicating that demographic characteristics had a relatively low overall explanatory power for the HSPS-C20 total score and dimensional scores (Table 8).

2.5 IPW Sensitivity Analysis

Women accounted for only 9.5% of the sample ($n = 191$), far lower than the proportion of women in the national age-appropriate population, 48.8% (National Bureau of Statistics 2022 data: men 51.2%, women 48.8%). To test whether this sex composition bias affected the conclusions, the sample was calibrated using IPW based on the marginal probability of sex ($w_{\text{male}} = 0.566$, $w_{\text{female}} = 5.148$). The results showed that, before and after weighting, the directions and significance of the coefficients of the four independent variables in the total-score dimension were completely consistent (Table 9): the coefficient for sex decreased from 0.591 to 0.543, the coefficient for place of birth increased from 0.154 to 0.214, and age and educational attainment remained nonsignificant. At the dimensional level, fluctuations occurred mainly for age (3 groups), educational attainment (2 groups), and place of birth: in the EOE dimension, age and educational attainment changed from nonsignificant to significant; in the AES dimension, age and educational attainment changed from significant to nonsignificant; and in the LST dimension, age and place of birth changed from nonsignificant to significant. The effects of sex and place of birth on the total score and EOE remained significant and in the same direction before and after weighting, and were the most robust effects across all comparisons. After weighting, independent variables with small effect sizes, such as age and educational attainment, were more sensitive to the weight distribution; the judgment of significance at the dimensional level was easily affected by sampling fluctuations and should be interpreted together with the total-score analysis. However, sex composition bias did not alter the core conclusions of the study.

3 Discussion

In this study, the original HSPS-27 was culturally adapted for the target population to obtain the HSPS-C20. Its good

Table 6 Test-retest reliability stability test results ($n = 267$)

Dimension	*ICC*	95%*CI*	*F* value	*P* value
HSPS-C20	0.707	0.640–0.760	5.949	<0.001
EOE	0.620	0.540–0.690	4.275	<0.001
LST	0.529	0.410–0.630	3.501	<0.001
AES	0.442	0.340–0.530	2.581	<0.001

Note: ICC = intraclass correlation coefficient; according to the Cicchetti criterion, $ICC \geq 0.75$ is good, 0.60–0.74 is moderate, and <0.60 is low.

Table 7 Analysis of intergroup differences in the total score and scores across individual dimensions of HSPS-C20 ($n = 2\ 015$)

Variable	HSPS-C20 total score	EOE	LST	AES
Sex	$t = -7.785$, $P < 0.001$, $g = 0.617$	$t = -7.804$, $P < 0.001$, $g = 0.607$	$t = -5.668$, $P < 0.001$, $g = 0.438$	$t = -6.983$, $P < 0.001$, $g = 0.525$
Educational attainment	$t = 0.524$, $P = 0.600$, $g = 0.028$	$t = 1.756$, $P = 0.079$, $g = 0.092$	$t = -0.119$, $P = 0.905$, $g = 0.006$	$t = -1.080$, $P = 0.281$, $g = 0.057$
Place of birth	$t = -3.543$, $P < 0.001$, $g = 0.173$	$t = -3.380$, $P < 0.001$, $g = 0.168$	$t = -2.293$, $P = 0.022$, $g = 0.111$	$t = -3.444$, $P < 0.001$, $g = 0.166$

Note: Comparisons between two groups used Welch' s t test; effect sizes are expressed as Hedges' g.

Table 8 Results of multiple linear regression analysis on factors associated with total and dimensional scores of HSPS-C20 scale (n = 2 015)

Predictor variable	HSPS-C20 total score β	HSPS-C20 total score P value	EOE β	EOE P value	LST β	LST P value	AES β	AES P value
Age	0.005	0.819	0.014	0.558	0.024	0.327	−0.033	0.166
Sex	0.172	<0.001	0.164	<0.001	0.119	<0.001	0.161	<0.001
Educational attainment	0.005	0.846	−0.027	0.256	0.008	0.738	0.052	0.028
Place of birth	0.066	0.003	0.065	0.003	0.043	0.056	0.061	0.006
R^2	0.036		0.035		0.019		0.030	
Adjusted R^2	0.034		0.033		0.017		0.028	

The psychometric properties include internal consistency, discriminant validity, convergent validity, and test-retest stability; all reliability and validity indicators met psychometric standards. The common-method bias test results were satisfactory, and the obtained data were reliable. Therefore, HSPS-C20 is suitable for SPS assessment in age-appropriate working populations and can also serve as an assessment tool for clinical medicine and mental health research.

