A psychometric evaluation of the Highly Sensitive Person Scale in ethnically and culturally heterogeneous South African samples



Andrew K. May 1 . Shane A. Norris 2,3 . Linda M. Richter 6 . Michael M. Pitman 4 .

Published online: 6 August 2020 © The Author(s) 2020

Abstract

Sensory Processing Sensitivity (SPS) is a personality trait that describes highly neurosensitive individuals who, for better and for worse, are permeable to their environmental context. Recently, these individuals have been noted for their amenability to positive intervention efforts - an observation that may have important psychosocial value. SPS is currently assessed through the 27-item Highly Sensitive Person Scale (HSPS). However, this instrument has not been adequately scrutinised in cross-cultural samples, and has generated varied factor solutions that lack consistent support. We assessed the HSPS in South African university student samples which were ethno-culturally diverse, across four academic years (n = 750). The HSPS demonstrated strong reliability across samples ($\alpha > .84$). Factor analysis revealed a novel five-factor solution (Negative Affect, Neural Sensitivity, Propensity to Overwhelm, Careful Processing and Aesthetic Sensitivity). As per previous reports, latent class analysis suggested a three class solution. We validated these findings in a general population sample that was part of the longitudinal Birth to Twenty Plus cohort (n = 1400). In conclusion, we found the HSPS to be reliable in culturally diverse samples. The instrument remains a robust tool for identifying sensitive individuals and may be an important addition to psychosocial studies in low-to-middle income countries.

Keywords Sensory Processing Sensitivity · Highly Sensitive Person Scale · Psychometric · Cross-cultural testing · Birth to Twenty Plus · South Africa

Sensory Processing Sensitivity (SPS) is a personality trait that describes concurrent heightened awareness and greater cognitive processing of sensory stimulation (Aron and Aron 1997). These features stem from a highly active and receptive central nervous system, confirmed through several fMRI studies

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12144-020-00988-7) contains supplementary material, which is available to authorized users.

- Andrew K. May andrewkmay@gmail.com
- Division of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand and the National Health Laboratory Service, Johannesburg, South Africa
- Developmental Pathways for Health Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
- ³ DSI-NRF Centre of Excellence in Human Development, University of the Witwatersrand, Johannesburg, South Africa
- Department of Psychology, School of Human and Community Development, University of the Witwatersrand, Johannesburg, South Africa

(Acevedo et al. 2017, 2018). Individuals in possession of the trait (termed highly sensitive persons - HSPs) are characterised by a number of behavioural responses, including: strong emotional reactions (positive and negative) which typically foster improved learning (Aron et al. 2005, 2012); a "pause-to-check" response in the presence of novel stimuli (Aron and Aron 1997); and a propensity to become overwhelmed by excessive stimulation and/or cognitive processing, leading to frequent withdrawal into solitude to afford regeneration of an overtaxed nervous system (Aron and Aron 1997; Aron et al. 2012). Thus HSPs, who constitute 20—35% of the general population (Greven et al. 2019), adopt a strategy of quiet vigilance rather than active exploration, paying careful attention to environmental detail in order to guide their behavioural responses (Aron and Aron 1997; Aron et al. 2012).

Recently, SPS was merged into a larger theoretical metaframework known as Environmental Sensitivity (ES; Greven et al. 2019; Pluess 2015). The central claim of ES theorising is that humans can be classified according to the extent to which they register and process environmental stimulation (i.e. their "neurosensitivity"). At one end of the spectrum exist 'dandelion' individuals who are characterised by genotypes, endophenotypes and phenotypes that render them



relatively unresponsive to environmental fluctuations. These individuals often exhibit stable characteristics (e.g. securely attached; Belsky and Pluess 2013) regardless of the quality of the developmental environment in which they mature. At the opposite end of the spectrum are 'orchid' individuals who, for better and for worse, are highly responsive to their social milieu, thriving or deteriorating in direct proportion to the calibre of their surrounds. Between these extremes lie 'tulip' individuals, who are moderately sensitive to their surrounds (Lionetti et al. 2018). In addition to a burgeoning literature on humans (Acevedo et al. 2018; Pluess et al. 2018), environmental sensitivity is supported by research in over 100 animal species (Aron et al. 2012).

Individual differences in sensitivity to environmental influence carry important practical implications (Belsky and Pluess 2013; Belsky and van IJzendoorn 2015). For example, orchid individuals may be especially receptive targets for intervention efforts. Several recent studies have demonstrated that orchid individuals respond far *better* to positive intervention than their dandelion counterparts (Morgan et al. 2017; Nocentini et al. 2018; Pluess and Boniwell 2015), but appear to fare far *worse* in the absence of intervention. Ethical issues aside (Belsky and van IJzendoorn 2015), this implication alone motivates for continued research into neurosensitive individuals and the mechanisms and markers that underpin their sensitivity.

Although several possible markers have already been identified (e.g. difficult temperament - Ramchandani et al. 2010; impulsivity - Slagt et al. 2015; and a range of genetic markers -Belsky et al. 2009), SPS, as operationalised through the Highly Sensitive Person Scale (HSPS), remains the most direct attempt to measure levels of neurosensitivity in humans (Pluess et al. 2018). The HSPS is a 27-item instrument that taps both general sensitivity to stimulation, along with a propensity to become easily overwhelmed in situations applying substantial stress to the nervous system. During its development, Cronbach's alphas ranging from .64 to .75 were reported across six different, age-diverse samples drawn from the US population (Aron and Aron 1997). Subsequent research has consistently reported alphas greater than .80, suggesting adequate levels of internal consistency reliability (Aron and Aron 2013). Regarding convergent validity, the HSPS correlated well (r = .64) with Mehrabian's (1977) measure of sensory screening. Lower correlations were observed with items measuring social introversion, which suggests good discriminant validity, given arguments that sensitivity is partially independent from introversion (Aron and Aron 1997).

Despite these strengths, questions still surround the psychometric properties of the HSPS. Aron and Aron (1997) offer few details on the psychometric principles and analyses that informed the scale's development. Of note is the absence of any detailed justification of the item phrasing employed, other than that, "items were based on observations from [an

interview-based study] and previous theory and research that seemed relevant to the construct of sensory processing sensitivity" (Aron and Aron 1997, p.352). An examination of the final HSPS item set (see Table Supp-1) reveals two potential concerns about response sets and biases. Firstly, the scale contains no items that are negatively/reverse phrased and reverse-scored. The use of reverse-scored items is a recommended practice in psychometric test construction, and assists in discouraging or detecting acquiescence and dissent response sets, and certain patterns of random responding (Kaplan and Saccuzzo 2012; Murphy and Davidshofer 2005). Secondly, a number of items are evaluatively loaded, and could be thought to encourage a 'faking good' response set (Tourangeau et al. 2009).

A further psychometric issue is the debate that surrounds the factor structure of the HSPS. Based on a significant decline between the first and second eigenvalues in their principle component analyses of the scale, Aron and Aron (1997) argued for a single-factor solution best conceptualised as "sensitivity". However, a subsequent analysis (Smolewska et al. 2006) demonstrated that a three-factor solution fits significantly better - a finding replicated by others (Booth et al. 2015; Lionetti et al. 2018; Sobocko and Zelenski 2015) and incorporated into the development of a child version of the HSPS (Pluess et al. 2018). These three factors have been labelled Ease of Excitation (EOE), Low Sensory Threshold (LST) and Aesthetic Sensitivity (AES). Meanwhile, for theoretical reasons, Evans and Rothbart (2008) preferred a two-factor solution with Orienting Sensitivity and Negative Affect as two orthogonal constructs. Lastly, Meyer and associates (Meyer et al. 2005) arrived at a four-factor solution when using the instrument as part of a larger examination of adults with borderline and avoidant features.

A limitation of these factor analyses is the lack of sample heterogeneity (Lionetti et al. 2018; Smolewska et al. 2006). Results were primarily based on participants of Caucasian ancestry of a generally Western culture. In response to this, Şengül-İnal and Sümer (2017) administered a translated HSPS to Turkish individuals of a relatively collectivist culture, and arrived at a unique four-factor solution (Sensitivity to Overstimulation, Sensitivity to External Stimulus, Aesthetic Sensitivity and Harm Avoidance) that outperformed twoand three-factor models. Part of the uniqueness of this solution was ascribed to cultural differences, which have been recommended as a focal point for further investigation (Pluess et al. 2018). The HSPS was not developed within culturally diverse samples, and there have been few subsequent attempts to confirm its reliability and freedom from cultural biases. Notable in this regard is that the instrument relies on language that is both colloquial and technical/sophisticated. An a priori analysis of the scale's language conducted by the authors revealed a number of phrases whose accessibility and interpretability were questionable outside of Anglo-American populations, and in



populations where English is not necessarily the home language of participants. Examples include "frazzled" (item 11), "rattled" (item 14), "shake you up" (item 21), "overwhelmed" (items 1 and 7), "subtleties" (item 2), and "conscientious" (item 12).

