Journal of Personality Assessment

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hjpa20

The DOES Scale: Measuring Sensory Processing Sensitivity as a Trait Constellation

Danièle Anne Gubler, Tobias Janelt, Marcus Roth, Katja Schlegel, Jasmin Guggisberg & Stefan Johannes Troche

To cite this article: Danièle Anne Gubler, Tobias Janelt, Marcus Roth, Katja Schlegel, Jasmin Guggisberg & Stefan Johannes Troche (25 Sep 2024): The DOES Scale: Measuring Sensory Processing Sensitivity as a Trait Constellation, Journal of Personality Assessment, DOI: 10.1080/00223891.2024.2405536

To link to this article: https://doi.org/10.1080/00223891.2024.2405536

JOURNAL OF PERSONALITY ASSESSMENT
https://doi.org/10.1080/00223891.2024.2405536

Routledge
Taylor & Francis Group

The DOES Scale: Measuring Sensory Processing Sensitivity as a Trait Constellation

Danièle Anne Gubler[1], Tobias Janelt[2], Marcus Roth[2], Katja Schlegel[1], Jasmin Guggisberg[1] and Stefan Johannes Troche[1]

[1]Department of Psychology, University of Bern, Bern, Switzerland; [2]Department of Psychology, University of Duisburg-Essen, Duisburg-Essen, Germany

ABSTRACT

Based on Aron's (2020) DOES model, we developed the DOES Scale to measure Sensory Processing Sensitivity (SPS) with four dimensions: Depth of Processing, Overstimulation, Emotional Reactivity, and Sensing the Subtle. Using interview data from the study by Roth et al. (2023), we created a 20-item questionnaire (5 items per dimension) in German and English. In three studies with 1,365 subjects from Switzerland, Germany, Austria, and the UK, we evaluated the psychometric properties of the scale using confirmatory factor analysis and examined construct validity with the established Highly Sensitive Person Scale (HSPS) and different personality measures. The results confirmed each subscale's unidimensionality and good psychometric properties. Considering the four subscales together indicated that they could be best described as correlated factors rather than in terms of a second-order factor. Convergent validity was confirmed, especially for Overstimulation in its association with the HSPS total score and its subscales EOE and LST. Regarding discriminant validity, the Sensing the Subtle dimension exhibited clear distinctiveness, while the other three subscales overlapped with neuroticism, extraversion, empathy, and rumination, aligning with theoretical expectations. The DOES Scale emerges as a reliable, valid tool for assessing SPS, recommending its four dimensions be interpreted as a trait constellation.

ARTICLE HISTORY
Received 29 February 2024
Revised 20 August 2024
Accepted 4 September 2024

In an increasingly fast-paced world characterized by an abundance of sensory stimuli – from busy cities and hectic streets to the never-ending stream of digital notifications – some people feel a heightened sensitivity to their surroundings and experiences. As such, these individuals report perceiving their environment with a depth and richness that often eludes others, but at the same time, they are also more susceptible to temporary sensory overload and emotional exhaustion (Aron et al., 2012). Individual differences in people's perceived sensitivity to the environment are studied under the label of Sensory Processing Sensitivity (SPS) and were first described by Aron and Aron (1997). Since then, SPS has not only become a topic of considerable interest in research but has also gained a lot of popularity in society, as evidenced by numerous popular science books, self-help guides, online groups, and conferences (Roth et al., 2023).

Aron and Aron (1997) and Aron et al. (2012) initially defined SPS as a personality trait characterized by greater depth of information processing, increased emotional reactivity and empathy, greater awareness of environmental subtleties, and ease of overstimulation. Later, this definition was compiled by Aron (2020) into the acronym DOES, which represents the four core characteristics: D = Depth of processing, O = Overstimulation, E = Emotional Reactivity and Empathy, and S = Sensing the Subtle. Most recently, Roth et al. (2023) obtained empirical support for these four core characteristics in an interview study in which individuals who described themselves as high on SPS reported attributes pertaining to all four DOES components.

Nevertheless, the SPS construct has not remained unchallenged. Some researchers contend that fundamental questions remain unanswered, such as how SPS is distinct from established personality traits and how it can be reliably and validly measured (see Hellwig & Roth, 2021; Roth et al., 2023). The present research addresses commonly discussed psychometric issues regarding the by far most frequently used SPS questionnaire, the Highly Sensitive Person Scale (HSPS; Aron & Aron, 1997), and aims to overcome these issues by developing a new instrument to measure SPS in the sense of its previously specified DOES substructure.

The HSPS was initially developed based on interviews with people who described themselves as either "highly introverted" (preferring the company of one or two people rather than several) or "easily overwhelmed by stimulation" (e.g., by noisy places or evocative or shocking entertainment; Aron & Aron, 1997, p. 350). Both aspects are typical characteristics of the Big Five dimensions introversion/extraversion and neuroticism. Based on the interview responses, the authors developed the

CONTACT Danièle Anne Gubler daniele.gubler@unibe.ch Institute of Psychology, University of Bern, Fabrikstrasse 8, CH-3012 Bern, Switzerland.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/00223891.2024.2405536.

27-item HSPS. The HSPS has been translated and adapted into several languages (Bordarie et al., 2022; Chacón et al., 2021; Ershova et al., 2018; Grimen & Diseth, 2016; Konrad & Herzberg, 2017; Şengül-İnal & Sümer, 2020). Despite its widespread use, however, research on the psychometric properties of the HSPS revealed a number of significant shortcomings.

First, in the initial development of the HSPS, Aron and Aron (1997) originally posited a unidimensional factorial structure for their scale. Subsequent research, however, has yielded diverse multifactorial solutions (Evans & Rothbart, 2008; Lionetti et al., 2018; May et al., 2022; Smolewska et al., 2006). Among these studies, the three-factor solution proposed by Smolewska et al. (2006) has been the most consistently supported factorial structure (Greven et al., 2019; Grimen & Diseth, 2016; Konrad & Herzberg, 2017; Sobocko & Zelenski, 2015). The three components that were extracted from 25 of the 27 items by Smolewska et al. (2006) were labeled Ease of Excitation (EOE; twelve items measuring being overwhelmed due to external and internal demands), Low Sensory Threshold (LST; six items measuring unpleasant sensory arousal), and Aesthetic Sensitivity (AES; seven items measuring aesthetic awareness). Nevertheless, as these three components emerged empirically and were not originally based on theory, their interpretation is potentially problematic (Greven et al., 2019). Moreover, the uneven distribution of items on these subscales results in an imbalance. In particular, the EOE subscale is over-represented in calculating the total score, which lacks a theoretical justification.

Second, at the item level, the HSPS overlaps with at least four of the Big Five personality traits, reflecting a conceptual merging of SPS with established personality constructs (Lionetti et al., 2019). Specifically, items tap into behaviors and characteristics associated with introversion (e.g., "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?"), neuroticism (e.g., "Do you find it unpleasant to have a lot going on at once?"), and openness to experience (e.g., "Are you deeply moved by the arts or music?), while one item specifically asks about conscientiousness (e.g., "Are you conscientious?). This content overlap at the item level is also reflected in the medium to high correlations of the HSPS subscales with the Big Five, showing substantial overlap of EOE with neuroticism, as well as LST with neuroticism and introversion and AES with openness to experience (Lionetti et al., 2019; Roth et al., 2023). In a study examining these variables on a latent level (Hellwig & Roth, 2021), EOE was indistinguishable from the neuroticism facet of self-conscientiousness, LST exhibited high overlap with neuroticism and introversion, and AES was identical to openness to experience. The question, therefore, arises as to whether the construct measured by the HSPS reflects SPS as a fundamental personality trait (e.g., Pluess, 2015) or may be better described as a trait constellation of three well-known personality traits (high neuroticism, low extraversion, and high openness to experience; Bröhl et al., 2021).

Third, many of the HSPS items are negatively connotated, with Evans and Rothbart (2008) finding that 18 of the 27 items primarily contain negative affect. Although the DOES dimension Overstimulation mainly reflects negative aspects of SPS, greater Depth of Processing, Emotional Reactivity and Empathy, and Sensing the Subtle could also represent neutral or even positive characteristics. For example, sensing subtle nonverbal cues could represent an advantage in interpersonal communication (Palese & Mast, 2020). Therefore, negative item wording may affect the nature of the construct being measured and contribute to the high correlations with neuroticism.

Finally, the scale shows limitations in content validity. As noted by Aron and colleagues, Depth of Processing might manifest in deeper and longer cognitive processing of information and inhibited behavior ("pause to check in novel situations," Aron et al., 2012), which is not represented in the items. Moreover, Aron et al. (2012) acknowledged that the HSPS does not capture higher emotional reactivity – a feature that did emerge in their original interviews but was removed in favor of a shorter questionnaire when the HSPS was developed. This is consistent with Evans and Rothbart's (2008) notion that the questionnaire does not fully reflect Aron and Aron's (1997) theory. Thus, the HSPS fails to capture two of the four core characteristics proposed in the DOES model of Aron (2020). Consequently, Aron and Aron's theory of an overarching unidimensional SPS factor cannot be conclusively investigated, given that the HSPS only measures part of the overall construct.

The reason why the HSPS is widely used despite its shortcomings can be explained by the lack of alternative scales. Newer scales, such as the Sensory Processing Sensitivity Questionnaire (SPSQ) developed by De Gucht et al. (2022) or the HSP-Test for Highly Sensitive Persons by Satow (2022), have addressed some of these existing problems. While these newer questionnaires offer a more balanced measurement of SPS and represent an improvement in the operationalization of the construct, none of them fully represents Aron's four postulated core characteristics of the DOES model. Consequently, a questionnaire that captures all four dimensions of the DOES model in a balanced and reliable way is still missing. Given these limitations and the ongoing methodological and conceptual debates surrounding the SPS construct (Hellwig & Roth, 2021), we propose to take a step back and develop a new scale based on an inductive approach that relies on qualitative data from individuals who identify with SPS. This approach aims to equally and accurately capture the four core characteristics postulated by Aron (2020). Such a questionnaire would provide theory-based insights for individuals with high SPS and allow for a more comprehensive understanding of their experiences and behaviors across the different dimensions of SPS. This development is critical to advance research and improve diagnostic accuracy and personalized interventions for individuals with high SPS.

The present studies

The goal of the present research was to develop a new scale to measure SPS that appropriately reflects all four theoretical characteristics of SPS according to the DOES model, aligning with the current state of research (Aron, 2020; Roth et al., 2023) and addressing the above-outlined shortcomings of older questionnaires measuring SPS. Thus, we aimed to

minimize overlap with the Big Five and acknowledge the potential neutral or positive facets of SPS while formulating the items. As a basis for item formulation, we used the interview transcripts and the results from the coding system from Roth et al.'s study (2023). Unlike Aron and Aron's (1997) interview study that focused on people who described themselves as introverted and/or high in neuroticism, Roth et al. (2023) conducted semi-structured interviews with individuals who self-identified as "highly sensitive", taking advantage of the growing awareness of the SPS construct in public. In these interviews, participants were asked, among other things, about their definitions of SPS, how it manifests in their everyday behaviors, thoughts, etc., and how it affects their lives. These statements were already sorted in the study by Roth et al. (2023) and categorized based on the number of frequencies mentioned.