3.1 Localization Adaptation and Research Design Features for Age-Appropriate Working Populations

3.1.1 Stratified purposive sampling Previous studies have often pursued precise matching between samples and national demographic proportions. This study instead adopted purposive sampling, with priority given to including age-appropriate working populations with college education or below and non-urban backgrounds. Following the ideas of the ITC Guidelines for Test Adaptation, if a scale can achieve good reliability, validity, and measurement invariance among groups with relatively low education and non-urban populations, then it is reasonable to infer that it is likewise applicable to populations with higher education and more complex urban backgrounds[^29]. Therefore, the educational distribution of this study sample was matched, and the proportion of non-urban participants was relatively high; thus, the sample was also representative.

3.1.2 Adaptation testing scheme targeting grassroots groups To fully examine the scale' s fit among grassroots working populations, this study selected frontline employees from labor-intensive enterprises in multiple regions for the survey. In the final sample, non-urban household registration accounted for 71.0%, higher than the national proportion of 53%; education at college level or below accounted for 5.3%, roughly equivalent to the educational distribution of the national age-appropriate working population, 77%. The appropriate increase in the proportion of non-urban household-registered workers was a preset research design intended to test the scale' s applicability in a way that was closer to the target population, thereby ensuring that the tool could be stably applied to grassroots working populations.

3.1.3 Alignment between theoretical design and practical evidence Unlike previous scale Sinicization studies that mainly focused on direct translation and back-translation, this study took interference factors such as educational level and urban–rural differences into prior consideration: item wording was first optimized to address comprehension barriers among grassroots populations, and a face-to-face pretest was conducted among workers with low education and non-urban backgrounds ($n = 30$), achieving understandability and answerability for all items. The content validity index ($I\text{-}CVI \geq 0.833$) also confirmed the appropriateness of the item wording. The key evidence verifying the effectiveness of this strategy came from the measurement invariance test: both the urban–rural grouping and the educational-level grouping passed strict equivalence ($\Delta CFI < 0.01$), indicating that the measurement meaning of the scale items was not affected by place of origin or educational level.

3.1.4 Comparison with similar domestic studies Compared with existing Chinese adaptation studies of the HSPS, this study conducted a test–retest reliability examination and further completed multidimensional measurement invariance tests for gender, place of origin, and educational background. At present, similar validation studies at home and abroad are rarely reported. In addition, this study conducted a large-sample survey of frontline workers, face-to-face interviews, medium- and long-term retesting, and comparisons of multiple demographic characteristics, providing evidence for SPS assessment among grassroots working populations and making up for the limitation that existing scales are not adapted to grassroots workers.

3.2 Evaluation of Scale Reliability and Validity Results

3.2.1 Model structure The revised HSPS-C20 is consistent with the mainstream theoretical framework of SPS, retaining a three-dimensional bifactor structure and corresponding to the dimensional allocation of items in the original HSPS-27[^23,^30]. In this study, several items from the LST dimension of the original HSPS-27 (e.g., 16, 23, 25, 26, 27) were assigned to the EOE dimension in the Chinese worker population sample. This may be related to the ambiguity of item boundaries in parts of the original HSPS-27, and may also be related to cross-cultural adaptation. Under different—

Table 9 Comparison of regression coefficients before and after inverse probability weighting (IPW) ($n = 2,015$; weighted effective $n \approx 719$)

Dependent variable	Independent variable	OLS_B	OLS_P	IPW_B	IPW_P	Conclusion consistent
HSPS-C20 total score	Age	−0.002	0.609	0.007	0.066	Yes
HSPS-C20 total score	Gender	0.591	<0.001	0.543	<0.001	Yes
HSPS-C20 total score	Educational level	0.030	0.570	−0.080	0.193	Yes
HSPS-C20 total score	Place of origin	0.154	0.001	0.214	<0.001	Yes
EOE	Age	0.002	0.656	0.013	0.002	No
EOE	Gender	0.600	<0.001	0.543	<0.001	Yes
EOE	Educational level	−0.052	0.384	−0.164	0.015	No
EOE	Place of origin	0.179	0.001	0.275	<0.001	Yes