Given the potential of the HSPS to distinguish between individuals most susceptible to the potentially advantageous and disadvantageous aspects of their environment (including intervention efforts), it could serve as an ideal tool in low-to-middle income countries (LMICs) in which poverty and inequality pose threats to early childhood development (Black et al. 2017), and where psychosocial interventions are critically needed. Assessing levels of SPS might also serve as a more socially acceptable way of categorising individuals, rather than based on genetic differences that are more distally connected to phenotypic sensitivity (Belsky and van IJzendoorn 2015; Ellis et al. 2011). However, in the absence of resources for developing a locally appropriate equivalent (Laher and Cockcroft 2017), researchers in LMICs will need assurance that the HSPS is applicable across a variety of contexts.

Motivated by these concerns, the aim of the present study was to interrogate the HSPS, in a culturally- and linguistically-diverse setting, using multiple heterogeneous samples of South African origin. We sought to a) assess the factor structure of the scale, with and without researcher-generated reverse-scored items and b) confirm the proposed three-class distribution of sensitivity phenotypes. We began by recruiting a small pilot sample of university psychology students (Study 1) with the view of addressing cultural and/or language issues in the original HSPS items. We then administered the instrument to larger student samples (Study 2) and conducted factor and latent class analyses. For validation, we compared our findings from these student samples to a more general sample of residents in the Soweto-Johannesburg metropolitan area (Study 3).

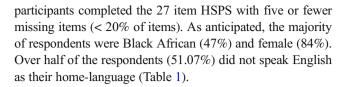
Study 1

Methods

Participants

Participants were drawn from psychology undergraduate classes at the University of the Witwatersrand, during the 2015 academic year. First-year psychology students were able to earn course credit through a student research participation programme. No exclusion criteria were applied, given that our primary aim was to gather commentary on comprehension and language-based issues. Demographically, the population of psychology students at the University is ethnically and linguistically diverse and predominantly female.

Of the 117 participants that responded to the invitation to participate (5% of psychology undergraduates), 94



Instrumentation

The pilot survey comprised the original HSPS (Aron and Aron 1997), along with three researcher-generated, reverse-scored items, which were scored along a 7-point scale (where 1 = Not at all, 4 = Moderately and 7 = Extremely). Following each item of the HSPS, participants were asked to note (via an open-ended response) any issues they had in comprehending the language of the item. This option was designed to give participants a chance to express their uncertainty about any of the items.

To address the lack of negatively-phrased items, three reverse-scored items were generated based on the "sensitivity-related variables" identified by Aron and Aron (1997).

Sensitivity-related variables were discerned from highly informative items originally used in the development of the HSPS, but that were not selected for the final instrument for various reasons (e.g. brevity; Aron and Aron 1997). These items related to the strong love intensity, country lifestyle preference and poor reaction to personal criticism commonly seen in HSPs (Aron and Aron 1997). Consequently, the reverse-scored items were: R1) "Do you often feel less emotional than other people (e.g. you cry less; you don't fall in love as easily etc.)?", R2) "Do you prefer your life to be very fast-paced (filled with activity)?" and R3) "Do you find personal criticism easy to deal with?"

Procedure

Following ethical approval from the University of the Witwatersrand Human Research Ethics Committee (Nonmedical; clearance certificate number H15/08/30), an invitation to research participation was extended to psychology undergraduates. After providing informed consent, participants completed the research survey electronically via SurveyMonkey®. Item order was randomised per participant to mitigate context effects. Prior to the survey items, participants were presented with optional demographic questions that elicited information on their age, sex, ethnicity and home language. No other identifying information was recorded. We removed participants that had over 20% missing data (remaining missing item responses were replaced with the arithmetic mean per item), but retained all responses to the open-ended questions regarding item comprehension. Data analysis was performed using R (version 3.5.1), RStudio (version 1.1.456) and the psych (version 1.8.4; Revelle 2018) and ggplot2 (version 3.0.0; Wickham 2016) packages.



Curr Psychol (2022) 41:4760-4774

Table 1 Descriptive statistics for all studies

Criterion	Pilot (2015)	2016	2017	2018 [†]	Combined	BTT+
n	94	262	168	240	750	1400
Age	22.36 (6.36)	19.33 (1.21)	19.24 (1.28)	19.32 (1.29)	19.38 (1.30)	28
Ethnicity (%)						
Black African	47.88	51.53	47.62	38.33	45.87	88.43
White	34.04	30.15	26.19	32.92	30.53	0.14
Indian	17.02	11.07	17.26	17.08	15.20	0.64
Mixed Ancestry	1.06	5.73	4.17	7.92	5.60	10.79
Other	0	1.53	4.76	3.33	2.67	0.00
Sex (% female)	84.04	81.68	77.98	84.17	82.00	52.71
Home language (%)						
English	48.93	50.00	55.95	59.17	54.27	NA
Non-English	51.07	50.00	44.05	40.83	45.60	NA
Total HSPS score *	121.93 (20.87)	121.74 (20.62)	124.21 (21.08)	122.41 (21.64)	122.61 (20.84)	105.19 (23.14)
Females	122.22 (20.77)	123.40 (20.54)	126.52 (20.08)	123.87 (21.75)	124.30 (20.70)	109.98 (22.45)
Males	120.53 (22.04)	114.33 (19.51)	116.05 (22.79)	114.68 (19.53)	114.97 (19.85)	99.85 (22.75)
Mean item score	4.52 (0.77)	4.51 (.76)	4.60 (.78)	4.53 (.80)	4.54 (.77)	3.90 (.86)
Cronbach's α	.87	.86	.85	.86	.86	.85
Guttman's λ^6	.93	.89	.90	.89	.88	.86

[†] In the 2018 sample, one participant chose not to report their ethnicity. Values reported as mean (± standard deviation). NA = not available

Results

The HSPS

In line with other published reports, the HSPS demonstrated good Cronbach's α and Guttman's λ^6 values (.87 and .93 respectively). Inter-item correlations were positive (average r = .20), as were item by scale correlations (average r = .47). Only three items had scale correlations < .30, namely items 2 (Do you seem to be aware of subtleties in your environment?), 4 (Do you tend to be more sensitive to pain?) and 18 (Do you make a point to avoid violent movies and TV shows?).

Researcher-generated reverse-scored items all had negative or near-zero correlations with HSPS items, except where a social desirability bias may have been in effect (e.g. items 10 [Are you deeply moved by the arts or music?], 15 [When people are uncomfortable in a physical environment do you tend to know what needs to be done to make it more comfortable?] and 22 [Do you notice and enjoy delicate or fine scents, tastes, sounds, works of art?], and items R2 and R3), creating areas of discrepant colouring on the correlation heat map (Figure Supp-1; Online Supplementary Material). On their own, reverse-scored items attained an α = .45, and were weakly correlated to each other (.16–.29). When scored together with HSPS items, the 30-item scale maintained an α = .87, but item-by-scale correlations for the reverse-scored items were weak (–.21 to –.29).

Total HSPS scores were normally distributed (W = .99, p = .43), with an average of 121.93. Mean item score and standard deviation (4.52 ± 77) were comparable to United States (3.96 ± .71) and German samples (4.54 ± .94; Aron and Aron 2013). There was no significant difference in HSPS score between male and female respondents (W = 682, p = .36, effect size r = .10), although the number of male respondents (n = 15) was notably small.

Overviewing comments provided by the participants highlighted several instances of comprehension difficulty. Firstly, in accord with our a priori analysis, three words proved problematic for students: "subtleties" (item 2), "frazzled" (item 11) and "conscientious" (item 12). There was also confusion regarding the scope of certain phrases. Students were unsure how to interpret "strong sensory input" (item 1), often remarking that the term was too broad. Many wondered what this term encompassed (physical input? auditory input? emotional input?) and were conflicted on how to respond if they found only very particular types of input overwhelming, but not others. Similar complaints were raised against the items "Do you tend to be more sensitive to pain?" and "Do changes in your life shake you up?", with many students querying the boundaries in defining "pain" (physical pain? emotional pain?) and "changes" (severity of change? positive or negative changes?). The most queried term was "inner life" (item 8). Students seemed unfamiliar with the concept, or otherwise understood the intent of the phrase, but questioned its exact definition. All aforementioned issues were uniformly raised



across participants, with no apparent trends according to home language or ethnicity.