To this end, we extracted all statements that reflected thoughts, experiences, behaviors, and feelings related to SPS. These statements were gradually reduced, grouped, and combined into superordinate categories. In the process, care was taken to ensure that the categories could be clearly distinguished from each other and that they appeared in most of the interview statements. Based on this qualitative analysis, we identified four dimensions that closely aligned with the DOES characteristics proposed by Aron (2020). This confirmed the suitability of the interview transcripts as a basis for item generation for the new questionnaire. In line with Aron et al.'s (2012) and Aron's (2020) definitions, the new questionnaire was developed to cover the following dimensions: Sensing the Subtle describes a lower perceptual threshold for stimuli with a more unfiltered perception of stimuli. Emotional Reactivity involves feeling and experiencing moments profoundly and strongly, evident in any experience, such as nature-related, artistic, or interpersonal interactions. Overstimulation refers to the tendency to be rapidly inundated and exhausted by stimuli and experiences, whereas Depth of Processing refers to thinking deeply, thoroughly, and intensely about experiences. The statements extracted from the interviews on these four critical characteristics of SPS formed the starting point for developing the new scale, which we called the DOES Scale.

Three studies across diverse adult community samples in Switzerland, Germany, Austria, and the UK were conducted to evaluate the scale's psychometric properties. In Study 1, we started with an item pool of 28 items (7 per subscale) and compared different factor models to describe the data. This investigation resulted in a final 20-item version with five items covering each aspect of the core SPS characteristics. In Study 2, we modified one item of the Emotional Reactivity subscale and validated the revised scale in a German-speaking sample and its translation in an English-speaking sample. For the German scale, we further examined test-retest reliability at a 30-day interval and determined convergent validity with the HSPS and discriminant validity with different personality measures. For the English scale, we determined convergent validity with the HSPS and measurement invariance with the German scale. Finally, in Study 3, we confirmed the factorial structure of the DOES Scale in two different samples: first, in individuals who self-identified with high SPS, and second, in individuals who did not self-identify with high SPS.

Study 1

Method

Development of the DOES Scale

Item development was conducted in the context of a master-level seminar at the University of Bern. Students and the study authors created a first item pool by reformulating the statements extracted from the interview contents of Roth et al.'s (2023) study, applying the following criteria: The items should cover neutral, positive, as well as negative aspects of the four DOES dimensions, their formulation should be on a similar level of abstraction, and content overlap between the four dimensions should be minimized. Specifically, for the Sensing the Subtle subscale, items encompassing all sensory modalities, including hearing, smelling or tasting, visual experiences, and bodily sensations, should be included. The aim was to neutrally describe a lower perceptual threshold across different senses without addressing Overstimulation (e.g., "During a walk, I notice sounds around me clearly."). For the Emotional Reactivity subscale, items attempted to describe profound experiences of positive and negative events in the realms of art, nature, interpersonal interactions, and other people's emotions. Importantly, items should focus on the experiential aspect, avoiding inquiries about actively seeking these encounters in order to minimize potential overlap with openness to experience. An example item is, "When I listen to beautiful music, I can become completely absorbed in it." For the Overstimulation subscale, negatively valenced phrasing was permitted, as the interviewees had described this aspect of SPS as a negative characteristic of being highly sensitive. The items represented overstimulation of sensory modalities (e.g., "I am very disturbed by sounds that occur at the same time.") as well as emotional overstimulation caused by social interactions (e.g., "I often feel exhausted after being out with a lot of people."). Finally, for the Depth of Processing subscale, we aimed to create items that distinguished profound thinking from mere rumination. The items attempted to reflect broad cognitive absorption, encompassing a wide range of thoughts and experiences, to ensure a clear distinction from repetitive or negative patterns of thinking (e.g., "It often happens to me that I lose myself for hours in new ideas."). The statements developed went through multiple rounds of discussion and revision. This process resulted in 28 items (7 items per subscale). A four-point response scale was chosen ranging from 1 ("strongly disagree"), 2 ("disagree"), 3 ("agree") to 4 ("strongly agree").

Participants

Participants were recruited through the social networks of students participating in the master-level seminar mentioned above. A total of 471 individuals fully completed the online study. They ranged in age from 18 to 65 years, with a mean age of 33.0 (± 13.6) years. Among these, 282

(59.9%) participants were women, 186 (39.5%) were men, and three (0.6%) were non-gender specific. Regarding educational level, three (0.6%) individuals did not finish mandatory school, 12 (2.6%) finished mandatory school, 89 (18.9%) finished an apprenticeship, 75 (15.9%) finished technical college, 276 (58.6%) finished high school or higher education, and 16 (3.4%) specified a different educational path. The protocol for this study, as well as Studies 2 and 3, was approved by the local ethics committee of the Faculty of Human Sciences of the University of Bern (Nr. 2022-11-05). The survey was administered online through Qualtrics. In this study, as well as in Studies 2 and 3, participants were initially informed that the studies aimed to investigate the relationship between various personality traits without mentioning SPS. Detailed feedback about the studies' true purpose was provided only in the debriefing at the end of the survey.

Data analysis

Statistical analyses were conducted using RStudio version 4.3.3. To assess the factorial structure, a CFA was performed using the lavaan package (Rosseel, 2012). Diagonally weighted least squares estimation (DWLS), which is recommended for categorical data, was employed (Mindrila, 2010). Unidimensional models were individually tested for each subscale before assessing the overall factorial structure of the scale. These models were evaluated based on several fit indices (see below), the magnitude of their factor loadings, and potential high residual correlations within each subscale. Subsequently, residual correlations among items from different subscales were examined. With these analyses, we aimed to identify and exclude items with low factor loadings and high residual correlations. However, when excluding items, we also made sure that the full intended content range for each subscale was maintained.

After item exclusion, the factorial structure of the resulting overall scale was tested for three different models: a one-factor model, where all items loaded on a single factor; a four-factor model, where items were organized into four separate and correlated factors according to their respective, a priori-defined subscale; and a hierarchical model, where the four factors loaded on a second-order factor. These models were compared to determine the best-fitting structure for the items.

Model/data fit was evaluated using the comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Good model fit was indicated by CFI values ≥ .95, RMSEA < .05, and SRMR < .10, while CFI ≥ .90 and RMSEA < .08 suggested an acceptable fit (Schweizer, 2010). We further calculated the chi-square statistics for completeness but did not interpret them due to their dependence on sample size (Hu & Bentler, 1999). The chi-square difference test was used to compare the different overall models. As this test suffers from the same weaknesses as the chi-square statistics (sensitive to sample size), we further calculated omega hierarchical subscale (omegaHS, ω_{hs}) estimates in the second-order model, which is an effect size indicator of unique latent variable strength of the first-order factors and can be calculated as the ratio of first-order variance to total variance (Gignac & Kretzschmar, 2017; Reise et al., 2012). Simply put, omegaHS values become larger when the factor loadings of the first-order factor on the higher second-order factor become smaller, indicating the first-order latent factor's unique strength. These effect sizes were interpreted following Gignac and Kretzschmar (2017), with ω_{hs} < 0.20, a relatively small effect; ω_{hs} = 0.20 to 0.30, a moderate effect; and ω_{hs} > 0.30, a relatively large effect. Thus, values more than ω_{hs} > 0.30 indicate that the first-order factors can be interpreted as unique dimensions. To assess the reliability of the four subscales, ordinal α and ordinal ω were computed (Flora, 2020; Zumbo et al., 2007).

A model-free power analysis was conducted using the semPower package (Moshagen & Bader, 2023) to determine the sample size for the global model fit criteria. The required sample size to detect an effect of RMSEA = .08, with a power of 80% and an alpha error level of .05 for the unidimensional models with 14 degrees of freedom (one latent factor derived from seven manifest items), yielded a required sample of $N = 206$. For a unidimensional model with 5 degrees of freedom (one latent factor derived from five manifest items), the required sample size to detect an effect of RMSEA = .08, with a power of 80% and an alpha error level of .05, yielded a required sample of $N = 472$. For the higher order models, the required sample size to detect an effect of RMSEA = .08, with a power of 80% and an alpha error level of .05 for a model with 164 degrees of freedom (see Table 1), yielded a required sample of $N = 50$. For an effect of RMSEA = .05, the required sample was $N = 125$. Despite these calculations, it is recommended to adhere to the guideline of using a minimum sample size of $N = 200$ when employing DWLS estimators with ordinal data (Kyziazos, 2018).

Results and discussion

First, fit indices for each subscale were evaluated separately (see Table 1). For the Sensing the Subtle subscale, fit indices of the seven items showed an acceptable to good model fit and factor loadings were above λ = .500, with the exception of one item with a smaller factor loading of λ =.316. This item ("I react to certain foods, such as coffee or sugar, even in small amounts.") showed significantly lower associations with the other items, so it was discarded. Furthermore, two items exhibited high residual correlations with a modification index of χ^2 = 48.048. As these items represented smell and taste perception, we decided to remove the taste item ("I can easily recognize the different ingredients and flavors in a dish," λ =.542) and keep the smell item ("I can perceive smells very clearly," λ = .736) as smell perception was mentioned more often in the interviews than taste (see Roth et al., 2023). This process resulted in a 5-item subscale that showed a good model fit (see Table 1).

For the Emotional Reactivity subscale, fit indices of the seven items revealed an acceptable to poor model fit with factor loadings above λ = .400, except for one item with λ

Table 1. Fit indices and reliability estimates for each subscale of the DOES Scale, as well as their combination into higher-order models in Study 1 and Study 2.