Dependent variable	Independent variable	OLS_B	OLS_P	IPW_B	IPW_P	Conclusion consistent
LST	Age	0.001	0.876	0.011	0.013	No
LST	Gender	0.510	<0.001	0.451	<0.001	Yes
LST	Educational level	0.055	0.365	−0.095	0.176	Yes
LST	Place of origin	0.104	0.056	0.182	0.001	No
AES	Age	−0.014	0.011	−0.009	0.078	No
AES	Gender	0.672	<0.001	0.651	<0.001	Yes
AES	Educational level	0.146	0.025	0.092	0.220	No
AES	Place of origin	0.168	0.004	0.144	0.012	Yes

Note: Using gender as the weighting variable, inverse probability weights were calculated according to the reciprocal of the ratio between the national population proportion and the sample proportion (IPW; $w_{\text{male}} = 0.566$, $w_{\text{female}} = 5.148$). Weighted least squares (WLS) and unweighted ordinary least squares (OLS) coefficients (B) and P values were compared. For "conclusion consistent," $P < 0.05$ was used as the threshold for judging significance; if the judgment results before and after weighting were the same, it was recorded as "Yes," and if different, as "No." Conclusions at the total HSP score level were all consistent; at the subdimension level, 6 of 16 comparisons (37.5%) reversed, concentrated in the two independent variables with relatively small effect sizes, age and educational level. Gender and place of origin had the most stable effects on the total score and EOE.

Under the same cultural context, dimension aggregation differences and attribution drift of individual factors are relatively common.

When conducting exploratory factor analysis, items were deleted according to the standard item-removal criteria. Although item 15 had a communality of 0.186 in the EFA, slightly below 0.20, comparison of the two bifactor CFA models with item 15 retained versus deleted showed that the model fit indices were almost identical ($\Delta CFI = 0.001$), and item 15 had significant loadings on both the general factor ($\lambda = 0.460$) and the AES-specific factor ($\lambda = 0.357$); therefore, it was retained.

In addition, when this study used sample 2 ($n = 1\ 008$) to verify convergent validity, the AVE values of the dimensions were 0.449–0.466, not reaching the 0.5 convergence threshold; however, the three-dimensional CR values were 0.802–0.887, all higher than 0.70. According to Henseler et al. [26] and Diamantopoulos, for self-report personality scales, composite reliability can be used to assist in judging convergent validity: when $CR > 0.8$ and $HTMT < 0.85$, AVE in the range of 0.40–0.50 can still be considered acceptable for convergent validity. In this study, the $HTMT$ results among the three dimensions were all below the critical value of 0.85, demonstrating sufficient discriminant validity among dimensions [26].

At the same time, the overall ECV of the three-dimensional bifactor model in this study was 0.742, indicating that the core characteristics of the scale as a whole were stable and could compensate for the deficiency of weak convergence at the subdimension level. The difficulty in achieving an AVE of 0.5 for the subdimensions was not caused by the present localization adaptation. For example, cross-cultural validations of the HSPS in Lithuania and Spain have reported similar weak convergent validity results, the source of which lies in the inherent structural characteristics of the original scale—high cross-loadings among the three-dimensional items and dominance of the total-score unidimensional trait—which represent intrinsic measurement limitations of the original scale structure [10,15] and require further improvement.

3.2.2 Retest reliability

Influenced by objective factors, the retest interval in this study was relatively long (about 6 months). Therefore, the total-scale $ICC = 0.707$ (moderate), $EOE = 0.620$ (moderate), $LST = 0.529$ (relatively low), and $AES = 0.442$ (relatively low). Among these, the retest reliability of the total scale reached a moderate level; for the dimensions, EOE was at a moderate level, whereas LST and AES were relatively low. Although SPS is a relatively stable personality trait, it is also affected to some extent by many factors such as emotion, stress, and experience. Moreover, the sample in this study came from several enterprises, the working environments were similar, and more than half of the sample consisted of newly hired employees who were in the period of occupational adaptation. The environmental and psychological changes experienced over six months may have been significantly greater than those of veteran employees, thereby introducing more genuine state fluctuations, which to some extent reduced the estimated ICC values.