Discussion

Due to limited previous administration in culturally diverse settings, we piloted the HSPS on a group of South African psychology undergraduates to assess scale performance and identify possible language and comprehension issues. On balance, the scale performed comparably to other published reports. However, due to a small number of male participants, we were unable to detect known differences in SPS scores between females and males (Aron and Aron 1997).

Regarding item comprehension, students expressed difficulty in understanding several terms, either as a consequence of their broad nature or because of their presumable absence from routine English conversation in South Africa. None of these issues could be ascribed to particular ethnic or language groups. Instead, some of the wording of the HSPS could be argued as being imprecise (subtleties, changes) and colloquial (frazzled), but comprehension issues might have also spoken to the general education and maturity level of undergraduates. Nevertheless, this is the first attempt, to date to interrogate the comprehension of HSPS items, and suggests some scope for improvement, especially in cross-cultural contexts.

Similarly, there have, to the best of the authors' knowledge, been no prior attempts to address the lack of negativelyphrased HSPS items. The reverse-scored items generated for this study performed adequately, and appeared to integrate with the HSPS without adverse consequences. Two of these items (R2 and R3) raised suspicions of a socially desirable response set, given their unexpectedly positive correlation with select HSPS items, which possibly weakens the scale's ability to delineate sensitive from non-sensitive individuals. Although at the cost of a longer overall instrument, reversescored items could help to detect and mitigate response biases, and may prove useful as the scale finds application in diverse samples and contexts. Our items provide a starting point, but could admittedly be improved, especially because they are based only on sensitivity-related features (Aron and Aron 1997), possibly explaining their weak item-by-scale correlations.

In summary, our pilot results identify possible areas of improvement for the HSPS, which remains unchanged (except for briefer and translated versions - Aron and Aron 2013; Ershova et al. 2018; Konrad and Herzberg 2019) and unchallenged as the primary psychological instrument for gauging neurosensitivity. Regardless, in its current form, the HSPS performed robustly even in a culturally heterogeneous sample. For our second study, aimed at factor and latent class analyses, we decided to clarify language-based issues ("subtleties", "frazzled" and "conscientious") before administering the

HSPS to a larger number of undergraduates. We chose not to clarify broad terms ("changes", "pain", "sensory input" and "inner life") over concerns that this might make the instrument too prescriptive.

Study 2

Methods

Participants

Participants were drawn exclusively from first year psychology classes during the 2016 (n = 262), 2017 (n = 168) and 2018 (n = 240) academic years. An age range of 18–25 years was enforced to avoid possible age-related effects on levels of sensitivity. Students were able to claim 1% course credit for participation and were entered into a prize draw for a monetary voucher.

Participants in each annual sample were predominantly Black African and female. Demographic variance was within reasonable limits, and psychometric criteria remained stable across samples. Given this stability, a combined sample was prepared (n = 750). To maximise sample size, the combined sample also included responses from participants younger than 26 from Study 1 (n = 80). Descriptive statistics for each independent sample, and the combined sample (n = 750), are reported in Table 1.

Instrumentation

The HSPS (including the reverse-scored item set) was included as part of a larger survey assessing personality and university adjustment. Select clarifications were added to HSPS items 2, 11 and 12 based on Study 1 results. The items themselves were not modified in any way; rather, a bracketed statement was placed after the item to offer clarification of problematic words (Tourangeau et al. 2009).

Procedure

Following ethical clearance (certificate number H16/05/34), an invitation was extended to first-year undergraduates and the same administration procedure was followed as detailed above. Participants with over 20% missing item responses were removed. Remaining missing values were imputed with the arithmetic mean per item. To afford a more rigorous assessment of the factor structure, separate exploratory factor analyses (EFA) were conducted and compared across the annual and combined samples. Confirmatory factor analysis (CFA) was performed on the combined dataset, with results compared against other published factor solutions. Vuong's test (Vuong 1989) for non-nested model comparisons was



used to check for significant fit differences between models. Finally, one- to four-class models were explored using latent class analysis (LCA) on the combined sample. All data analyses were performed using R (version 3.5.1) and RStudio (version 1.1.456). Packages employed included *psych* (version 1.8.4; Revelle 2018), *nFactors* (version 2.3.3; Raîche et al. 2013), *lavaan* (version 0.6–3; Rosseel 2012), *nonnest2* (version 0.5–2; Merkle et al. 2016), and *poLCA* (version 1.5.0; Linzer and Lewis 2011).

Results

The HSPS

Each annual sample performed comparably and in line with observations from Study 1 (Table 1). In the combined sample, scale total scores were normally distributed (W = .99, p = .38).

There was a significant difference in average scale total between males and females (t=4.91, p<.001) of moderate effect size (Hedges' g=.47). Inter-item correlations (average = .18) and item-by-scale correlations (average = .46) were similar to Study 1; however, items exhibiting scale correlations below .30 were different (items 8,12 and 15 compared to 2, 4 and 18 from Study 1).

Exploratory Factor Analysis

Each of the four samples (three annual and one combined) passed preparatory checks, including the Kaiser-Meyer-Olkin test, Bartlett's test of sphericity and thresholds for the determinant of the correlation matrices. In line with other results (Aron and Aron 2013), a large first eigenvalue (6.31–6.60 across samples) was observed that accounted for $\pm 24\%$ of the variance, followed by a second, smaller eigenvalue of 1.99–2.44, accounting for $\pm 8\%$ of the variance.

Choosing the number of factors to extract was based on a weight of evidence across multiple methods (Ford et al. 1986), including parallel analysis, optimal coordinates, sample size adjusted Bayesian Information Criterion (BIC), Velicer Minimum Average Partial test and the distribution of residuals. A five-factor solution was most commonly suggested, except for the 2018 sample in which a four-factor solution received more support. Four- and five-factor solutions were subsequently explored, first without rotation, and then with factors extracted via a minimum residual approach and rotated obliquely using oblimin rotation (SPS is known to have environmental and genetic correlates, thus latent factors are unlikely to be independent - Aron and Aron 1997; Kline 1994). A cut-off value of .30 was used to identify significant item loadings but cross-loading items were removed. Across samples, the five-factor solution was easiest to interpret and cohered best with SPS theory. Factor loadings for the combined sample (n = 750) produced the clearest separation of factors and are reported in Table 2. Note that these loadings were derived from a Pearson correlation matrix - results were negligibly different when using a polychoric correlation matrix (Holgado-Tello et al. 2010).

Based on the theme of their member items, the five factors were named as: Negative Affectivity (NA), Neural Sensitivity (NS), Propensity to Overwhelm (PO), Aesthetic Sensitivity (AS) and Careful Processing (CP). The item composition of these factors, per sample, is displayed in Table Supp-1 (Online Supplementary Materials). Three of these factors, namely PO, AS and CP, demonstrated high cross-sample stability. The NS and NA factors were less consistent in their composition with several items alternating between these factors from sample to sample. The delineation between NA and NS improved in the larger, combined sample where all factors attained acceptable reliability estimates (Table 2).

We retained 20 items that loaded significantly, each onto only a single factor. Only item 1 was excluded due to crossloading, on both the PO (.40) and NS (.38) factors. The other six excluded items (2, 6, 15, 18, 27) failed to load onto any of the extracted factors. Inter-factor correlations ranged from .01 (between NA and AS) to .46 (between NA and NS). The lack of correlation between NA and AS supports other studies which have found these to be two distinct features of HSPs (Sobocko and Zelenski 2015). Meanwhile, the moderate correlation between NA and NS possibly explains why select items alternated between these two factors across different samples.

Reverse-scored items continued to perform in line with expectations from Study 1. To further interrogate their appropriateness, we re-ran EFA with the inclusion of reverse-scored items in the combined student sample (Table Supp-2). A five-factor solution was again suggested. Items R1 and R3 loaded negatively on to the NS factor, replacing HSPS items 11 and 5. Item R2 loaded negatively on the NA factor. These additions lowered the reliability of the NA (.74 down from .75) and NS (.61 down from .69) subscales, as well as the entire scale (.85 down from .86).