Model	*N*	χ^2	*df*	*p* value	CFI	RMSEA	SRMR	ordinal α	ordinal ω	r_{tt}
Study 1 – German-speaking sample										
Sensing the Subtle (7 items)	471	70.353	14	<. 001	.958	.093	.067	.76	.73	
Sensing the Subtle (5 items)	471	12.129	5	.033	.992	.055	.038	.76	.72	
Emotional Reactivity (7 items)	471	76.301	14	<. 001	.867	.097	.080	.65	.61	
Emotional Reactivity (5 items)	471	23.319	5	<. 001	.937	.088	.066	.61	.57	
Overstimulation (7 items)	471	42.536	14	<. 001	.989	.066	.054	.82	.79	
Overstimulation (5 items)	471	3.941	5	.558	1.000	.000	.023	.80	.79	
Depth of Processing (7 items)	471	255.569	14	<. 001	.897	.192	.125	.80	.79	
Depth of Processing (5 items)	471	6.042	5	.302	.999	.021	.027	.81	.76	
20-item Model 1 (one-factor model)	471	1041.757	170	<. 001	.897	.104	.101	.86	.85	
20-item Model 2 (four-factor model with correlated factors)	471	414.759	164	<. 001	.970	.057	.067			
20-item Model 3 (second-order model)	471	437.961	166	<. 001	.968	.059	.070			
Study 2 – German-speaking sample										
Sensing the Subtle (5 items)	242	6.159	5	.291	.998	.031	.036	.77	.73	.82
Emotional Reactivity (5 items)	242	4.931	5	.424	1.000	.000	.037	.71	.67	.77
Overstimulation (5 items)	242	2.726	5	.742	1.000	.000	.025	.85	.82	.89
Depth of Processing (5 items)	242	3.414	5	.636	1.000	.000	.029	.81	.76	.86
20-item Model (one-factor model)	242	739.500	170	<. 001	.898	.118	.118	.88	.87	.90
20-item Model (four-factor model with correlated factors)	242	308.453	164	<. 001	.974	.060	.079			
20-item Model (second-order model)	242	334.120	166	<. 001	.970	.065	.083			
Study 2 – English-speaking sample										
Sensing the Subtle (5 items)	232	7.303	5	.199	.992	.045	.046	.72	.66	
Emotional Reactivity (5 items)	232	16.666	5	.005	.968	.101	.069	.72	.69	
Overstimulation (5 items)	232	2.210	5	.819	1.000	.000	.026	.83	.81	
Depth of Processing (5 items)	232	12.081	5	.034	.991	.078	.050	.81	.77	
20-item Model (one-factor model)	232	431.385	170	<. 001	.959	.082	.092	.90	.89	
20-item Model (four-factor model with correlated factors)	232	269.212	164	<. 001	.984	.053	.075			
20-item Model (second-order model)	232	279.620	166	<. 001	.982	.054	.076			

Note. CFI = comparative fit index, RMSEA = root-mean-square error of approximation, SRMR = standardized root-mean-square residual, r_{tt} = retest reliability over a one-month interval.

=.340. This item ("I often lose track of time when I'm doing something interesting.") was discarded. To further improve model fit, we decided to exclude another item as it described emotional reactivity rather abstractly ("I perceive my mood intensely and clearly."), and, at the same time, had a relatively low factor loading, λ =.417, compared to the other items. The resulting 5-item subscale showed an acceptable model fit (except for RMSEA, which exceeded the threshold of .08).

For the Overstimulation subscale, fit indices of the seven items showed a good model fit with all factor loadings being above λ = .400. To ensure that all subscales were equally represented in the final scale, we decided to exclude two items. One item ("I am very susceptible to bright or glaring light," λ =.572) was excluded due to a high residual correlation with an item from the Sensing the Subtle subscale, and another item, which had the lowest factor loading ("There are materials whose touch I can hardly tolerate," λ =.407). The resulting 5-item subscale showed a good model fit.

For the Depth of Processing subscale, fit indices of the seven items revealed a poor model fit, with all factor loadings being above λ = .500. For this subscale, two items ("If something puzzles me, I keep looking into it until I understand it," λ =.611, and "I tend to get to the bottom of things," λ = .668) had very high residual correlations with each other with a modification index of χ^2 = 226.545. Compared to the other five items, these two items also tapped more into an intrinsic need for cognition. Based on these considerations, they were excluded. The resulting 5-item subscale showed a good model fit.

In summary, two items were removed from each subscale, resulting in four final subscales with five items each. These subscales demonstrated acceptable to good model fit and comprehensively covered the four DOES domains. Regarding connotation, the Emotional Reactivity and the Depth of Processing subscales contained only one item with an exclusively negative connotation, while the other items were specifically and carefully formulated to include neutral and positive aspects (see Table 2). The Sensing the Subtle subscale contained entirely neutral formulations. Consequently, with the five items of the Overstimulation subscale, only seven of the 20 items were negatively worded across the entire DOES Scale. Compared to the HSPS, which predominantly features negatively worded items, the DOES Scale captures neutral, positive, and negative aspects of SPS in a more balanced way.

In a next step, the fit indices for these final 20 items were evaluated for three different models (see Table 1): A one-factor model, a four-factor model with correlated factors, and a hierarchical four-factor model with a second-order factor to represent the overarching construct of SPS. Whereas the one-factor model did not meet the criteria for acceptable model fit, the four-factor model with correlated factors demonstrated good model fit. Correlations between the four subscales were all significantly positive (all *ps* < .001): *r* =. 406 (Sensing the Subtle and Emotional Reactivity), *r* =. 457 (Sensing the Subtle and Overstimulation), *r* =. 389 (Sensing the Subtle and Depth of Processing), *r* =. 539 (Emotional Reactivity and Overstimulation), *r* =.727 (Emotional Reactivity and Depth of Processing), and *r* =.716 (Overstimulation and Depth of Processing). These

significant intercorrelations justified examining a second-order model. The second-order model also revealed a good model fit. However, the chi-square difference test indicated that the second-order model described the data significantly worse than the model with four correlated factors, $\Delta\chi^2(2) = 23.202$, $p < .001$. OmegaHS values of the four subscales in the second-order model were as follows: Sensing the Subtle, $\omega_{hs} = .567$; Emotional Reactivity, $\omega_{hs} = .258$; Overstimulation, $\omega_{hs} = .299$; and Depth of Processing, $\omega_{hs} = .165$, suggesting that the Sensing the Subtle subscale exhibited a large portion of unique variance not captured by the second-order factor. The other three subscales exhibited unique variance with small to moderate effect sizes. Overall, the results suggest that the four-factor model with correlated factors more accurately depicts the underlying structure of SPS as measured by the DOES Scale compared to the second-order model. This finding challenges the notion of an overarching unidimensional SPS factor underlying the DOES categories, with Sensing the Subtle representing the most unique dimension.

The 20 items of the DOES Scale, their means, standard deviations, skewness, and kurtosis are provided in Table 2. Most items showed moderate to low difficulty (i.e., high agreement toward the item) with symmetric to left-skewed distributions. Only Item 6 from the Emotional Reactivity subscale exhibited very low difficulty and left-skewed distribution with a high kurtosis. The reliabilities of the four subscales are shown in Table 1. Ordinal alpha and ordinal omega were acceptable to good for all subscales except for the Emotional Reactivity subscale, where the RMSEA also exceeded the threshold of .08 as mentioned above. One item from this subscale (Item 6) was modified in Study 2 due to its suboptimal item statistics as well as the poor reliability and model fit of the subscale.

Study 2

We pursued four objectives with Study 2. First, we aimed to improve the quality of the newly developed DOES Scale by revising Item 6. Second, we examined the factorial structure and psychometric properties of the final 20 selected items when presented without the eight items removed in Study 1. This included the examination of the model fit of all subscales and their combination into combined models as in Study 1 and the investigation of the reliability of the subscales in terms of their internal consistency and test-retest reliability. Third, to make the new scale available to the international research community on SPS, we translated the items into English. The English version of the scale was examined regarding its psychometric properties in the same way as the German version. We further tested both scales (German vs. English version) for measurement invariance to determine the extent to which the measurement models hold across both language versions. Finally, we determined the convergent validity of the DOES Scale with Aron and Aron's (1997) HSPS and the discriminant validity with neuroticism, extraversion, openness to experience, empathy, rumination, and the behavioral inhibition and activation systems (BIS/BAS). For these constructs, substantial correlations with the HSPS have been reported in previous studies, especially with the Big Five personality traits. Due to the content overlap of the HSPS with the DOES subscales, we tested whether the newly developed subscales could be dissociated from these constructs or should be better understood as variants of these well-established constructs. For example, we examined whether the relationship between openness to experience and the DOES subscales would be less pronounced than with the HSPS. We further explored the extent to which Overstimulation can be distinguished from neuroticism and the BIS system, Emotional Reactivity from empathy, and Depth of Processing from rumination. Lastly, we investigated how Sensing the Subtle maps into all of these constructs.

Method

Participants and procedures

The study comprised two samples with a total of 474 participants recruited online through Prolific (https://www.prolific.com). The first sample initially involved 250 respondents from Germany, Switzerland, and Austria who were paid £3.00 for participation. Of these, eight individuals were excluded due to incorrect responses to control items. The remaining 242 participants (116 women, 123 men, 3 non-gender specific) had a mean age of 32.7 years ($SD = 9.6$). Of these, 101 participants took part in the survey twice, 30 days apart, to determine the test-retest reliability of the new DOES Scale. These participants were additionally paid £1.00. The second sample originally comprised 240 native-English-speaking participants from the UK who were paid £0.80 to fill out the DOES Scale and the HSPS in English. Eight of them were excluded due to conspicuous response behavior (very fast response behavior or no variance or very high variance in response behavior). The resulting 232 participants (119 women; 113 men) had a mean age of 38.3 years ($SD = 11.0$).

Measures

DOES Scale. For the German sample, the items selected in Study 1 were used, with the exception of Item 6 ("I enjoy deep conversations very much.") which was changed to "I'm intensely moved by deep conversations." The phrase "intensely moved" implies a strong emotional response, not merely enjoyment, associated with deep conversations, which better aligns with the emotional reactivity described by the interviewees in Roth et al. (2023). The final German DOES Scale is available in the supplementary material.

For the English sample, the twenty items of the DOES Scale were translated following the ITC Guidelines for Translating and Adapting Tests (International Test Commission, 2017). Initially, a native English speaker with excellent knowledge of German translated the items into English. Subsequently, a native German speaker, proficient in

Table 2. Means, Standard Deviations, Skewness s, and Kurtosis κ in the German-speaking sample of Study 1 as well as the German-speaking and English-speaking sample of Study 2.