3.3 Effects of demographic characteristics on total scale and dimension scores

Through invariance testing, differential analysis, and multiple linear regression analysis, this study found that the HSPS-C20 showed strict measurement invariance across groups by sex, educational level, place of birth, and age. Meanwhile, the R^2 values of the regression models were relatively low (0.019–0.036), suggesting that demographic variables could explain only about 1.9%–3.6% of the variance in HSPS-C20 total and dimension scores. This is similar to common findings in studies of comparable personality scales, indicating that demographic variables have limited explanatory power for SPS. However, in the regression analysis, place of birth had significant predictive effects on the total score and multiple dimensions, and educational level also had a significant predictive effect on AES. These findings also confirm the necessity of using place of birth and educational level as key variables in scale adaptation. If these effects are ignored, subtle differences between groups on the scale may be masked. Future studies may further include non-demographic variables such as personality, emotions, and environmental exposure to explore their relationships with SPS.

Regarding sex differences, women scored higher than men on all HSPS-C20 dimensions. The effect size Hedges' g was 0.438–0.617, overall reaching a moderate level; moreover, this result held under conditions of strict measurement invariance, suggesting that the difference is more likely to reflect true SPS differences between male and female groups rather than measurement bias, consistent with conclusions of similar studies in China and abroad [31]. Although the sex ratio in the sample of this study was severely imbalanced, IPW sensitivity analysis based on the marginal probability of sex showed that after correcting for sex imbalance, the sex effect remained significant, and the regression conclusions at the HSPS-C20 total-scale level did not change. This indicates that women's higher SPS levels show a relatively stable statistical difference, and that sex composition bias did not substantively affect the conclusions of this study.

The analysis by place of birth showed that the urban group scored slightly higher than the non-urban group on the HSPS-C20 total score and all dimension scores. Regression analysis further indicated that place of birth did not have a statistically significant predictive effect on LST scores ($P = 0.056$), but had significant predictive effects on the HSPS-C20 total score, EOE, and AES scores ($P = 0.003$, 0.003, 0.006), with relatively small overall effect sizes ($\beta = 0.061$–0.066; LST was marginally significant, $\beta = 0.043$). Because most items of the HSPS-27 mainly target responses and perceptions related to external noise, sensory load, and environmental discomfort, objective differences between urban and rural environments may be potential factors leading to slight differences in the scores of the two groups; however, the specific mechanisms remain unclear and require further verification in subsequent studies. Since the urban-rural grouping passed strict measurement invariance testing, the above score differences represent inherent group trait differences rather than measurement non-equivalence of the scale, and therefore do not affect the core conclusions regarding the overall structure of the scale and its cross-population application. It is worth noting that IPW sensitivity analysis showed that the place-of-birth effect remained stable after weighting according to the sex ratio, suggesting that its effect does not completely depend on differences in the sex composition of the sample. Considering that the proportion of urban participants in the female sample was relatively high, there may be some degree of structural interaction between sex and place of birth; however, even after correcting for the sex ratio, the place-of-birth effect remained significant. This result supports the relatively independent interpretive significance of place of birth for the HSPS-C20 total-scale score.

Because no significant differences were found across educational levels in the HSPS-C20 total score or dimension scores ($P > 0.05$), together with the measurement invariance results, differences in educational level did not substantially interfere with the scale test results. However, in the regression analysis, educational level had a significant positive predictive effect on AES ($\beta = 0.052$, $P = 0.025$), suggesting that slightly higher AES scores were associated with higher educational level, but the effect was very small and had limited explanatory power. At the HSPS-C20 total-scale level, age and educational level did not reach significance either before or after weighting in the IPW sensitivity analysis, suggesting that the independent explanatory effects of the two on the total-scale score were relatively limited. At the dimension level, however, the judgment of significance for the two was relatively sensitive to sample weighting.

Regarding age, the overall differences in HSPS-C20 total score and dimension scores among age groups were limited. Only the 18–22-year-old group and the 33–37-year-old group showed significant differences in some dimensions, and in regression analysis, the effect of age on the total-scale score was

the predictive effect did not reach statistical significance. Regarding the total HSPS-C20 score, in the IPW sensitivity analysis, neither age nor education level reached a significant level before or after weighting, suggesting that the independent explanatory effects of the two variables on the total scale score were relatively limited. At the dimensional level, however, the significance judgments for the two variables were comparatively sensitive to sample weighting. It is worth noting that the mean age of the female sample (28.14 ± 6.94) years was higher than that of the male sample (23.76 ± 5.31) years, suggesting that there may also be a certain association between sex and age. However, because age had a limited predictive effect at the total-score level and still did not reach significance after IPW weighting, this association is insufficient to alter the core conclusions of the study.