Confirmatory factor analysis

To interrogate the validity of the five-factor solution, the model fit was compared with other published factor solutions using CFA. The one-factor (Aron and Aron 1997), two-factor (Evans and Rothbart 2008), three-factor (Smolewska et al. 2006), and four-factor (Şengül-İnal and Sümer 2017) solutions were forced on the dataset. Furthermore, we tested a bifactor model whereby all items were allowed to load onto a general factor of sensitivity as well as the factors from the three-factor model (Smolewska et al. 2006). For the bifactor model, factors were constrained to be orthogonal (Lionetti et al. 2018). Model fit was compared to our five-factor



 Table 2
 Factor loadings based on the pattern matrix

Item	NA	NS	РО	AS	CP *	_h 2
23. Do you find it unpleasant to have a lot going on at once?	.78					.67
16. Are you annoyed when people try to get you to do too many things at once?	.65					.47
14. Do you get rattled when you have a lot to do in a short amount of time?	.58					.49
26. When you must compete or be observed while performing a task, do you become so nervous or shaky that you do much worse than you would otherwise?	.37					.22
21. Do changes in your life shake you up?	.32					.31
20. Does being very hungry create a strong reaction in you, disrupting your concentration or mood?	.30					.16
13. Do you startle easily?		.57				.40
4. Do you tend to be more sensitive to pain?		.56				.31
3. Do other people's moods affect you?		.38				.23
11. Does your nervous system sometimes feel so frazzled that you just have to go off by yourself?		.38				.41
5. Do you find yourself needing to withdraw during busy days, into bed or into a darkened room or any place where you can have some privacy and relief from stimulation?		.30				.35
9. Are you made uncomfortable by loud noises?			.77			.56
25. Are you bothered by intense stimuli, like loud noises or chaotic scenes?			.73			.63
7. Are you easily overwhelmed by things like bright lights, strong smells, coarse fabrics, or sirens close by?			.57			.41
19. Do you become unpleasantly aroused when a lot is going on around you?			.31			.39
10. Are you deeply moved by the arts or music?				.76		.57
22. Do you notice and enjoy delicate or fine scents, tastes, sounds, works of art?				.74		.55
12. Are you conscientious?					.60	.35
24. Do you make it a high priority to arrange your life to avoid upsetting or overwhelming situations?					.59	.38
17. Do you try hard to avoid making mistakes or forgetting things?					.59	.43
Rotated sum of squared loadings	2.41	2.11	2.24	1.48	1.43	
Explained variance (%)	9	8	8	5	5	
Cronbach's α	.75	.69	.76	.74	.62	

^{*} NA Negative Affectivity, NS Neural Sensitivity, PO Propensity to Overwhelm, AS Aesthetic Sensitivity, CP Careful Processing

solution using several indices. These included: the Standardised Root Mean square Residual (SRMR), the Root Mean Square Error of Approximation (RMSEA), the Tucker-Lewis Index (TLI), and the BIC and Akaike Information Criterion (AIC). Since the chi-square statistic is sensitive to sample size, we followed the suggestion of dividing the statistic by the degrees of freedom (*df*) to produce a ratio for which values between 2 and 3 are considered as acceptable fits (Schermelleh-Engel et al. 2003). These fit indices, per factor solution, are compared in Table 3.

The five-factor model displayed good fit metrics across all indices. Moreover, its fit (for the combined student dataset) was noticeably better than other published solutions, enjoying a 10,000–20,000 point reduction in AIC/BIC values depending on the particular model comparison. The five-factor solution was the only model to achieve "good" and "marginal" fits for the RMSEA and TLI indices respectively. Similarly, only the five-factor model achieved an acceptable χ^2/df ratio. The five-factor model was a significantly better fit than the bifactor model (Vuong test: p < .001). Lastly, conducting CFA per language group revealed that the five-factor model fit best amongst both English and non-English home language

speakers, with a slightly stronger fit for non-English participants (Table Supp-3).

Latent Class Analysis

Measures of fit for one- to four-class solutions are displayed in Table 4. Several fit statistics supported a four-class solution, but this solution made little theoretical sense. Three classes were suggested by both a low BIC and high entropy score. Based on combined density plots (Fig. 1a), these three classes reflected low-, medium- and high-sensitivity groups akin to "dandelion", "tulip" and "orchid" theorising (Lionetti et al. 2018). Accordingly, 24.13% of students ranked as dandelions, 39.07% as tulips, and 36.80% as orchids, with a cut-off average item score of 4.03 separating dandelions from tulips and 4.82 separating tulips from orchids.

Discussion

The HSP scale was designed and psychometrically assessed in largely homogeneous samples. We aimed to address this



Table 3 Fit comparison for different HSPS factor solutions

Model	Reference	χ^2	df	χ^2/df	p value	SRMR	RMSEA	TLI	AIC	BIC *
Student										
One factor	Aron and Aron 1997	1908.54	324	5.89	< .001	.070	.081	.644	74,786	75,161
Two factor	Evans and Rothbart 2008	1428.51	274	5.21	< .001	.076	.073	.722	68,908	69,260
Three factor	Smolewska et al. 2006	1138.35	227	5.02	< .001	.081	.073	.703	64,219	64,551
Four factor	Şengül-İnal and Sümer 2017	1282.47	293	4.38	< .001	.054	.062	.767	74,009	74,490
Bifactor	Lionetti et al. 2018	1085.23	301	3.61	< .001	.054	.062	.810	74,009	74,490
Five factor	Study 2	454.68	160	2.84	< .001	.047	.050	.909	53,977	54,300
General populat	ion									
One factor		2173.39	324	6.71	< .001	.060	.064	.713	149,646	150,071
Two factor		1636.24	274	5.97	< .001	.057	.060	.773	138,321	138,719
Three factor		1335.49	227	5.88	< .001	.060	.059	.765	128,289	128,667
Four factor		1412.22	293	4.82	< .001	.056	.052	.820	143,194	143,635
Bifactor		1137.01	301	3.77	< .001	.045	.045	.861	148,656	149,201
Five factor	Study 3	613.38	160	3.83	< .001	.040	.045	.902	109,740	110,107

^{*} SRMR Standardised Root Mean square Residual, RMSEA Root Mean Square Error of Approximation, TLI Tucker-Lewis Index, AIC Akaike Information Criterion, BIC Bayesian Information Criterion

limitation by assessing the HSPS in South Africa, which is both ethnically diverse and has good representation of idiocentric and allocentric cultures (Burgess et al. 2002; Johnston 2015). While most of our findings served to support previously published claims, EFA of our data revealed a five-factor solution that has, to our knowledge, never been previously reported. Moreover, based on several fit measures, we found our five-factor solution to be a better fit than other published solutions.

A possible advantage of the five-factor solution is that it appears to overcome perceived weaknesses in the theoretical interpretations of other factor solutions. In the three-factor solution (EOE, AES and LST; Smolewska et al. 2006), which has been most widely adopted (e.g. Gerstenberg 2012; Lionetti et al. 2018; Liss et al. 2008; Sobocko and Zelenski 2015), several items seem incongruous with the latent factor

they supposedly measure. Some examples include: items 17 (Do you try hard to avoid making mistakes or forgetting things?) and 27 (When you were a child, did parents or teachers seem to see you as sensitive or shy?) as measures of Ease of Excitation; and items 5 (Do you find yourself needing to withdraw during busy days?) and 12 (Are you conscientious?) as measures of Aesthetic Sensitivity. Moreover, the AES factor is known to have low reliability ($\alpha = .55 - .61$; Liss et al. 2008; Sobocko and Zelenski 2015). Meanwhile, the two-factor solution (Negative Affectivity and Orienting Sensitivity; Evans and Rothbart 2008) makes arguably unfair assumptions that items 4 (Do you tend to be more sensitive to pain?), 9 (Are you made uncomfortable by loud noises?), and 13 (Do you startle easily?) reflect an individual's level of negative affectivity, rather than simply reflecting greater physiological sensitivity. In addition, the reliability of the Orienting

Table 4 Latent class analysis results

Model	residual df	log-likelihood	likelihood ratio	AIC	BIC	Entropy*
Student						
One class	588	-36,301.56	57,444.19	72,927.12	73,675.57	
Two class	425	-34,961.80	55,000.32	70,573.59	72,750.12	.892
Three class	262	-34,310.62	53,740.04	69,597.23	71,851.83	.904
Four class	99	-33,978.18	53,132.32	69,258.37	72,266.04	.900
General population	on					
One class	1238	-36,915.20	53,390.30	74,154.03	75,003.60	-
Two class	1075	-35,109.92	49,784.34	70,869.84	72,574.21	.850
Three class	912	-33,728.12	47,048.87	68,432.25	70,991.43	.866
Four class	749	-33,412.70	46,415.62	68,127.41	71,541.40	.848