	Study 1 – German-speaking sample ($N = 471$)				Study 2 – German-speaking sample ($N = 242$)				Study 2 – English-speaking sample ($N = 232$)			
Items	*M*	*SD*	*S*	κ	*M*	*SD*	*s*	κ	*M*	*SD*	*s*	κ
1. During a walk, I notice sounds around me clearly. (S)	3.14	0.82	−0.68	−0.15	3.14	0.65	−0.32	0.07	3.20	0.57	−0.29	1.09
2. The moods of other people are highly contagious to me. (E)	2.64	0.87	−0.13	−0.69	2.80	0.79	−0.03	−0.72	2.86	0.74	−0.09	−0.56
3. Events where I am exposed to multiple impressions overwhelm me. (O)	2.23	0.96	0.28	−0.90	2.66	1.01	−0.23	−1.05	2.69	0.85	0.12	−0.86
4. It is difficult for me to switch off mentally. (D)	2.89	0.97	−0.52	−0.69	2.94	0.91	−0.50	−0.60	2.90	0.82	−0.24	−0.66
5. I clearly perceive the kind of textiles I wear on my skin. (S)	2.33	1.02	0.18	−1.10	2.46	0.89	0.09	−0.75	2.62	0.79	−0.09	−0.43
6. I am intensely moved by deep conversations.* (E)	3.55	0.68	−1.57	2.36	3.17	0.73	−0.52	−0.21	2.89	0.71	−0.06	−0.52
7. I often feel exhausted after being out with a lot of people. (O)	2.83	1.00	−0.39	−0.95	3.14	0.94	−0.86	−0.23	3.16	0.81	−0.54	−0.62
8. I have intense thoughts about situations that I will experience in the future. (D)	2.87	0.95	−0.41	−0.81	3.01	0.86	−0.52	−0.49	2.78	0.81	−0.12	−0.63
9. I am aware when the light in a room is glaring or bright. (S)	3.17	0.86	−0.75	−0.29	3.36	0.78	−1.12	0.79	3.26	0.72	−0.63	−0.10
10. When others quarrel, it greatly affects me. (E)	2.87	0.88	−0.31	−0.74	2.79	0.81	−0.15	−0.59	2.86	0.78	−0.08	−0.74
11. When there is a lot going on around me, I am easily overwhelmed. (O)	2.36	0.94	0.17	−0.87	2.70	0.94	−0.10	−0.96	2.84	0.85	−0.07	−0.97
12. After experiencing something positive or negative, I think about what happened for a long time. (D)	3.15	0.75	−0.53	−0.29	3.13	0.72	−0.47	−0.10	3.11	0.70	−0.45	−0.04
13. I notice slight differences in temperature quickly. (S)	2.45	0.94	0.07	−0.90	2.59	0.86	0.01	−0.68	2.73	0.73	0.20	−0.71
14. Emotional film scenes touch me deeply. (E)	3.19	0.79	−0.73	0.02	2.88	0.88	−0.39	−0.60	2.96	0.80	−0.33	−0.51
15. I am very disturbed by sounds that occur at the same time. (O)	2.85	0.96	−0.35	−0.91	2.90	0.93	−0.40	−0.78	2.28	0.84	0.39	−0.37
16. I spend a lot of time thinking about everyday experiences. (D)	2.45	0.93	0.15	−0.85	2.60	0.83	0.12	−0.68	2.79	0.71	−0.03	−0.44
17. I can perceive smells very clearly. (S)	2.91	0.84	−0.37	−0.51	2.86	0.87	−0.36	−0.60	2.80	0.79	−0.22	−0.41
18. When I listen to beautiful music, I can become completely absorbed in it. (E)	3.31	0.79	−0.95	0.24	3.23	0.75	−0.69	−0.03	3.03	0.75	−0.42	−0.17
19. I often find odors like perfume or strong food smells unbearable. (O)	2.36	1.01	0.22	−1.04	2.33	1.01	0.26	−1.03	2.38	0.80	0.19	−0.41
20. It often happens to me that I lose myself for hours in new ideas. (D)	2.51	0.92	0.11	−0.68	2.61	0.96	−0.11	−0.95	2.58	0.77	0.05	−0.45

Note. * Modified item that was formulated as "I enjoy deep conversations very much" in Study 1.

English, conducted a back-translation of the questionnaire. In the final step, the original German, translated English, and back-translated German versions were compared and any discrepancies were resolved to ensure that the English items accurately represented the meaning of the original German items. The 20 English items are available in Table 2 and the supplementary material.

Highly Sensitive Person Scale (HSPS). The German version of the Highly Sensitive Person Scale (HSPS), introduced by Konrad and Herzberg (2017), was administered to the German-speaking sample and comprised three factors: Ease of Excitation (EOE; 10 items, e.g., "Do you get rattled when you have a lot to do in a short amount of time?"), Aesthetic Sensitivity (AES; 5 items, e.g., "Are you deeply moved by the arts or music?"), and Low Sensory Threshold (LST; 11 items, e.g., "Are you easily overwhelmed by things like bright lights, strong smells, coarse fabrics, or sirens close by?). Participants rated their responses on a 5-point response scale, ranging from "strongly disagree" (0) to "strongly agree" (4). The English version of the HSPS, developed by Aron and Aron (1997), was used for the English sample. The English version comprised 27 items, of which 25 items can be used to extract the three subscales: EOE (12 items), AES (7 items), and LST (6 items), according to Smolewska et al. (2006). Items were answered on a seven-point response scale ranging from "not at all" (1) to "extremely" (7).

HEXACO. The German version of the HEXACO-60 (Moshagen et al., 2014) was used to assess the personality dimensions of honesty-humility (e.g., "I wouldn't use flattery to get a raise or promotion at work, even if I thought it would succeed."), emotionality (neuroticism, e.g., "I sometimes can't help worrying about little things."), extraversion (e.g., "In social situations, I'm usually the one who makes the first move."), agreeableness (e.g., "I rarely hold a grudge, even against people who have badly wronged me."), conscientiousness (e.g., "I always try to be accurate in my work, even at the expense of time."), and openness to experience (e.g., "If I had the opportunity, I would like to attend a classical music concert."). The 10 items of each scale were rated on a five-point response scale ranging from "strongly disagree" (1) to "strongly agree" (5).

Toronto Empathy Questionnaire. To measure empathy, a German version (Janelt et al., 2024) of the Toronto Empathy Questionnaire (TEQ) was used (Spreng et al., 2009). The TEQ comprises a total of 16 items (e.g. "I can tell when others are sad even when they do not say anything.") to be answered on a five-point response scale ranging from "never" (1) to "always" (5).

Rumination-Reflection Questionnaire. Rumination was assessed using König's (2012) German version of the

Rumination-Reflection Questionnaire (RRQ, Trapnell & Campbell, 1999). The scale consists of 12 items (e.g., "Sometimes it is difficult for me to turn off thoughts about myself."), which were answered on a five-point response scale ranging from "do not agree at all" (1) to "fully agree" (5).

BIS/BAS Scale. To measure individual differences in the Behavioral Inhibition System (BIS) and the Behavioral Approach System (BAS), a German adaptation of Carver and White's (1994) BIS/BAS scales was used (Strobel et al., 2001). The BIS Scale consists of seven items (e.g., "Criticism or scolding hurts me quite a bit."), measuring the anticipation of or sensitivity to negative experiences. The BAS Scale consists of 13 items (e.g., "When I want something, I usually go all-out to get it."), measuring drive, reward responsiveness, and fun-seeking. Answers were given on a four-point response scale ranging from "does not apply to me at all" (1) to "applies exactly to me" (4).

Statistical analysis

The factorial structure of both the German and the English versions of the DOES Scale was tested by means of CFA using the same software, estimation method, fit statistics, and criteria for good/acceptable model fit as in Study 1. Analyses on the subscale and the overall scale levels followed the same steps as described in Study 1. Convergent and discriminant validity were assessed through zero-order correlations of the DOES subscales with the subscales and/or total scores of all other questionnaires (see Table 3 for descriptive statistics and internal consistency of all variables). The measurement invariance between the two language versions of the DOES Scale was tested up to scalar invariance to determine whether the relationships of the subscales to other scales (metric invariance) or the mean differences (scalar invariance) in the subscales between the two language versions can be interpreted in terms of the underlying constructs. Invariance analyses were based on DWLS estimations, and model fit was evaluated with difference scores like ΔCFI (≤ .010) and ΔRMSEA (≤ .015; Chen, 2007).

For Study 2, a power analysis was conducted regarding the correlation analyses using G*Power (Faul et al., 2007). To detect a small to medium effect size of r = .20 with an alpha error of .05 and a power of 80%, a sample of N=193 is required.

Results and discussion

Descriptive statistics for the DOES Scale items in German and English are provided in Table 2. Fit indices and reliability estimates for both versions are shown in Table 1. In the German version, all subscales demonstrated a good model fit, indicating each subscale's unidimensionality. Modifying Item 6 improved the Emotional Reactivity subscale from an acceptable model fit in Study 1 to a good model fit in Study 2. The internal consistencies of the four subscales were acceptable to good. Furthermore, the test-retest reliability, which was measured at a one-month interval, was acceptable to good for all four subscales, indicating that the four subscales capture relatively stable traits.

Good model fits were also found in the English version of the DOES Scale for the Sensing the Subtle, the Overstimulation, and the Depth of Processing subscales. For the Emotional Reactivity subscale, CFI and SRMR suggested a good model fit, while the RMSEA exceeded the threshold

Table 3. Descriptive Statistics and Cronbach's Alpha of test scores from all Measures used in Study 2 for the German-speaking and the English-speaking sample.

Measures	Mean	SD	Min-Max	Skewness	Kurtosis	Cronbach's α
German-speaking sample N=242						
1.Sensing the Subtle (1–4)	2.88	0.56	1.20–4.00	−0.38	−0.02	.77*
2.Emotional Reactivity (1–4)	2.97	0.52	1.40–4.00	−0.24	−0.46	.71*
3.Overstimulation (1–4)	2.75	0.72	1.00–4.00	−0.29	−0.41	.85*
4.Depth of Processing (1–4)	2.86	0.61	1.00–4.00	−0.31	−0.38	.81*
5.HSPS Total Score (0–4)	2.05	0.76	0.35–3.81	0.02	−0.57	.94
6.EOE Score (0–4)	2.22	0.86	0.00–3.90	−0.40	−0.60	.90
7.LST Score (0–4)	1.69	1.01	0.00–4.00	0.27	−0.73	.94
8.AES Score (0–4)	2.47	0.74	0.40–4.00	−0.36	−0.37	.71
9.Honesty/Humility (1–5)	3.38	0.66	1.20–4.90	−0.20	−0.01	.77
10.Emotionality (Neuroticism) (1–5)	3.21	0.67	1.50–4.80	−0.08	−0.40	.80
11.Extraversion (1–5)	2.91	0.73	1.10–5.00	−0.04	−0.32	.85
12.Agreeableness (1–5)	3.16	0.59	1.30–4.60	−0.17	−0.23	.77
13.Conscientiousness (1–5)	3.56	0.57	2.20–4.80	−0.13	−0.49	.77
14.Openness to experiences (1–5)	3.57	0.63	2.00–4.90	−0.24	−0.69	.74
15.Empathy (1–5)	3.59	0.55	1.81–4.94	−0.27	−0.27	.88
16.Rumination (1–5)	3.64	0.78	1.25–5.00	−0.41	−0.41	.91
17.BIS Score (1–4)	3.06	0.64	1.14–4.00	−0.44	−0.50	.89
18.BAS Score (1–4)	2.99	0.40	1.92–4.00	−0.30	−0.16	.79
English-speaking sample N=232						
19.Sensing the Subtle (1–4)	2.92	0.47	1.40–4.00	−0.05	−0.12	.72*
20.Emotional Reactivity (1–4)	2.92	0.50	1.80–4.00	0.06	−0.45	.72*
21.Overstimulation (1–4)	2.67	0.61	1.20–4.00	0.01	−0.68	.83*
22.Depth of Processing (1–4)	2.83	0.55	1.40–4.00	0.08	−0.56	.81*
23.HSPS Total Score (1–7)	4.32	0.96	1.56–6.44	−0.04	−0.40	.93
24.EOE Score (1–7)	4.62	1.05	1.83–6.92	−0.14	−0.49	.88
25.LST Score (1–7)	3.54	1.24	1.00–6.83	0.04	−0.60	.81
26.AES Score (1–7)	4.54	0.95	1.71-7.00	−0.03	−0.18	.74

Note. *Ordinal alphas calculated from CFA.

of .08 (RMSEA = .101). However, the RMSEA improved to a good model fit (.078) when the residual correlation between Items 2 and 10 was allowed, both of which capture aspects of intense experience in interpersonal interactions. Overall, the results of Study 1 regarding psychometric properties were replicated in both the German and English versions of the DOES Scale.