3.4 Comparison with Previous Studies

Compared with previous research findings, the present study is applicable to a broader population. For example, Zhang Yiman et al.

4

rigorously revised the full version of the HSPS-27 and proposed a six-factor structure, whereas Li Chongyi et al.

5

developed a 12-item ultra-brief scale and established a two-factor structure; moreover, the conclusion that women have higher SPS levels than men has remained highly consistent across scales in different samples and versions. However, the samples differed substantially among studies. Zhang Yiman et al. mainly recruited 846 participants through online channels, with the sample consisting primarily of students and urban online respondents, and with frontline workers almost absent. Li Chongyi recruited 867 online participants, among whom the proportion with a junior-college education or below was low, and the study did not cover a large number of workers who grew up outside urban settings.

3.5 Research Significance and Scope of Application

This study is the first to rigorously and systematically validate the three-dimensional bifactor structural model of the HSPS-C20 among the working-age population in China, thereby expanding the cross-cultural application scenarios and population boundaries of SPS theory. It confirms that the three-dimensional structure of the scale is also applicable among grassroots populations in non-urban settings and with educational attainment below junior college, and also verifies that, in the Chinese working-age population, women have significantly higher SPS levels than men. These findings provide new empirical evidence for the application of SPS theory in an Eastern cultural context. Although the age distribution of this sample skewed toward young adults (18–27 years accounting for 81.3%), the regression analysis showed that age had no significant predictive effect on the total SPS score ($\beta = 0.005,\ P = 0.819$), and the scale passed the measurement invariance test across different age groups ($\Delta CFI < 0.01$), indicating that age bias had a limited effect on the evaluation of scale performance. The scale showed stable psychometric properties across age groups from 18 to 58 years.

Therefore, the HSPS-C20 is mainly applicable to China's working-age population aged 18–60 years, including employees in various urban and rural enterprises, community workers, and grassroots laborers. It is especially suitable for psychological screening among populations from non-urban backgrounds and those with educational attainment below junior college.

3.6 Study Limitations and Future Directions

The limitations of this study are as follows: (1) Data on anxiety, depression, neuroticism, and other variables were not collected simultaneously; therefore, criterion-related validity testing is lacking, and the relationships between SPS and negative emotions as well as psychosomatic symptoms cannot be fully clarified. (2) The retest interval was relatively long, and the likelihood of experiencing life events or environmental changes during the period increased, affecting the evaluation of the temporal stability of the scale. (3) A balanced sex structure was not used in sampling. Although the IPW weighting method verified that sex bias had no substantive influence on the core conclusions of this study, it still limits the external generalizability of the conclusions.

Therefore, future studies should further use sex as a stratified sampling indicator and systematically and stratifiedly complete follow-up research; supplement clinical emotional indicators to improve criterion-related validity verification;

shorten the retest interval to enhance the temporal stability of the scale; combine other-rating scales and objective physiological indicators to compensate for the limitations of single self-report measures; conduct prospective cohort studies to verify the long-term predictive value of the scale; and further optimize the wording of HSPS-C20 items to reduce interference caused by differences in life background in item comprehension and to improve the scale' s cross-population universality.

4 Conclusion

This study is the first to complete the localization revision and psychometric validation of a high-sensitivity scale among grassroots populations in China, filling the gap in the application of existing scales, which has been limited to students, urban populations, or online groups. At the same time, through rigorous measurement invariance testing, it compared perceived sensitivity levels among people of different sex, urban/rural background, and educational attainment, thereby providing a standardized local tool for occupational mental health screening and follow-up. It also improves the empirical evidence on the trait of sensitivity to sensory processing among China' s working population. The revised Chinese Highly Sensitive Person Scale HSPS-C20 has a stable structure, good reliability and validity, and clearly described items. It is applicable to the working-age population in China and can lay a foundation for research on highly sensitive personality in the general population, serving as an assessment tool for clinical psychological evaluation, occupational health monitoring, mental health research, and general-practice medical research.