^{*} df = degrees of freedom; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion. Bolded values represent the best fit per criterion



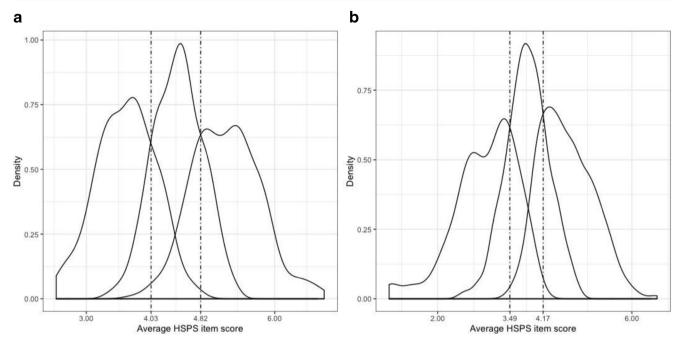


Fig. 1 Combined density plots for the three latent classes. Average HSPS item score densities are plotted for each of the three classes predicted by latent class analysis using *poLCA*. a. Amongst psychology students, relatively high cut-off values separated dandelions from tulips (4.03) and tulips from orchids (4.82). This is in keeping with the previous findings that psychology students rate themselves more sensitive than

members of the general population. b. Cut-off scores were notably lower for members of the BTT+ cohort, who better represent a general population sample. On balance, cut-off values were similar to those reported by Lionetti and colleagues (Lionetti et al. 2018) for a random sample of US psychology students, namely 3.71 and 4.66

Sensitivity subscale was found to be weak ($\alpha = .51-.56$) in samples of Canadian undergraduates (Sobocko and Zelenski 2015). Equivalent issues appear to affect the four-factor structure identified in a Turkish sample (Şengül-İnal and Sümer 2017). For example, a modified item 4 ("I tend to be very sensitive to pain") loads incongruously on the Aesthetic Sensitivity factor; and the weakly reliable ($\alpha = .55$) Harm Avoidance factor applies a seemingly negative label to an item set reflecting behaviour that might also be wisely cautious.

In contrast, our five-factor solution better aligns to the theoretical conceptualisation of a highly sensitive person. Aron (2019) cites four fundamental characteristics of HSPs, namely:

1) stronger emotional reactions, both positive and negative; 2) heightened awareness of sensory information; 3) deep cognitive processing and 4) high susceptibility to overstimulation. Similarly, our factor solution of the HSPS appears to make distinctions along these lines, in that NA and AS reflect positive and negative emotionality (1); NS relates to heightened sensory awareness (2); CP indicates deep processing (3) whilst PO encapsulates the tendency for overwhelm (4). Encouragingly, our five factors attained robust reliability estimates, with the possible exception of CP (α = .62). Despite these strengths, however, there are notable concerns with the five-factor solution. It is perhaps problematic that two factors are comprised of a small number of items. Delineation between the NA and NS factors was

weak, and appeared to fluctuate from sample to sample. Finally, the extraction of five factors is less parsimonious than solutions with fewer factors.

For what exact reasons our samples generate a unique factor solution remains unclear, although several possibilities bear mention. We are most inclined to attribute the unique factor structure to our sample diversity, knowing that ethnic/ cultural differences have been noted to affect factor structure in other instrumentation (e.g. Allen et al. 2007; Asner-Self et al. 2006; Mylonas 2009). South Africa is highly multicultural, with strong collectivist and individualist influences. On one hand, the African philosophy of "Ubuntu" has fostered collectivist trends evident amongst the value orientations of all residents, whilst on the other hand, forces of migration, urbanisation and affluence have been driving shifts towards individualism (Burgess et al. 2002). This is especially pertinent to urban Black South Africans (the majority of our respondents) who juggle the more collectivist influences of African culture and the individualism characteristic of urban living. These contrasting cultural views are known to value individuals of sensitive disposition in opposite fashion - sensitivity is mostly frowned upon amongst individualists, but highly praised amongst collectivists (Aron 2004). However, sensitivity is an umbrella construct that encompasses an array of different physiological features and behaviours (Liss et al. 2008). It is conceivable that more nuanced relationships exist between culture and sensitivity - sub-facets of sensitivity may be



differentially valued in a manner that isn't captured by merely distinguishing opinion on sensitive and non-sensitive individuals. That SPS might encompass a larger number of latent factors may point towards a finer discrimination among attributes of sensitivity, and how these are differentially viewed and valued by people of different backgrounds. Similar reasoning was offered by Şengül-İnal and Sümer (2017) for the lack of fit of one-, two- and three-factor models in their Turkish sample.

An additional explanation for the larger factor structure is that language issues inflated item-response variance, prompting the need for more factors to be extracted. However, we made special effort to try address language issues beforehand (Study 1), and we added supporting statements in order to aid the comprehension of some items. Of the three items that did require some clarification, two (items 11 and 12) loaded on to logical factors, whilst the remaining item (item 2) was discarded due to poor loadings. Correlations with reverse-scored items were also in the expected direction and their inclusion as part of EFA did not destabilise the fivefactor structure. Although many South Africans do not have English as their first or home language (Statistics South Africa 2012), our sample of students are proficient enough to manage tertiary-level education in English. It thus seems reasonable to assume that the HSPS was interpreted as intended, especially given that all psychometric properties were comparable to other studies. Nevertheless, we may have failed to identify other sample specific characteristics that explain the larger factor structure.

A final possibility is that our factor solution reinforces cautions already expressed by Aron and Aron (2013). For several reasons, the HSPS is not completely amenable to factor analysis. Most importantly, the instrument is targeted towards only 20–35% of the population, meaning that the bulk of the sample variance in scale item response is attributable to less sensitive individuals. Furthermore, there exists a sex bias in that females tend to score higher than males (although the overall proportion of male and female HSPs is equal; Aron and Aron 1997). This bias is likely a consequence of cultural differences in the ways males and females are socialised (Aron 2006; Aron and Aron 2013). In support of sex differences, we found a significant difference between scale total averages for males and females. Lastly, some items are prone to a social desirability bias (which we noted in our analyses), but this effect is possibly balanced out by items that assess how easily overaroused participants are (which is not a socially desirable characteristic; Aron and Aron 1997). Our unique factor solution might be testament to these issues, which manifest differently across different samples, perhaps exacerbated by our sample's heterogeneity. Combining these issues with the common acknowledgement that factor analysis tends to be poorly performed and interpreted (Ford et al. 1986; Kline 1994)

might further justify why there has been such inconsistency in factor analytic results for this instrument.

Encouragingly, the results of our LCA align with those previously published. In subsamples of US psychology students, Lionetti and associates (Lionetti et al. 2018) also obtained varied support for either three- or four-class solutions, with the balance of evidence favouring a three class solution. Our cut-off scores were marginally higher (4.03 and 4.82 vs 3.71 and 4.66) and our sample comprised fewer dandelions but more orchids. This possibly reflects the larger percentage of female participants in our sample (82.00% vs 62.30%).

Study 3

Methods

Participants

To validate our findings in an independent sample of urban South African residents, we arranged for the HSPS to be administered to members of a longitudinal cohort known as Birth to Twenty Plus (BTT+). The BTT+ is South Africa's largest longitudinal cohort, with 3273 enrolments. Initiated in 1990, cohort members were recruited, during pregnancy and birth, from hospitals in the Soweto-Johannesburg region during a seven-week period. The growth, development and well-being of these members has since been tracked through near-annual data collection waves (as described elsewhere; see Richter et al. 2007).

A total of 1400 cohort members (52.71% female) completed the HSPS with fewer than 20% of missing responses. Descriptive statistics are summarised in Table 1. Data on the home language of participants was not available for this study, but according to figures from the 2011 census (Statistics South Africa 2011), English first-language speakers account for just 2.34% of Soweto residents, although many are able to adequately converse in English, including all BTT+ members. Similarly, we did not have access to detailed records of the educational background of our subset of participants, but summary statistics on the overall cohort indicate that, by age 22, 38% had not matriculated, 46% had matriculated and 16% had completed one or more years of university education. Education levels were thus more varied than within our undergraduate samples.