Fit indices for the three different CFA models on the structure of the DOES Scale are shown in Table 1, separately for the German and English samples. For the German scale, the one-factor model did not describe the data well according to all fit indices. The four-factor model with correlated factors and the four-factor model with a second-order factor demonstrated good model fits. However, as in Study 1, the four-factor model with correlated factors described the data significantly better than the second-order model according to the chi-square difference test, $\Delta\chi^2(2) = 25.674$, $p < .001$. In the second-order model, the omegaHS values were as follows: Sensing the Subtle, $\omega_{hs} = .547$; Emotional Reactivity, $\omega_{hs} = .453$; Overstimulation, $\omega_{hs} = .313$; and Depth of Processing, $\omega_{hs} = .092$, indicating that except for the Depth of Processing subscale, all other subscales exhibited unique variance with a large effect not captured by the second-order factor. These results align with the finding of Study 1 and support the assumption that four correlated factors are better suited to describe the data than one overarching factor.

For the English version of the DOES Scale, we found a similar pattern. While the four-factor model with correlated factors and the four-factor model with a second-order factor exhibited good model fit according to all indices, the one-factor model did not meet the criteria for an acceptable model fit. The chi-square difference test indicated that the four-factor model with correlated factors described the data significantly better than the second-order factor model, $\Delta\chi^2(2) = 10.408$, $p = .005$. For this second-order model, again, the Sensing the Subtle scale exhibited unique variance with a large effect, $\omega_{hs} = .414$, while the other three subscales only exhibited a small amount of unique variance: Emotional Reactivity, $\omega_{hs} = .198$; Overstimulation, $\omega_{hs} = .187$; and Depth of Processing, $\omega_{hs} = .144$.

To sum up, across three independent samples, results revealed that four correlated factors best represented the underlying structure of the DOES Scale. Although the model fit of a second-order model was good across all three samples, it described the data significantly worse than the correlated four-factor model. This result was mainly driven by the Sensing the Subtle subscale in all three samples, which exhibited a large amount of unique variance not captured by the second-order factor. For this reason, analyzing the four subscales separately provides more accurate information about SPS than combining them into a second-order factor, although all latent factors were significantly positively correlated in the German and the English samples ($p < .001$; see Figure 1). The results of the measurement invariance analyses are shown in Table 4. The fit indices indicated metric invariance and partial scalar invariance. This result suggests that the relationships of the subscales to other scales and potential mean differences in test scores across both language versions can be interpreted and attributed to the underlying constructs.

Regarding the convergent validity of the DOES subscales with the HSPS, moderate to strong positive associations with the HSPS total score and its three subscales (EOE, LST, and AES) emerged in both the German and the English versions (see Table 5). Notably, among the four DOES subscales, the Overstimulation subscale revealed the highest correlation with the HSPS total score, EOE, and LST, corroborating the idea that the HSPS and its subscales mainly record SPS *via* the concept of overstimulation and thus predominantly capture the negative aspects of SPS.

In terms of discriminant validity (see Table 5), the Overstimulation and Depth of Processing subscales exhibited very similar correlation patterns. They were positively related, with moderate to strong effect sizes, to the emotionality (neuroticism) subscale of the HEXACO, the rumination scale, and the BIS scale. Furthermore, they were negatively related, also with moderate effect sizes, to the extraversion subscale of the HEXACO. The Depth of Processing subscale exhibited a particularly strong association with the rumination scale ($r = .77$, corrected for attenuation due to measurement error). The HSPS total score and the EOE subscale demonstrated a similar pattern concerning these personality traits. In this regard, the Overstimulation and Depth of Processing subscales exhibit similarly strong correlations with existing personality traits (high neuroticism and low extraversion) as the HSPS and its subscale EOE.

Regarding the discriminant validity with openness to experience, from the four DOES subscales, the Emotional Reactivity subscale exhibited the strongest association ($r = .34$); however, according to Steiger's (1980) z value, this association was significantly smaller ($z = 4.135$, $p < .001$) than the relationship between the HSPS subscale AES and the openness to experience subscale ($r = .56$). We have thus achieved a more distinct separation of the DOES subscales from the construct of openness to experience than was previously accomplished with the HSPS. The Emotional Reactivity subscale further displayed significant positive associations of moderate effect sizes with the emotionality (neuroticism) subscale, the rumination scale, and the BIS

Table 4. Measurement Invariance across the German and English Versions of the DOES Scale in Study 2.

	χ^2	*df*	$\Delta\chi^2$	Δdf	*p* value	CFI	ΔCFI	RMSEA	ΔRMSEA	SRMR
Configural	577.665	328				.979		.057		.077
Metric	626.505	344	48.84	16	< .001	.976	.003	.059	−0.002	.080
Scalar	794.254	380	167.75	36	< .001	.965	.011	.068	−0.009	.078
*Partial Scalar	750.351	379	123.85	35	< .001	.969	.007	.064	−0.005	.078

Note. Number of observations per group: German sample: 242, Englisch sample: 232. Subsequential restriction of factorial structure (configural), factor loadings (metric), and intercepts (scalar). *For partial scalar invariance, the intercept threshold t2 of item 15 was freely estimated between the two groups.

and BAS scales. Moreover, the relation to the empathy scale was very high, with r = .81 (corrected for attenuation due to measurement error). Lastly, the Sensing the Subtle subscale demonstrated significant, albeit weak to moderate, positive correlations with the emotionality (neuroticism) subscale and the empathy, rumination, and BAS scales. This scale's unique aspect is its focus on individuals' neutrally phrased perceptual thresholds, setting it apart from other personality traits. This distinction is crucial as Sensing the Subtle does not include the emotional responses often associated with these lower thresholds mentioned by individuals identifying with SPS.

Taken together, the results of Studies 1 and 2 question the concept of an overarching unidimensional SPS factor when assessed through the DOES Scale. Instead, the results suggest that the four subscales are better understood as correlated factors. The Sensing the Subtle subscale is particularly noteworthy, as it explains unique variance with a strong effect size not accounted for by a second-order factor. These findings fit into ongoing debates about whether SPS should be considered as a constellation of personality traits rather than an unidimensional personality trait (Bröhl et al., 2021; Hellwig & Roth, 2021). This idea raises a further question: If SPS is a composite of various personality traits, then the factorial structure of the DOES Scale might vary based on whether people self-identify as highly sensitive or not.

Study 3

In light of these considerations, with Study 3, we aimed to examine the factorial structure of the DOES Scale in two different samples: individuals identifying with high SPS and individuals who did not. We hypothesized that the factorial structure might reveal less integration among the four subscales in individuals not identifying as highly sensitive, highlighting the distinct nature of these traits outside the SPS experience. Conversely, for those identifying as highly sensitive, a more cohesive factorial structure may be expected, as a high score in one subscale would likely come along with higher scores on the others.

Method

Participants and procedure

Participants were native German speakers primarily recruited from the participant pool of the Institute of Psychology at the

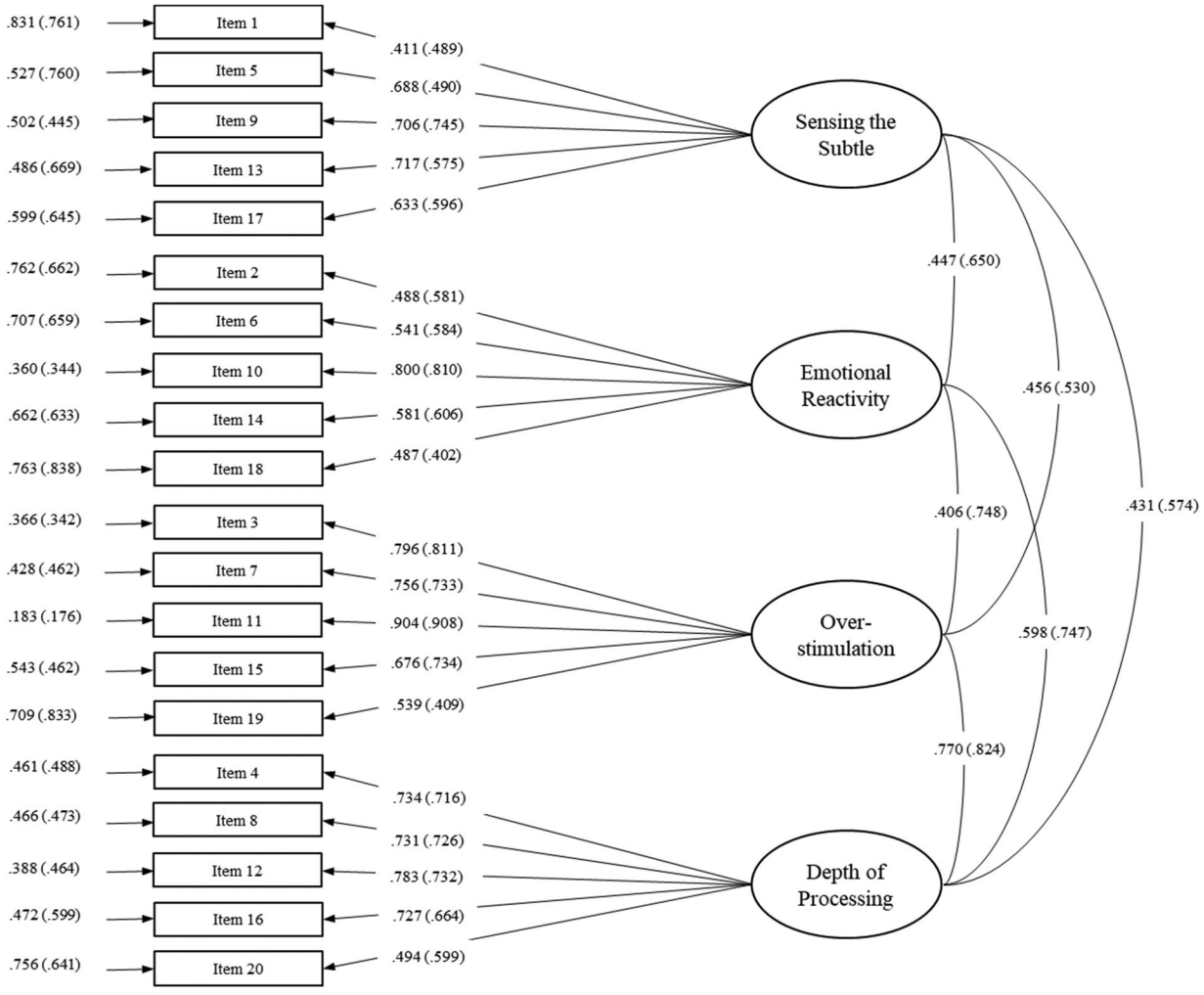

Figure 1. CFA of four correlated factors of the German and English (numbers in brackets) versions of the DOES Scale in Study 2.

Table 5. Zero-order correlations between all measures used in Study 2 in the German-speaking sample (below the diagonal) and the English-speaking sample (above the diagonal).