Author contributions: Zhang Yijun was responsible for the overall study design, proposed the research ideas and research plan, completed the statistical analysis and final revision of the manuscript, and is responsible for the paper; Peng Dingjiang was responsible for data collection, acquisition, cleaning, and part of the statistical analysis; Yang Mei was responsible for collection of part of the data for the paper; Liao Jingjing was responsible for reviewing materials, writing part of the paper, conducting part of the statistical analysis, and proofreading the paper.

The authors declare no conflicts of interest.

References

1

Aron E N, Aron A. Sensory-processing sensitivity and its relation to introversion and emotionality

J

. J Pers Soc Psychol, 1997, 73(2): 345–368. DOI: 10.1037/0022-3514.73.2.345.

2

Kisla Ekinci R M, Balci S, Mart O O, et al. Is Henoch-Schönlein Purpura a susceptibility factor for functional gastrointestinal disorders in children?

J

. Rheumatol Int, 2019, 39(2): 317–322. DOI: 10.1007/s00296-018-4129-7.

3

Midenfjord I, Grinsvall C, Koj P, et al. Central sensitization and severity of gastrointestinal symptoms in irritable bowel syndrome, chronic pain syndromes, and inflammatory bowel disease

J

. Neurogastroenterol Motil, 2021, 33(12): e14156. DOI: 10.1111/nmo.14156.

4

Zhang Yiman, Zhang Ya. Revision and reliability and validity testing of the Chinese version of the Highly Sensitive Person Scale

J

. Chinese Journal of Clinical Psychology, 2023, 31(4): 836–842. DOI: 10.16128/j.cnki.1005-3611.2023.04.014.

5

Li Chongyi. Reliability and validity testing of the Chinese brief version of the Highly Sensitive Person Scale

J

. Advances in Psychology, 2025, 15(4): 418–425. DOI: 10.12677/ap.2025.154222.

6

Fu Linghui, Liu Aihua. China Statistical Yearbook 2022

Z

. Beijing: China Statistics Press Co., Ltd., 2022.

[7] Wang Gui-rong. *China Rural Statistical Yearbook 2023* [Z]. Beijing: China Statistics Press Co., Ltd., 2023.

[8] Brislin R W. Translation and back-translation procedure[M]//Lonner W J, Berry J W. *Field methods in cross-cultural research.* Thousand Oaks: Sage Publications, 1986: 137–164.

[9] Zhao Y X, Summers R, Gathara D, et al. Conducting cross-cultural, multilingual or multi-country scale development and validation in health care research: a 10-step framework based on a scoping review[J]. *J Glob Health*, 2024, 14: 04151. DOI: 10.7189/jogh.14.04151.

[10] Li Jian-feng, Zhang Jing-yi, Li Li-kang, et al. Discussion of methods for handling missing values in clinical trials[J]. *Chinese Journal of Evidence-Based Medicine*, 2024, 15(5): 1165–1172. DOI: 10.12290/xhyxzz.2023-0415.

[11] Caycho-Rodríguez T, Vilca L W, Valencia P D, et al. Is the meaning of subjective well-being similar in Latin American countries? A cross-cultural measurement invariance study of the WHO-5 well-being index during the COVID-19 pandemic[J]. *BMC Psychol*, 2023, 11(1): 102. DOI: 10.1186/s40359-023-01149-8.

[12] Tsegaye B S, Andegiorgish A K, Amhare A F, et al. Construct validity and reliability Amharic version of perceived stress scale (PSS-10) among Defense University students[J]. *BMC Psychiatry*, 2022, 22(1): 691. DOI: 10.1186/s12888-022-04345-9.

[13] Goretzko D, Heumann C, Bühner M. Investigating parallel analysis in the context of missing data: a simulation study comparing six missing data methods[J]. *Educ Psychol Meas*, 2020, 80(4): 756–774. DOI: 10.1177/0013164419893413.

[14] Lynn M R. Determination and quantification of content validity[J]. *Nurs Res*, 1986, 35(6): 382???386. DOI: 10.1097/00006199-198611000-00017.

[15] Ngai S S, Cheung C K, Ng Y H, et al. Compassion Scale: factor structure and scale validation in Hong Kong adolescents[J]. *Front Psychol*, 2025, 16: 1508402. DOI: 10.3389/fpsyg.2025.1508402.