Procedure

As part of the 2018/2019 data collection wave, cohort members completed a battery of personality and cognitive tests, including the HSPS (and reverse-scored items) via interview with a trained research assistant (ethical clearance certificate number H16/05/34). The assistant read each question in turn,



and asked each cohort member to provide a response along the 7-point scale on a digital tablet interface.

Scale responses were cleaned, scored and analysed as per the student studies. CFA was conducted using the five-factor model from Study 2 and results were compared to other published factor models. LCA for one to four classes was also carried out.

Results

The HSPS

The HSPS continued to perform reliably (α = .85), but scores were not normally distributed (W = .99, p = < .01) - an issue that seemed to be driven by male respondents, since female respondents had normally distributed scores. Original HSPS item correlations closely mirrored those in Figure Supp-1, with an average item correlation of .18. However, against expectations, correlations between reverse-scored items and original HSPS items were near zero or weakly positive (r = 0.0–0.2). The average item-by-scale correlation was .42, with items 6 and 18 having values < .30. Compared to psychology students, the members of the BTT+ cohort had substantially lower total and average item scores (Table 1), although not beyond the range of what has been previously reported. Females continued to score significantly higher than males (W = 185,989, p < .001, r = .21).

To interrogate the skewed distribution of male respondents, outliers were identified following graphical inspections of the data. At least 16 male respondents (and one female respondent) had apparently "faked non-sensitive" by responding "1" (on the 7-point Likert scale) for almost every item, including the reverse-scored items. Exclusion of these individuals restored the distribution of responses to normal for male individuals.

Roughly 15% of participants restricted the overwhelming majority of their item responses to just 1, 4, and 7 on the 7point scale. These happen to be the three scale points that are descriptively labelled (1 = not at all, 4 = moderately, 7 = extremely; Aron and Aron 1997). We can think of two possible, non-mutually exclusive explanations for this behaviour: 1) cohort members were experiencing test fatigue given the battery of other assessments they were expected to complete and/or 2) participants may have found it easier to work within a three-point scoring system, especially if there was any anxiety about answering personal questions posed in a language that was not their default. To investigate if these responses were a source of bias, we examined them separately and confirmed that they still had a similar distribution (W = .99, p = <.01) and were of adequate internal consistency reliability $(\alpha = .72)$. For the purposes of CFA, we chose to retain these individuals and responses in order to maximise sample size.

However, to avoid the artefactual identification of these individuals as a latent class, we collapsed all item responses to a 3-point scale for the latent class analysis (original item responses were coded as follows: 1-2=1; 3-5=2 and 6-7=3).

Confirmatory Factor Analysis

The five-factor model identified in Study 2 displayed strong fit statistics and remained, on average, the best performing model (Table 3). Other factor solutions, especially the bifactor, also showed adequate to good fit statistics and posed more of a challenge to the five-factor solution than was observed in the student samples. However, based on the Vuong test, the five-factor model fit significantly better than the bifactor model (p < .001).

Latent Class Analysis

Outcomes for the LCA closely mirrored those from Study 2 (Table 4). Again, a four-class solution received some support but was not easily interpretable. The BIC and entropy values favoured a three-class solution. When plotting density curves (Fig. 1b) for the three predicted classes, there was further evidence for dandelion, tulip and orchid groups. In line with the lower total scores, the cut-off scores between these classes were smaller when compared to the student sample (3.49 and 4.17 vs 4.03 and 4.82). Regarding the sample composition, 28.43% of respondents were dandelions, 35.57% tulips and 36.00% orchids.

There were notable differences in class frequency when stratified by sex. Amongst males, 36.86% were dandelions, 33.84% were tulips and 29.31% were orchids. Amongst females, 20.97% were dandelions, 37.13% were tulips and 42.01% were orchids.

Discussion

Our psychometric investigations of the HSPS in the BTT+ sample provided reasonable support for the findings of the previous student studies. The HSPS maintained good reliability, and inter-item and item-by-scale correlations were only moderately weaker than those observed in Study 2. The five-factor model retained an adequate fit, although it was more closely challenged by the bifactor model. The fit of different latent class solutions displayed an identical pattern to those in Study 2, even though we had collapsed participant responses to a three point scale for consistency. Understandably, cut-off scores separating the different sensitivity groups shifted downwards in accordance with the lower average HSPS scores. These mostly consistent findings were encouraging, given that the age, demographics and likely educational background of cohort members were markedly



different from our student studies. Also, the BTT+ sample had a better balance of males and females, which was a significant limitation in the student samples. Although almost 90% of Study 3 participants were Black African, the Apartheid-era category 'Black' does not (and never did) mark out a culturally homogeneous grouping, and former townships such as Soweto are typically marked by considerable ethno-cultural and linguistic diversity, as well as more and less urbanized individuals (Burgess et al. 2002). This would seem to suggest that the HSPS works sufficiently well in heterogeneous samples.

However, there were important concerns raised on closer inspection of the results. Unlike in Study 2, reverse-scored items correlated positively (albeit weakly) with original HSPS items. This finding could imply that the majority of respondents did not adjust their response pattern when item wording switched from positive to negative. Several factors may have contributed to this issue. Firstly, the item order was not randomised per participant. Reverse-scored questions were always asked after original HSPS items, which could have resulted in context effects (Tourangeau et al. 2009). Possibly, respondents were overly primed by the first 27 items, and failed to adjust their response set accordingly. Secondly, reverse-scored items are known to increase cognitive burden and respondent frustration (Kulas et al. 2018). Such issues were possibly exacerbated by test fatigue, explaining both the preference for descriptively labelled points on the Likert scale (1, 4 and 7), and the lack of response adjustment for reverse-scored items. Lastly, the generally lower levels of education achievement, and more limited familiarity with the English language of BTT+ cohort members might have additionally confounded the proper interpretation of negatively-worded items. On the other hand, it is possible that response trends on the reverse-scored items imply that they do not track core aspects of SPS (such that an option for future research would be to generate and evaluate at least one reverse-phrased item for each factor identified in the HSPS).

A second concern was the small percentage of male respondents that appeared to have "faked non-sensitive". These individuals (who were mainly Black African, save for two mixed-ancestry participants) were identified because of their outlying average item scores, and because of their handling of reverse-scored items. Why the issue was almost exclusively limited to male individuals may be a consequence of Black African culture (broadly considered), which is heavily patriarchal (Morrell et al. 2013). There is substantial pressure amongst African males (and males more generally) to exhibit masculinity, to the extent that some males resort to forceful displays of violence, against both women and other men (Morrell et al. 2013). This pressure may discourage African males from admitting to the sort of sensitive behaviour tapped by the HSPS (especially when administered by interviewers,

all of whom were female), more so than the typical tendency of men to score lower than females (Aron and Aron 2013). In support of this notion, there was a notably high percentage of dandelion males (36.86%), but a lower percentage of orchids (29.31%). The inverse pattern was true for females, despite arguments of an equal male:female ratio for highly sensitive persons (Aron and Aron 2013). Although limited to a small enough handful of participants so as not to noticeably interfere with the psychometric results of the HSPS, this response set bias potentially highlights the instrument's need for reverse-scored items, especially in contexts where denying, or admitting to sensitivity is socially/instrumentally desirable.

Note that although a large and diverse sample of urban South Africans, our BTT+ sub-sample was not reflective of South Africa's current demographic which limits the generalisability of these findings. However, the sample is still fairly representative of a low-to-middle income bracket which is where many psychosocial interventions are targeted and thus the audience for which the HSPS might be particularly pertinent.

General Discussion

A growing body of research continues to highlight the importance of individual differences in neurosensitivity (Greven et al. 2019). Not only does neurosensitivity provide a novel means of stratifying research participants (Keers and Pluess 2017; Morgan et al. 2017), it also carries significant implications for the psychological care and treatment of individuals (Belsky and van IJzendoorn 2015). As a LMIC, South Africa stands to greatly benefit from any research that might enhance intervention strategies. To this end, an inexpensive, socially acceptable psychological instrument assessing neurosensitivity, such as the HSPS, would be an ideal addition to future psychosocial research, provided that the instrument performs reliably and with high validity.

Across all of our studies, we found evidence to support the use of the HSPS in the South African context, despite the extensive heterogeneity of the population. Reliability remained sufficiently high, and meaningful factor and latent class structures were evident. So long as participants are conversant in English, the instrument appears to maintain comparable psychometric performance levels to other reported samples.