Variables	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
1. Sensing the Subtle		.44***	.39***	.40***	**.45*****	.36***	.39***	.50***									
2. Emotional Reactivity	.31***		.52***	.54***	**.65*****	.60***	.48***	.62***									
3. Overstimulation	.33***	.25***		.67***	**.79*****	.73***	.75***	.48***									
4. Depth of Processing	.32***	.43***	.60***		**.70*****	.67***	.56***	.54***									
5. HSPS Total Score	**.52*****	**.40*****	**.82*****	**.64*****		.93***	.87***	.77***									
6. EOE Score	.29***	.36***	**.77*****	.62***	.86***		.71***	.57***									
7. LST Score	.50***	.25***	**.77*****	.52***	.93***	.67***		.58***									
8. AES Score	.60***	.54***	.29***	.42***	.56***	.30***	.43***										
9. Honesty/Humility	.00	.17**	.04	.01	−0.02	−0.02	−0.03	.02									
10. Emotionality (Neuroticism)	**.23*****	**.54*****	**.43*****	**.50*****	.51***	.57***	.36***	.30***	.02								
11. Extraversion	**.06**	**−0.04**	**−0.51*****	**−0.37*****	−0.33***	−0.50***	−0.23***	.06	−0.12	−0.25***							
12. Agreeableness	−0.12	−0.02	−0.14*	−0.15*	−0.13	−0.08	−0.12	−0.14*	.15*	−0.13*	.13*						
13. Conscientiousness	.07	.06	−0.13*	.01	−0.16*	−0.20**	−0.13*	.01	.16*	.02	.15*	−0.06					
14. Openness to experiences	**.13***	**.34*****	**−0.03**	**.21****	.12	−0.01	.04	.56***	.07	.15*	.14*	−0.13*	.06				
15. Empathy	**.24*****	**.64*****	**.09**	**.29*****	.25***	.21**	.14*	.44***	.22***	.45***	.18**	.11	.04	.39***			
16. Rumination	**.22*****	**.35*****	**.51*****	**.66*****	.50***	.58***	.35***	.27***	−0.05	.51***	−0.38***	−0.27***	−0.04	.07	.25***		
17. BIS Score	**.11**	**.39*****	**.56*****	**.54*****	.55***	.74***	.35***	.17**	.06	.71***	−0.48***	−0.16*	−0.05	.07	.29***	.61***	
18. BAS Score	.29***	.18**	−0.17**	.04	.00	−0.12	.00	.29***	−0.21**	.11	.46***	−0.04	.18**	.19**	.32***	.01	−0.08

Note. *$p < .05$, **$p < .01$, ***$p < .001$; Number of observations = German-speaking sample 242; English-speaking sample 232. Values in bold refer to the hypothetically assumed relationships of the DOES subscales with established constructs.

University of Bern, who received course credit for their participation. Furthermore, the link for the study was circulated on various SPS networks to recruit people who self-identified with high SPS. Participants were informed that the purpose of the study was to investigate the relationship between different personality traits without mentioning SPS. The study comprised a total of 447 participants, of whom 27 were excluded due to incorrect answers to control items. The final sample included 420 individuals (344 women, 72 men, 4 non-gender specific) with a mean age of 27.39 years (*SD* = 11.5).

Participants first completed the German version of the DOES Scale and the German version of the HSPS (Konrad & Herzberg, 2017; see Study 2). After completing the questionnaires, participants were presented with the following description of SPS:

> Sensory Processing Sensitivity (SPS) is characterized by an increased sensitivity to external (light, noise, etc.) and internal (pain, hunger, etc.) stimuli. Highly sensitive people (HSP) differ from other people in their reactions to negative and positive environmental influences. SPS is often accompanied by a deeper processing of information and a greater perception of the subtleties of the environment. On the downside, SPS can also lead to overstimulation by too many environmental influences.

After reading this definition, participants were asked whether they self-identify with high SPS, which they could answer with either yes or no. Of the 420 participants, 251 (204 women, 44 men, 3 non-gender specific) did not identify with high SPS (Non-HSP group), whereas 169 (140 women, 28 men, 1 non-gender specific) did identify with high SPS (HSP group). As these two samples did not differ significantly in terms of gender, $\chi^2(2) = 0.471$, $p = .790$, age, $t(418) = -0.021$, $p = .983$, and education, $\chi^2(7) = 2.347$, $p = .938$, we carried out further statistical analyses with all individuals of these two groups.

Statistical analysis

First, the mean scores on all subscales and the HSPS were compared between the two groups (HSP vs. Non-HSP) using *t* tests. Effect sizes were examined with Cohen's *d*. The factorial structure of the DOES Scale was further tested by means of CFA using the procedures as in Studies 1 and 2.

Results and discussion

Descriptive statistics for the DOES subscales and the HSPS total score and its subscales for the whole sample, as well as separated for HSP and Non-HSP individuals, are given in Table 6. As expected, HSP individuals had significantly higher scores across all dimensions of the DOES Scale as well as the HSPS total score and its subscales compared to non-HSP individuals. Effect sizes were moderate to strong. Notably, the Overstimulation subscale demonstrated the most significant group difference within the DOES Scale, underscoring its particular relevance in characterizing SPS self-identification. Interestingly, as shown in Table 6, the self-categorization into HSP and non-HSP individuals did not substantially reduce the standard deviations within the

subscales of the two groups compared to the overall sample. This indicates that both groups exhibit similar levels of variability in their responses to the subscales. To summarize, all four DOES subscales effectively distinguished between HSPs and non-HSPs. Furthermore, the observed variability in responses within both groups underscores the need for further research into the nuanced manifestations of sensitivity in individuals. However, it should be noted that some individuals may have been unaware that their heightened sensitivity is called SPS, while others may have over-identified with the construct. These differences in self-perception and identification may have influenced the observed results.

Fit indices for the three different models, including all items for all participants and for the two groups (HSP vs. Non-HSP), are given in Table 7. Fit indices for all participants were similar to Studies 1 and 2. Confirming previous results, the one-factor model did not describe the data well. The four-factor model with correlated factors and the four-factor model with a second-order factor exhibited good model fits. According to the chi-square difference test, the four-factor model with correlated factors described the data significantly better than the second-order model, $\Delta\chi^2(2) = 27.097$, $p < .001$. In the second-order model, the omegaHS values were as follows: Sensing the Subtle, $\omega_{hs} = .573$; Emotional Reactivity, $\omega_{hs} = .306$; Overstimulation, $\omega_{hs} = .244$; and Depth of Processing, $\omega_{hs} = .333$, supporting again the assumption of four correlated factors that explain unique variance (especially Sensing the Subtle) rather than a single second-order factor.

When examining the fit indices separately for the HSP and the non-HSP group, a slightly different picture emerged. For the HSP group, both the four-factor model with correlated factors and the four-factor model with a second-order factor showed good model fit. When comparing both models in the HSP group, the four-factor model with correlated factors did not describe the data significantly better than the second-order model, $\Delta\chi^2(2) = 5.935$, $p = .051$, indicating that the four factors can be subsumed under a second-order factor. The omegaHS values in the HSP group were as follows: Sensing the Subtle, $\omega_{hs} = .564$; Emotional Reactivity, $\omega_{hs} = .379$; Overstimulation, $\omega_{hs} = .282$; and Depth of Processing, $\omega_{hs} = .256$, with strong effect sizes for the first two subscales.

In the non-HSP group, the four-factor model with correlated factors and the four-factor model with a second-order factor showed only acceptable to good model fits. The chi-square difference test was significant, $\Delta\chi^2(2) = 15.488$, $< .001$, indicating, similar to Studies 1 and 2, that the four-factor model with correlated factors explained the data better than the second-order model. The omegaHS values in this group exhibited a large effect size for Sensing the Subtle, $\omega_{hs} = .703$; Overstimulation, $\omega_{hs} = .342$; and Depth of Processing, $\omega_{hs} = .349$, and moderate effect size for Emotional Reactivity, $\omega_{hs} = .257$. The OmegaHS value of the Sensing the Subtle subscale was particularly high, emphasizing its uniqueness in the non-HSP group. This is also reflected in the latent correlations of this subscale with the other three subscales (see Table 8), which are relatively low in the non-HSP group. Essentially, Sensing the Subtle, which is characterized as a subjectively rated perceptual threshold, has relatively little in common with the other three subscales in non-HSP individuals. This divergence probably contributes to the fact that the four subscales in this group cannot be summarized into a single second-order SPS factor. In contrast, the data suggest that the four factors better merge into one overarching factor when examining the four subscales in HSP individuals. These results indicate that Sensing the Subtle plays a unique role in the SPS experience, which is further discussed below.

General discussion

This article introduces the DOES Scale, a novel instrument developed to capture the multifaceted trait of SPS with the four dimensions Depth of Processing, Overstimulation, Emotional Reactivity, and Sensing the Subtle, with five items each. In three studies, we examined the psychometric properties and factorial structure of the German and English versions of the DOES Scale. The scale showed good psychometric properties in terms of internal consistency and retest reliability, and analyses of the factorial structure suggest that SPS is best described as a set of four correlated factors. Notably, we observed different factorial structures among the four subscales when comparing groups of individuals identifying with SPS (HSP) to those who did not (non-HSP). While the four correlated factor model explained the data best in the non-HSP group, with primarily the Sensing the Subtle subscale accounting for a considerable amount of unique variance, the four subscales could be subsumed under a second-order factor in the HSP group. Regarding convergent validity, all DOES subscales were strongly correlated with the total score and subscales of the HSPS, especially the Overstimulation subscale. Regarding discriminant

Table 6. Means, Standard Deviations, t-tests, and Cohen's ds for the four subscales of the DOES Scale and the HSPS total score and its subscales for 420 individuals of Study 3, and separated for 251 individuals not identifying with HSP (Non-HSP) and 169 individuals identifying with HSP, respectively.

	All ($N = 420$)			Self-identified Non-HSP ($N = 251$)			Self-identified HSP ($N = 169$)					
	Mean	*SD*	Min-Max	Mean	*SD*	Min-Max	Mean	*SD*	Min-Max	*t* (418)	*p* value	Cohen's *d*
Sensing the Subtle (1–4)	2.81	0.58	1.00–4.00	2.66	0.54	1.00–4.00	3.04	0.55	1.60–4.00	−7.013	< .001	−0.698
Emotional Reactivity (1–4)	3.20	0.45	1.80–4.00	3.10	0.41	1.80–4.00	3.35	0.46	2.00–4.00	−5.646	< .001	−0.562
Overstimulation (1–4)	2.68	0.65	1.00–4.00	2.42	0.58	1.00–3.80	3.08	0.54	1.60–4.00	−11.740	< .001	−1.170
Depth of Processing (1–4)	2.90	0.63	1.20–4.00	2.75	0.61	1.20–4.00	3.13	0.59	1.60–4.00	−6.350	< .001	−0.632
HSPS Total Score (0–4)	2.09	0.68	0.42–4.00	1.79	0.56	0.42–3.19	2.54	0.59	1.04–4.00	−13.215	< .001	−1.310
EOE Score (0–4)	2.22	0.76	0.20–4.00	1.97	0.69	0.20–3.80	2.58	0.72	0.50–4.00	−8.668	< .001	−0.862
LST Score (0–4)	1.77	0.91	0.00–4.00	1.38	0.74	0.00–3.36	2.36	0.82	0.36–4.00	−12.760	< .001	−1.270
AES Score (0–4)	2.56	0.70	0.60–4.00	2.33	0.66	0.60–3.80	2.88	0.63	1.20–4.00	−8.549	< .001	−0.851

Table 7. Fit indices for the combination of the four DOES subscales into higher-order models across all 420 participants and for 251 individuals not identifying with HSP (Non-HSP) and 169 individuals identifying with HSP, separately.