[16] Hernández Encuentra E, Robles N, Angulo-Brunet A, et al. Spanish and Catalan versions of the eHealth literacy questionnaire: translation, cross-cultural adaptation, and validation study[J]. *J Med Internet Res*, 2024, 26: e49227. DOI: 10.2196/49227.

[17] Kock F, Berbekova A, Assaf A G. Understanding and managing the threat of common method bias: Detection, prevention and control[J]. *Tour Manag*, 2021, 86: 104330. DOI: 10.1016/j.tourman.2021.104330.

[18] Lee J G J, Ng Q X, Luo N, et al. Validation of the consumer health activation index (CHAI) among community-dwelling adults in primary care clinics in Singapore[J]. *Popul Health Metr*, 2025, 23(1): 49. DOI: 10.1186/s12963-025-00402-z.

[19] Nunes C, Pérez-Padilla J, Martins C, et al. The brief COPE: measurement invariance and psychometric properties among community and at-risk Portuguese parents[J]. *Int J Environ Res Public Health*, 2021, 18(6): 2806. DOI: 10.3390/ijerph18062806.

[20] Graves B S, Hall M E, Dias-Karch C, et al. Gender differences in perceived stress and coping among college students[J]. *PLoS One*, 2021, 16(8): e0255634. DOI: 10.1371/journal.pone.0255634.

[21] Curran P J, West S G, Finch J F. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis[J]. *Psychol Meth*, 1996, 1(1): 16–29. DOI: 10.1037/1082-989x.1.1.16.

[22] Del Pilar Díaz Gamarra P, Rosario Quiroz F J, Alomía E R E, et al. Psychometric evidence of the WHO-5 well-being index in a sample of participants from hospitals and older adults care centers in Peru[J]. *Front Public Health*, 2025, 13: 1670429. DOI: 10.3389/fpubh.2025.1670429.

[23] Smolewska K A, McCabe S B, Woody E Z. A psychometric evaluation of the Highly Sensitive Person Scale: The components of sensory-processing sensitivity and their relation to the BIS/BAS and "Big Five" [J]. *Personality Individ Differ*, 2006, 40(6): 1269–1279. DOI: 10.1016/j.paid.2005.09.022.

[24] Zare H, Tagharrobi Z, Zare M. Cross-cultural adaptation and psychometric evaluation of the social frailty scale in Iranian older adults[J]. *BMC Geriatr*, 2024, 24(1): 368. DOI: 10.1186/s12877-024-04940-3.

[25] Hu L T, Bentler P M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives[J]. *Struct Equ Model A Multidiscip J*, 1999, 6(1): 1–55. DOI: 10.1080/10705519909540118.

[26] Henseler J, Ringle C M, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling[J]. *J Acad Mark Sci*, 2015, 43(1): 115–135. DOI: 10.1007/s11747-014-0403-8.

[27] Podsakoff P M, MacKenzie S B, Lee J Y, et al. Common method biases in behavioral research: a critical review of the literature and recommended remedies[J]. *J Appl Psychol*, 2003, 88(5): 879–903. DOI: 10.1037/0021-9010.88.5.879.

[28] Koo T K, Li M Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research[J]. *J Chiropr Med*, 2016, 15(2): 155–163. DOI: 10.1016/j.jcm.2016.02.012.

[29] International Test Commission. The ITC guidelines for translating and adapting tests (2nd ed.)[EB/OL]. [2026-01-10]. https://www.intestcom.org.

[30] Lionetti F, Aron A, Aron E N, et al. Dandelions, tulips and orchids: evidence for the existence of low-sensitive, medium-sensitive and high-sensitive individuals[J]. *Transl Psychiatry*, 2018, 8(1): 24. DOI: 10.1038/s41398-017-0090-6.

[31] Trå H V, Volden F, Watten R G. High Sensitivity: Factor structure of the highly sensitive person scale and personality traits in a high and low sensitivity group. Two gender-matched studies[J]. *Nord Psychol*, 2023, 75(4): 328–350. DOI: 10.1080/19012276.2022.2093778.

(Received: 2026-05-28; Revised: 2026-07-09)

(Editor of this article: Kang Yao-hui)

Note: Figure translations are in progress. See original paper for figures.