However, our factor solution was different to what has been previously reported. The continued lack of factor structure consistency for the HSPS likely reinforces the notion that psychometric evaluation can be a challenging endeavour that is further complicated by rich cultural and ethnic diversity (Byrne and van de Vijver 2010; Laher and Cockcroft 2017; Mylonas et al. 2014). This seems particularly relevant to an instrument that broadly measures "sensitivity" - a word with



an array of different meanings; may have differential cultural interpretations and value, and has historically been tied (erroneously) to other psychological constructs such as introversion (Eysenck 1992), inhibitedness (Kagan 1994) and shyness (Aron and Aron 1997; Cheek and Buss 1981). Participants in our pilot study spoke clearly to this issue, frequently questioning the items and terms that try to accommodate the many ways in which one might be "sensitive".

Cultural heterogeneity may thus explain the unique factor structure, but it did not appear to undermine the instrument, possibly because its underlying theory is grounded in evolutionary and comparative biology (Aron et al. 2012) rather than in cultural understandings of sensitive behaviour. Indeed, regardless of how many sub-factors the construct comprises, we are inclined to agree with Smolewska et al. (2006), and Lionetti et al. (2018), that the HSPS still functions as an appropriate measure for the higher-order construct of SPS. To date, there has been little evidence to suggest that the instrument does not make the meaningful distinction between sensitivity levels (i.e. orchids, tulips and dandelions) that it sets out to achieve. Rather, the SPS construct continues to gain support from genetic (Chen et al. 2011; Licht et al. 2011) and brain imaging (Acevedo et al. 2018; Acevedo et al. 2017) investigations, and embeds well within the larger framework of Environmental Sensitivity (albeit in mostly Caucasian samples thus far; Greven et al. 2019).

Our experimentation with reverse-scored items did reveal a possible means for strengthening the HSPS. These items worked as intended amongst student participants, but were less effective for BTT+ cohort members, likely because of study design limitations, but perhaps also because the items did not tap core SPS features. Nevertheless, the items did appear to help highlight a small proportion of Study 3 respondents who were faking their sensitivity level. The extent of this sort of response set bias has not been properly addressed before, and seems an important consideration for continued use of the HSPS in South Africa and other settings where there is strong pressure to view sensitivity in particular, socially-driven ways.

Some limitations to our studies warrant discussion. Firstly, our student samples were biased towards females which may have unfairly skewed the results and limits the generalisability of the findings. Observations from the more general South African sample, however, were mostly in keeping with the student results, although the five-factor structure had a less convincing fit and the average total HSPS score was appreciably lower. Secondly, we relied solely on participant self-report, the results of which would have been better supported through additional genetic, physiological and/or behavioural data. Thirdly, we did not attempt to exclude participants based on any history of disorder (e.g. sensory processing disorder) or medication use that may have impacted upon measures of sensitivity. Such individuals may have biased our results, but were presumably

minimal in number. Lastly, we did not compare factor solutions by first returning all excluded items to their highest-loading factor (Şengül-İnal and Sümer 2017).

In conclusion, the HSPS remains a strongly reliable instrument although with room for improvement. The instrument has generated varied factor solutions, none of which have received consistent support. This likely reflects the complex relationship between culture and sensitivity as well as the historical confusion in defining "sensitivity" and its attendant features and behaviours. Additionally, item phrasing has not been significantly interrogated for flaws, which might contribute to factor structure instability. The current lack of reversescored items limits identification of possible issues in respondent comprehension or response set bias. These issues will need to be considered as the HSPS finds application in diverse settings. Nevertheless, SPS theory is underpinned by strong cross-disciplinary support and the stratification of participants by levels of sensitivity continues to offer unique insights. Short of a locally developed equivalent, some minor clarifications can help make the HSPS more accessible to crosscultural samples and so researchers should consider seriously the addition of the HSPS in future psychosocial studies, especially where crossover interactions of the sort predicted by SPS and ES might be evident.

Acknowledgements The authors would like to thank Feziwe Mpondo and her team of research assistants for administering the HSPS to BTT+cohort members, along with all cohort members and students that participated in this study. We thank Francesca Lionetti and Michael Pluess for their guidance and statistical assistance. The reviewers are thanked for their constructive feedback. The DSI-NRF Center of Excellence in Human Development is thanked for their funding contribution.

Author contributions All authors contributed to the study conception and design. Data collection was facilitated by Shane A. Norris and Linda M. Richter. Data analysis was performed by Andrew May. The first draft of the manuscript was written by Andrew May and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding information The support of the DSI-NRF Centre of Excellence in Human Development towards this research/activity is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the CoE in Human Development

Availability of data and code Raw data and R code are available as a Mendeley dataset (https://doi.org/10.17632/32p7btkgr4.1.)

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to declare.

Ethical Approval This study was approved by the University of the Witwatersrand's Human Research Ethics Committee (Non-medical). Clearance certificate number: H16/05/34.



Informed Consent All study participants provided written informed consent before completing the study survey.

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

- Acevedo, B. P., Jagiellowicz, J., Aron, E. N., Marhenke, R., & Aron, A. (2017). Sensory Processing Sensitivity and Childhood Quality's Effects on Neural Responses To Emotional Stimuli. *Clinical Neuropsychiatry*, 14(6), 359–373.
- Acevedo, B. P., Aron, E. N., Pospos, S., & Jessen, D. (2018). The functional highly sensitive brain: A review of the brain circuits underlying sensory processing sensitivity and seemingly related disorders. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(20170161).
- Allen, K. D., DeVellis, R. F., Renner, J. B., Kraus, V. B., & Jordan, J. M. (2007). Validity and factor structure of the AUSCAN Osteoarthritis Hand Index in a community-based sample. *Osteoarthritis and Cartilage*, 15(7), 830–836.
- Aron, E. N. (2004). Revisiting Jung's concept of innate sensitiveness. *Journal of Analytical Psychology*, 49, 337–367.
- Aron, E. N. (2006). The Clinical Implications of Jung's Concept of Sensitiveness. *Journal of Jungian Theory and Practice*, 8(2), 11–43.
- Aron, E. N. (2019). The Highly Sensitive Person. Retrieved September 25, 2018, from http://hsperson.com/faq/evidence-for-does/
- Aron, E. N., & Aron, A. (1997). Sensory-processing sensitivity and its relation to introversion and emotionality. *Journal of Personality and Social Psychology*, 73(2), 345–368.
- Aron, E. N., & Aron, A. (2013). Tips for SPS research (revised November 21, 2013). Retrieved March 22, 2018, from http://hsperson.com/pdf/ Tips%7B%5C%7Dfor%7B%5C%7DSPS%7B%5C%7DSPS%7B%5C%7DResearchers%7B%5C%7DNov21%7B%5C%7D2013.pdf.
- Aron, E. N., Aron, A., & Davies, K. M. (2005). Adult shyness: the interaction of temperamental sensitivity and an adverse childhood environment. *Personality & Social Psychology Bulletin*, 31(2), 181–197.
- 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.
- Asner-Self, K. K., Schreiber, J. B., & Marotta, S. A. (2006). A crosscultural analysis of the brief symptom Inventory-18. *Cultural Diversity and Ethnic Minority Psychology*, 12(2), 367–375.
- Belsky, J., & Pluess, M. (2013). Beyond risk, resilience, and dysregulation: Phenotypic plasticity and human development. *Development and Psychopathology*, 25, 1243–1261.
- Belsky, J., & van IJzendoorn, M. H. (2015). What works for whom? Genetic moderation of intervention efficacy. *Development and Psychopathology*, 27(1), 1–6.