Model	*N*	χ^2	*df*	*p* value	CFI	RMSEA	SRMR
Study 3 – All							
20-item Model 1 (one-factor model)	420	1150.742	170	< .001	.865	.117	.115
20-item Model 2 (four-factor model with correlated factors)	420	478.656	164	< .001	.957	.068	.080
20-item Model 3 (second-order model)	420	505.753	166	< .001	.953	.070	.082
Study 3 – Self identified as HSP							
20-item Model 1 (one-factor model)	169	493.812	170	< .001	.868	.106	.120
20-item Model 2 (four-factor model with correlated factors)	169	240.557	164	< .001	.969	.053	.089
20-item Model 3 (second-order model)	169	246.492	166	< .001	.967	.054	.090
Study 3 – Self identified as Non-HSP							
20-item Model 1 (one-factor model)	251	818.064	170	< .001	.754	.123	.128
20-item Model 2 (four-factor model with correlated factors)	251	382.003	164	< .001	.917	.073	.094
20-item Model 3 (second-order model)	251	397.521	166	< .001	.912	.075	.096

Note. CFI = comparative fit index, RMSEA = root-mean-square error of approximation, SRMR = standardized root-mean-square residual.

Table 8. Latent Correlations of the four subscales for the HSP (below the diagonal and the Non-HSP (above the diagonal) group.

Measures	1.	2.	3.	4.
1. Sensing the Subtle		.258***	.206***	.056
2. Emotional Reactivity	.404***		.448***	.540***
3. Overstimulation	.380***	.490***		.568***
4. Depth of Processing	.343***	.558***	.678***	

Note. ***p < .001; Number of observations: HSP group = 169; Non-HSP group = 251.

validity, three of the four DOES subscales (except for Sensing the Subtle) exhibited moderate to strong correlations with established personality traits, including neuroticism, extraversion, empathy, and rumination. Still, no complete construct overlap was found even when corrected for attenuation due to measurement error. In addition, the DOES Scale can be better dissociated from openness to experience than the HSPS, which remedies one of the measurement issues previously raised by researchers (Hellwig & Roth, 2021).

Factorial structure of the DOES Scale

The factorial structure of the DOES Scale suggests that SPS should be better conceptualized as a specific constellation of four correlated dimensions rather than one overarching personality trait that underlies these four characteristics. This challenges the notion of a higher-order unidimensional SPS trait originally posited by Aron and Aron (1997). Previous studies, especially those using the HSPS, have postulated that the three subscales, EOE, LST, and AES, can be subsumed under a second-order factor (Konrad & Herzberg, 2017; Smolewska et al., 2006). However, this assumption could never be formally tested due to the statistical equivalence (i.e., same degrees of freedom and model fit) between second-order models with three lower-order factors and models with correlated factors of the same number (Gignac & Kretzschmar, 2017; Reise, 2012). By addressing this gap, the present research offers new insights into the conceptualization of SPS.

The analysis of OmegaHS values in our three studies revealed that the Sensing the Subtle subscale consistently exhibited unique variance, with a large effect size not captured by a second-order factor, emphasizing the complexity in the DOES structure beyond an overarching unidimensional factor. Study 3 particularly highlighted this, revealing that the Sensing the Subtle subscale accounted for a high amount of unique variance in the non-HSP group. Interestingly, within the HSP group, the four subscales could be combined into a second-order SPS factor, which illustrates that the interplay among the four subscales differs across groups.

This result could be interpreted as follows: While three of the four DOES subscales primarily capture various emotional and cognitive reactions to everyday experiences, the "Sensing the Subtle" scale deviates by focusing on a nuanced perception of the environment, i.e., it captures more sensory manifestations. This characteristic appears to be more independent of emotional and cognitive processing in non-HSP individuals. In HSP individuals, in contrast, all four characteristics seem to be more closely connected. Thus, a lowered perceptual threshold of the environment may lead to more intensive emotional and cognitive processing, resulting in a vulnerability to overstimulation. However, the extent to which these processes build on each other (if they do at all) or why a lower perceptual threshold is associated with increased cognitive and emotional processing in HSP but not in non-HSP individuals remains an open question, which can be investigated in future studies using the DOES Scale.

Psychometric properties

The four subscales of the DOES Scale exhibited consistently acceptable to good internal consistencies across the three presented studies and both language versions. The retest reliability of the German version of the DOES Scale additionally exhibited high test-retest reliabilities for all four subscales, indicating stability over one month. Importantly, Boyle (1991) points out that high internal consistency often results from item content redundancy and does not necessarily indicate high reliability. Instead, acceptable to good internal consistency, as observed with the DOES Scale, could mean a broader content coverage of the respective subscales, which can avoid the pitfalls of overly narrow scales. Furthermore, McCrae (2015) emphasized that test-retest reliability is a better measure of reliability than internal consistency as it additionally captures item-specific variance, which reflects the stability of a construct over time. Therefore, the

relatively high test-retest reliability of the DOES Scale confirms its stability and reliability.

Convergent and discriminant validity

The consistent positive relationship between the DOES subscales and the HSPS supports the convergent validity of the new instrument in both the German and English versions. In particular, the Overstimulation subscale of the DOES Scale exhibited strong associations with the HSPS total score and its subscales EOE and LST, emphasizing the primary focus of the HSPS on measuring the negative aspects of SPS, as criticized in previous studies (Evans & Rothbart, 2008; Hellwig & Roth, 2021). Roth et al.'s (2023) study revealed that Overstimulation addresses only one aspect of SPS, ignoring its potential neutral and positive aspects. In the HSPS, positive aspects are mainly represented in the AES subscale. However, this subscale is underrepresented with the number of items in the overall scale and displays near complete overlap with the personality trait "Openness to experience" (Hellwig & Roth, 2021), questioning its discriminant validity. In contrast, the DOES Scale, especially the Sensing the Subtle and Emotional Reactivity subscales, successfully covers positive and neutral aspects of SPS and thus represents the entire range of experiences associated with high SPS.

When examining the discriminant validity of the DOES subscales, the Sensing the Subtle subscale stands out most clearly from existing personality constructs, measuring individuals' subjectively rated perceptual threshold. This subscale contributes significantly to the DOES Scale's ability to capture the entire spectrum of SPS, closing a crucial gap in the measurement of SPS by providing insights into the subtle ways in which individuals perceive their environment independent of their emotional and cognitive reactions. In evaluating the discriminant validity against well-established personality constructs, particularly the Big Five traits of neuroticism, extraversion, and openness to experience, the Overstimulation and Depth of Processing subscales of the DOES Scale displayed correlation patterns with neuroticism and extraversion that were closely aligned with those observed for the HSPS and its EOE subscale. However, the DOES subscales achieved a more pronounced differentiation from openness to experience compared to the HSPS, with the Emotional Reactivity subscale showing only a moderate positive correlation. These results suggest that, as with the HSPS, overlaps exist with high neuroticism and low extraversion within specific DOES subscales. However, these overlaps can now be more accurately attributed to SPS itself, as the DOES subscales were developed based on interviews with individuals who identify as highly sensitive. These correlations likely represent aspects of the experience of high sensitivity (e.g., when overstimulated, one is more often stressed and may be more likely to want to withdraw and be alone), thus enriching the construct's definition and conceptual framework. As a result, the DOES subscales allow for a more nuanced categorization within established personality constructs and eliminate the concerns that the observed correlations are merely the by-product of a predominantly negative item wording.

However, further exploration of discriminant validity revealed that the Emotional Reactivity subscale exhibited strong correlations with empathy and Depth of Processing with rumination. These results evoke concerns related to the 'jangle fallacy' (Kelley, 1927), which suggests that different labels are used for constructs that capture similar underlying personality dimensions. Thus, critical voices could point out that the subscales of the DOES Scale – Depth of Processing, Emotional Reactivity, and Overstimulation – are proxies for rumination, empathy, neuroticism, and introversion. It should be noted that the observed correlations do not indicate complete overlaps between the constructs, even when corrected for measurement error. Although the Emotional Reactivity subscale correlates significantly with empathy, it captures aspects of emotional experience that are not only restricted to interpersonal contexts. Furthermore, empathy encompasses a more nuanced trait of understanding, adaptively responding to, and effectively communicating the emotions of others, thereby encouraging prosocial behavior (Spreng et al., 2009). This broadens the concept beyond simple emotional reactions. Consequently, Emotional Reactivity could be viewed as a contributing yet distinct factor to empathy. The Depth of Processing subscale, in turn, includes deeper processing of both positive and negative experiences (e.g., "After experiencing something positive or negative, I think about what happened for a long time.") and thus differs from the rumination scale, which generally focuses on persistent and frequently negative thoughts (Nolen-Hoeksema, 1991).

Despite the discussion about whether the dimensions of the SPS represent unique constructs or a constellation of existing traits (Hellwig & Roth, 2021), it becomes clear that the four extracted dimensions of the DOES Scale play a crucial role in defining high SPS, contributing to the content validity of the scale. This emphasizes how important their inclusion is for understanding the full spectrum of the high SPS experience. In particular, the Sensing the Subtle subscale introduces a new component to the study of SPS that distinguishes it from other aspects of high sensitivity and other personality traits. Whether studied independently or in conjunction, the four DOES subscales enrich our understanding of SPS as a complex personality constellation. The question of whether some of the four DOES subscales correspond to or go beyond established personality constructs remains a topic for future research to deepen the conceptual clarity of SPS.

Limitations

This study has several limitations. Firstly, the samples in Study 1 and Study 3 were predominantly composed of young, female, well-educated individuals, which limits their representativeness. While the samples of Study 2 were somewhat more diverse, they were drawn from the online platform Prolific, raising concerns about including careless respondents and potential non-human respondents (bots; Douglas et al., 2023). Despite implementing attention checks, it is possible that not all such instances were identified. Furthermore, the

self-selection bias inherent in these samples may have influenced the results, particularly in how participants responded to the measures. Additionally, recruiting participants from SPS forums in Study 3 might have introduced confirmation bias, as these individuals were likely more familiar with the concept of SPS and could have responded in ways that aligned with existing theories and their self-concept. Efforts were made to mitigate these biases by withholding the specific purpose of the study from participants until after they completed the questionnaires. However, some individuals may have recognized the questionnaires, which does not entirely eliminate the potential for confirmation bias. To address these issues in future studies, it would be valuable to recruit participants from more diverse backgrounds to ensure broader representativeness. Additionally, randomizing the order of questionnaires or adding more subtle instructions during recruitment could be effective strategies to minimize response bias. Another limitation lies in the item construction process, which resulted from interview statements of participants who self-identified with high SPS. These statements may be influenced by Aron and colleagues' (1997, 2012) work, aligning with their theory as it is quite popular in both academics and society. Therefore, it is almost impossible to understand SPS without referring to Aron's work. Nonetheless, the interview study by Roth et al. (2023) asked about specific experiences and behaviors related to SPS. The statements thus reflect the aspects of Aron's theory that are relevant to individuals identifying with high SPS.