- Belsky, J., Jonassaint, C., Pluess, M., Stanton, M., Brummett, B., & Williams, R. (2009). Vulnerability genes or plasticity genes? *Molecular Psychiatry*, 14(8), 746–754.
- Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., ... Grantham-McGregor, S. (2017). Early childhood development coming of age: Science through the life course. *The Lancet*, 389(10064), 77–90.
- Booth, C., Standage, H., & Fox, E. (2015). Sensory-processing sensitivity moderates the association between childhood experiences and adult life satisfaction. *Personality and Individual Differences*, 87, 24–29.
- Burgess, S. M., Harris, M., & Mattes, R. B. (2002). SA tribes: Who we are, how we live and what we want from life in the new South Africa. Cape Town: New Africa Books.
- Byrne, B. M., & van de Vijver, F. J. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*, 10(2), 107–132.
- Cheek, J. M., & Buss, A. H. (1981). Shyness and sociability. *Journal of Personality and Social Psychology*, 41(2), 330–339.
- 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.
- Ellis, B. J., Boyce, W. T., Belsky, J., & Bakermans-Kranenburg, M. J. (2011). Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. *Development and Psychopathology*, 23, 7–28.
- Ershova, R. V., Yarmotz, E. V., Koryagina, T. M., Shlyakhta, D. A., & Tarnow, E. (2018). Operationalization of the Russian Version of Highly Sensitive Person Scale. *RUDN Journal of Psychology and Pedagogics*, 15(1), 22–37.
- Evans, D. E., & Rothbart, M. K. (2008). Temperamental sensitivity: Two constructs or one? *Personality and Individual Differences*, 44, 108–118.
- Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality* and *Individual Differences*, 13(6), 667–673.
- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39(2), 291–314.
- Gerstenberg, F. X. (2012). Sensory-processing sensitivity predicts performance on a visual search task followed by an increase in perceived stress. *Personality and Individual Differences*, 53(4), 496–500.
- Greven, C. U., Lionetti, F., Booth, C., Aron, E. N., Fox, E., Schendan, H. E., Pluess, M., Bruining, H., Acevedo, B., Bijttebier, P., & Homberg, J. (2019). Sensory processing sensitivity in the context of environmental sensitivity: A critical review and development of research agenda. *Neuroscience and Biobehavioral Reviews*, 98(January), 287–305.
- Holgado-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality and Quantity*, 44, 153–166.
- Johnston, E. R. (2015). South African clinical psychology's response to cultural diversity, globalisation and multiculturalism: A review. South Africa Journal of Psychology, 45(3), 374–385.
- Kagan, J. (1994). Galen's prophecy: Temperament in human nature. New York: Basic Books.
- Kaplan, R. M., & Saccuzzo, D. P. (2012). Psychological Testing: Principles, Applications, and Issues (8th ed.). San Diego: Wadsworth Publishing.
- Keers, R., & Pluess, M. (2017). Childhood quality influences genetic sensitivity to environmental influences across adulthood: A lifecourse gene x environment interaction study. *Development and Psychopathology*, 29(5), 1921–1933.
- Kline, P. (1994). An easy guide to factor analysis. London: Routledge.



- Konrad, S., & Herzberg, P. Y. (2019). Psychometric properties and validation of a German high sensitive person scale (HSPS-G). European Journal of Psychological Assessment, 35(3), 364–378.
- Kulas, J. T., Klahr, R., & Knights, L. (2018). Confound It!: Social Desirability and the "reverse-Scoring" Method Effect. European Journal of Psychological Assessment, 35(6), 855–867.
- Laher, S., & Cockcroft, K. (2017). Moving from culturally biased to culturally responsive assessment practices in low-resource, multicultural settings. *Professional Psychology: Research and Practice*, 48(2), 115–121.
- Licht, C. L., Mortensen, E. L., & Knudsen, G. M. (2011). Association between Sensory Processing Sensitivity and the 5-HTTLPR Short / Short Genotype. *Biological Psychiatry*, 69, 152S–153S Abstract 510.
- Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software*, 42(10), 1–18.
- Lionetti, F., Aron, A., Aron, E. N., Burns, G. L., Jagiellowicz, J., & Pluess, M. (2018). Dandelions, tulips and orchids: Evidence for the existence of low-sensitive, medium-sensitive and highsensitive individuals. *Translational Psychiatry*, 8(24), 24.
- Liss, M., Mailloux, J., & Erchull, M. J. (2008). The relationships between sensory processing sensitivity, alexithymia, autism, depression, and anxiety. *Personality and Individual Differences*, 45, 255–259.
- Mehrabian, A. (1977). Individual differences in stimulus screening and arousability. *Journal of Personality*, 45(2), 237–250.
- Merkle, E. C., You, D., & Preacher, K. J. (2016). Testing nonnested structural equation models. *Psychological Methods*, 21(2), 151– 163.
- Meyer, B., Ajchenbrenner, M., & Bowles, D. P. (2005). Sensory sensitivity, attachment experiences, and rejection responses among adults with borderline and avoidant features. *Journal of Personality Disorders*, 19(6), 641–658.
- Morgan, B., Kumsta, R., Fearon, P., Moser, D., Skeen, S., Cooper, P., Murray, L., Moran, G., & Tomlinson, M. (2017). Serotonin transporter gene (SLC6A4) polymorphism and susceptibility to a homevisiting maternal-infant attachment intervention delivered by community health workers in South Africa: Reanalysis of a randomized controlled trial. *PLoS Medicine*, 14(2), e1002237.
- Morrell, R., Jewkes, R., Lindegger, G., & Hamlall, V. (2013). Hegemonic Masculinity: Reviewing the Gendered Analysis of Men's Power in South Africa. South African Review of Sociology, 44(1), 3–21.
- Murphy, K. R., & Davidshofer, C. O. (2005). Psychological testing: principles and applications (6th ed.). Upper Saddle River: Pearson/Prentice Hall.
- Mylonas, K. (2009). Reducing Bias in cross-cultural factor analysis through a statistical technique for metric adjustment: Factor solutions for quintets and quartets of countries. In A. Gari & K. Mylonas (Eds.), Quod erat demonstrandum: From herodotus' ethnographic journeys to cross-cultural research. Athens: Pedio Books Publishing.
- Mylonas, K., Furnham, A., Divale, W., Leblebici, C., Gondim, S., Moniz, A., ... Boski, P. (2014). Bias in Terms of Culture and a Method for Reducing It: An Eight-Country "Explanations of Unemployment Scale" Study. Educational and Psychological Measurement, 74(1), 77–96
- Nocentini, A., Menesini, E., & Pluess, M. (2018). The Personality Trait of Environmental Sensitivity Predicts Children's Positive Response to School-Based Antibullying Intervention. *Clinical Psychological Science*, 6(6), 848–859.

- Pluess, M. (2015). Individual Differences in Environmental Sensitivity. Child Development Perspectives, 9(3), 138–143.
- 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.
- 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.
- Raîche, G., Walls, T., Magis, D., Riopel, M., & Blais, J.-G. (2013). Non-graphical solutions for Cattell's Scree test. *Methodology*, 9, 23–29.
- Ramchandani, P. G., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2010). Differential susceptibility to fathers' care and involvement: The moderating effect of infant reactivity. *Family Science*, 1(2), 93–101.
- Revelle, W. (2018). Psych: Procedures for personality and psychological research. Northwestern University, Evanston, Illinois, USA. Retrieved from https://cran.r-project.org/package=psych.
- Richter, L., Norris, S., Pettifor, J., Yach, D., & Cameron, N. (2007). Cohort profile: Mandela's children: The 1990 birth to twenty study in South Africa. *International Journal of Epidemiology*, 1–8.
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Schermelleh-Engel, K., Müller, H., & Moosbrugger, H. (2003). Evaluating the fit of structural equation models. *Methods of Psychological Research*, 8(2), 23–74.
- Şengül-İnal, G., & Sümer, N. (2017). Exploring the Multidimensional Structure of Sensory Processing Sensitivity in Turkish Samples. Current Psychology, 1(2007).
- Slagt, M., Dubas, J. S., & van Aken, M. A. (2015). Differential Susceptibility to Parenting in Middle Childhood: Do Impulsivity, Effortful Control and Negative Emotionality Indicate Susceptibility or Vulnerability? *Infant and Child Development*, 25(4), 302–324.
- Smolewska, K. A., McCabe, S. B., & Woody, E. Z. (2006). 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". Personality and Individual Differences, 40(6), 1269–1279.
- Sobocko, K., & Zelenski, J. M. (2015). Trait sensory-processing sensitivity and subjective well-being: Distinctive associations for different aspects of sensitivity. *Personality and Individual Differences*, 83, 44–49.
- Statistics South Africa. (2011). South African Census Community Profiles. DataFirst. Retrieved from https://www.datafirst.uct.ac.za/ dataportal/index.php/catalog/517/study-description
- Statistics South Africa. (2012). Census 2011: Census in brief. Pretoria. Retrieved from http://www.statssa.gov.za/census/census%7B%5C%7D2011/census%7B%5C%7Dproducts/Census%7B%5C%7D2011%7B%5C%7DCensus%7B%5C%7Din%7B%5C%7Dbrief.pdf.
- Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2009). *The psychology of survey response*. Cambridge: Cambridge University Press.
- Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, 57(2), 307–333.
- Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. New York: Springer-Verlag.

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