Conclusion and outlook

To summarize, the DOES Scale with its four subscales – Depth of Processing, Overstimulation, Emotional Reactivity, and Sensing the Subtle – not only aligns with the theoretical framework of SPS as proposed by Aron in her DOES model but also exhibits good psychometric properties across multiple studies and languages. This scale marks a pivotal development in the empirical assessment of SPS, offering a comprehensive tool that captures the full spectrum of characteristics of those identifying with high SPS. Unlike the HSPS, which primarily focuses on Overstimulation, the DOES Scale provides a more nuanced and neutral understanding of how individuals perceive and interact with their surroundings. Additionally, the four dimensions of the DOES Scale are grounded in the narratives of highly sensitive individuals compared to the empirical derivation of the HSPS, facilitating their more straightforward interpretation. Moreover, the results of this study argue in favor of a novel interpretation of SPS in that the four subscales of the DOES Scale should be understood as manifestations of the four trait constellations and not as a mere aggregation of the components into a unidimensional SPS factor. This approach not only promotes a deeper understanding of the individual characteristics of the four dimensions but also makes it possible to consider the interaction of the four dimensions instead of relying on an overarching SPS factor. This means that the DOES Scale allows for more nuanced examinations of the varying levels of expression within the four dimensions. In this sense, it is possible to go beyond the purely additive effects of the subscales and interpret SPS as a combination of trait constellations.

Declaration of interest

No potential conflict of interest was reported by the author(s).

Data availability statement

Data and analysis script can be accessed under the Open Science Framework: https://osf.io/2uj3c

ORCID

Danièle Anne Gubler http://orcid.org/0000-0003-4317-9939
Tobias Janelt http://orcid.org/0000-0002-3138-5818
Katja Schlegel http://orcid.org/0000-0003-3768-5544
Stefan Johannes Troche http://orcid.org/0000-0002-0961-1081

References

Aron, E. N. (2020). *The highly sensitive person. How to thrive when the world overwhelms you*. Citadel Press.

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. https://doi.org/10.1037/0022-3514.73.2.345

Aron, E. N., Aron, A., & Jagiellowicz, J. (2012). Sensory processing sensitivity: A review in the light of the evolution of biological responsivity. *Personality and Social Psychology Review*, *16*(3), 262–282. https://doi.org/10.1177/108886831143421

Bordarie, J., Aguerre, C., & Bolteau, L. (2022). Validation and study of psychometric properties of a French version of the Highly Sensitive Person Scale (HSPS-FR). *European Review of Applied Psychology*, *72*(4), 100781. https://doi.org/10.1016/j.erap.2022.100781

Boyle, G. J. (1991). Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality and Individual Differences*, *12*(3), 291–294. https://doi.org/10.1016/0191-8869(91)90115-R

Bröhl, A. S., Van Leeuwen, K., Pluess, M., De Fruyt, F., Van Hoof, E., Weyn, S., & Bijttebier, P. (2021). Personality profile of the self-identified highly sensitive person. *Journal of Individual Differences*, *43*(2), 95–104. https://doi.org/10.1027/1614-0001/a000363

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Chacón, A., Pérez-Chacón, M., Borda-Mas, M., Avargues-Navarro, M. L., & López-Jiménez, A. M. (2021). Cross-cultural adaptation and validation of the highly sensitive person scale to the adult Spanish population (HSPS-S). *Psychology Research and Behavior Management*, *14*, 1041–1052. https://doi.org/10.2147/PRBM.S321277

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

De Gucht, V., Woestenburg, D. H., & Wilderjans, T. F. (2022). The different faces of (high) sensitivity, toward a more comprehensive measurement instrument. Development and validation of the Sensory Processing Sensitivity Questionnaire (SPSQ). *Journal of Personality Assessment*, *104*(6), 784–799. https://doi.org/10.1080/00223891.2022.2032101

Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLoS One*, *18*(3), e0279720. https://doi.org/10.1371/journal.pone.0279720

Ershova, R. V., Yarmotz, E. V., Koryagina, T. M., Semeniak, I. V., Shlyakhta, D. A., & Tarnow, E. (2018). A psychometric evaluation of the highly sensitive person scale: The components of sensory-processing sensitivity. *Electronic Journal of General Medicine*, *15*(6), 1–7. https://doi.org/10.29333/ejgm/100634

Evans, D. E., & Rothbart, M. K. (2008). Temperamental sensitivity: Two constructs or one? *Personality and Individual Differences*, *44*(1), 108–118. https://doi.org/10.1016/j.paid.2007.07.016

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, *3*(4), 484–501. https://doi.org/10.1177/2515245920951747

Gignac, G. E., & Kretzschmar, A. (2017). Evaluating dimensional distinctness with correlated-factor models: Limitations and suggestions. *Intelligence*, *62*, 138–147. https://doi.org/10.1016/j.intell.2017.04.001

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*, 287–305. https://doi.org/10.1016/j.neubiorev.2019.01.009

Grimen, H. L., & Diseth, Å. (2016). Sensory processing sensitivity: Factors of the highly sensitive person scale and their relationships to personality and subjective health complaints. *Perceptual and Motor Skills*, *123*(3), 637–653. https://doi.org/10.1177/0031512516666114

Hellwig, S., & Roth, M. (2021). Conceptual ambiguities and measurement issues in sensory processing sensitivity. *Journal of Research in Personality*, *93*, 104130. https://doi.org/10.1016/j.jrp.2021.104130

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

International Test Commission. (2017). ITC guidelines for translating and adapting tests (second edition). *International Journal of Testing*, *18*, 101–134. https://doi.org/10.1080/15305058.2017.1398166

Janelt, T., Altmann, T., Spreng, R. N., & Roth, M. (2024). Analyzing the factor structure of the Toronto Empathy Questionnaire: Dimensionality, reliability, validity, measurement invariance and one-year stability of the German version. *Journal of Personality Assessment*, *106*(2), 230–241. https://doi.org/10.1080/00223891.2023.2224873

Kelley, T. L, Interpretation of educational measurements (1927). World Book Company. Mindrila, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society*, *1*(1), 60–66.

König, D. (2012). *Deutsche Version der Skala Rumination aus dem Rumination-Reflection Questionnaire (RRQ)[German version of the Rumination subscale from the Rumination-Reflection Questionnaire (RRQ)]*. http://dk.akis.at/RRQ_Rumination.pdf

Konrad, S., & Herzberg, P. Y. (2017). Psychometric properties and validation of a german high sensitive person scale (HSPS-G). *European Journal of Psychological Assessment*, *35*(3), 364–378. https://doi.org/10.1027/1015-5759/a000411

Kyziazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(8), 2207–2230. https://doi.org/10.4236/psych.2018.98126

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 high-sensitive individuals. *Translational Psychiatry*, *8*(1), 24. https://doi.org/10.1038/s41398-017-0090-6

Lionetti, F., Pastore, M., Moscardino, U., Nocentini, A., Pluess, K., & Pluess, M. (2019). Sensory processing sensitivity and its association with personality traits and affect: A meta-analysis. *Journal of Research in Personality*, *81*, 138–152. https://doi.org/10.1016/j.jrp.2019.05.013

May, A. K., Norris, S. A., Richter, L. M., & Pitman, M. M. (2022). A psychometric evaluation of the Highly Sensitive Person Scale in ethnically and culturally heterogeneous South African samples. *Current Psychology*, *41*(7), 4760–4774. https://doi.org/10.1007/s12144-020-00988-7

McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review*, *19*(2), 97–112. https://doi.org/10.1177/1088868314541857

Mindrila, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society*, *1*(1), 60–66.

Moshagen, M., & Bader, M. (2023). semPower: General power analysis for structural equation models. *Behavior Research Methods*, *56*(4), 2901–2922. https://doi.org/10.3758/s13428-023-02254-7

Moshagen, M., Hilbig, B. E., & Zettler, I. (2014). Faktorenstruktur, psychometrische Eigenschaften und Messinvarianz der deutschsprachigen Version des 60-item HEXACO Persönlichkeitsinventars. *Diagnostica*, *60*(2), 86–97. https://doi.org/10.1026/0012-1924/a000112

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology*, *100*(4), 569–582. https://doi.org/10.1037/0021-843X.100.4.569

Palese, T., & Mast, M. S. (2020). Interpersonal accuracy and interaction outcomes: Why and how reading others correctly has adaptive advantages in social interactions. In R. J. Sternberg & A. Kostić (Eds.), *Social intelligence and nonverbal communication* (pp. 305–331). Palgrave Macmillan.

Pluess, M. (2015). Individual differences in environmental sensitivity. *Child Development Perspectives*, *9*(3), 138–143. https://doi.org/10.1111/cdep.12120

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2012). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement*, *73*(1), 5–26. https://doi.org/10.1177/0013164412449831

Rosseel, Y. (2012). lavaan: An R Package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Roth, M., Gubler, D. A., Janelt, T., Kolioutsis, B., & Troche, S. J. (2023). On the feeling of being different–an interview study with people who define themselves as highly sensitive. *PLoS One*, *18*(3), e0283311. https://doi.org/10.1371/journal.pone.0283311

Satow, L. (2022). HSP. HSP-Test für Hochsensibilität [Verfahrensdokumentation, Fragebogen in Lang- und Kurzform, Testdokumentation]. Open Test Archive. Trier: ZPID. https://doi.org/10.23668/psycharchives.5419

Schweizer, K. (2010). Some guidelines concerning the modeling of traits and abilities in test construction. *European Journal of Psychological Assessment*, *26*(1), 1–2. https://doi.org/10.1027/1015-5759/a000001

Şengül-İnal, G., & Sümer, N. (2020). Exploring the multidimensional structure of sensory processing sensitivity in Turkish samples. *Current Psychology*, *39*(1), 194–206. https://doi.org/10.1007/s12144-017-9751-0

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. https://doi.org/10.1016/j.paid.2005.09.022

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. https://doi.org/10.1016/j.paid.2015.03.045

Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment*, *91*(1), 62–71. https://doi.org/10.1080/00223890802484381

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245–251. https://doi.org/10.1037/0033-2909.87.2.245

Strobel, A., Beauducel, A., Debener, S., & Brocke, B. (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *22*(3), 216–227. https://doi.org/10.1024//0170-1789.22.3.216

Trapnell, P. D., & Campbell, J. D. (1999). Private self-consciousness and the five-factor model of personality: Distinguishing rumination from reflection. *Journal of Personality and Social Psychology*, *76*(2), 284–304. https://doi.org/10.1037/0022-3514.76.2.284

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21–29. https://doi.org/10.22237/jmasm/1177992180