Measurement in Sensory Modulation: The Sensory Processing Scale Assessment

Sarah A. Schoen, Lucy J. Miller, Jillian C. Sullivan

OBJECTIVE. Sensory modulation issues have a significant impact on participation in daily life. Moreover, understanding phenotypic variation in sensory modulation dysfunction is crucial for research related to defining homogeneous groups and for clinical work in guiding treatment planning. We thus evaluated the new Sensory Processing Scale (SPS) Assessment.

METHOD. Research included item development, behavioral scoring system development, test administration, and item analyses to evaluate reliability and validity across sensory domains.

RESULTS. Items with adequate reliability (internal reliability > .4) and discriminant validity (p < .01) were retained. Feedback from the expert panel also contributed to decisions about retaining items in the scale.

CONCLUSION. The SPS Assessment appears to be a reliable and valid measure of sensory modulation (scale reliability > .90; discrimination between group effect sizes > 1.00). This scale has the potential to aid in differential diagnosis of sensory modulation issues.


Sensory modulation is the ability to regulate and grade responses to the sensory environment so that responses to sensory input are appropriate to the demands of daily life (Miller, Anzalone, Lane, Cermak, & Osten, 2007). Although different models exist to describe sensory modulation challenges, there appears to be some general agreement or consensus as to the behaviors associated with these subgroupings (Dunn, 1999; Dunn & Westman, 1997; Miller et al., 2007). Overresponsivity (also termed sensory sensitivity or sensory avoiding) is characterized by exaggerated, negative responses to typical sensory experiences in daily life. Underresponsivity (also termed low registration) is characterized by muted or delayed responses to daily sensory events, and sensory craving (SC; also termed sensory seeking) is characterized by an insatiable drive for enhanced sensory experiences. Similar categories exist in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM–5; American Psychiatric Association, 2013) for autism spectrum disorder (ASD) and in the Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood (Zero to Three, 2005; see also Postert, Averbeck-Holocher, Beyer, Müller, & Furniss, 2009). However, a reliable measure to identify these behavioral patterns is lacking (Postert et al., 2009).

Research and development of the Sensory Processing Scale (SPS) Assessment has been ongoing since 2002. Preliminary information on Version 1 was published in 2008 (Schoen, Miller, & Green, 2008), focusing on sensory overresponsivity (SOR). The expanded version of the SPS Assessment includes all three subtypes of sensory modulation dysfunction identified in the literature: SOR, sensory underresponsivity (SUR), and SC (Dunn, 1999; Miller et al., 2007; Parham, Ecker, Kuhaneck, Henry, & Glennon, 2007).
Need for the Sensory Processing Scale Assessment

Sensory modulation challenges affect social participation, play, and engagement in home and school routines (Bar-Shalita, Vatine, & Parush, 2008; Bundy, Shia, Qi, & Miller, 2007; Cohn, Miller, & Tickle-Degnen, 2000; Cosbey, Johnston, & Dunn, 2010). Symptoms are documented in children with a range of developmental disabilities (Ben-Sasson et al., 2009; Dar, Kahn, & Carmeli, 2012; Ghanizadeh, 2011; Lane, Reynolds, & Thacker, 2010), including as many as 90% of children with ASD (Tomchek & Dunn, 2007). Surprisingly, in spite of the relatively high rates reported in the clinical literature (Baranek, David, Poe, Stone, & Watson, 2006; Lane et al., 2010; Leekam, Nieto, Libby, Wing, & Gould, 2007; Parush, Sohmer, Steinberg, & Kaitz, 2007), a standard method of assessment that uses an examiner-administered scale does not exist.

Discriminating characteristics of sensory modulation dysfunction within and among clinical disorders is crucial (Reynolds & Lane, 2008). From a clinical perspective, classifying a specific set of symptoms facilitates the development of appropriate intervention strategies. From a research perspective, identifying a specific cluster of symptoms facilitates investigations into associated neurobiological, psychological, and psychosocial features (Postert, Averbeck-Holocher, Achtergarde, Müller, & Furniss, 2012; Postert et al., 2009).

The SPS Assessment uses direct observation to assess behavioral responses to sensory occurrences that are similar to daily life experiences. It is intended to contribute accuracy and consistency to the diagnosis of sensory modulation dysfunction by providing an additional source of data to informant-based reports and facilitating further objective exploration of sensory modulation phenotypes.

Benefit of Examiner-Administered Tools

Multiple methods of assessment are crucial to diagnosing childhood disorders (Gualtieri & Johnson, 2005; Jensen et al., 1999; Kerr, Lunkenheimer, & Olson, 2007; Sharp, Jaquess, & Lukens, 2013). Multiple sources provide the richest source of information about a person’s functioning and improve diagnostic accuracy (Woodard et al., 2012). Multidimensional assessment batteries maximize the validity of individual assessments (Meyer et al., 2001), especially when informant reports are combined with performance measures. Data from a meta-analysis of more than 125 studies concluded that multiple assessment methods provide unique information essential to a complete understanding of clients (Meyer et al., 2001).

Parent and caregiver questionnaires addressing sensory modulation are criticized because of their limited evidence of validity (Woodard et al., 2012). There is the potential for respondent bias and lack of agreement with therapist observations (Ben-Sasson et al., 2009). The need exists to develop quantitative methods of direct observation of sensory modulation challenges so as to provide clearer and more consistent criteria for determining clinical impairment (Reynolds & Lane, 2008; Rogers & Luby, 2011; Woodard et al., 2012).

The existing measures of sensory processing and integration that use direct observation do not offer the benefits of the SPS Assessment. The most commonly used performance measure of sensory processing is the Sensory Integration and Praxis Tests (SIPT; Ayres, 1989). However, the SIPT was not designed to systematically quantify modulation behaviors. Another performance measure, the Sensory Processing Assessment (SPA), is limited to auditory, visual, and tactile domains; was designed only for children with ASD; is not standardized; and is largely used in research rather than clinical practice (Baranek, Boyd, Poe, David, & Watson, 2007; Baranek et al., 2013). Both the SIPT and the SPA offer important augmentation of caregiver- and parent-report measures. Neither tool, however, is designed to assess modulation across all seven sensory domains (visual, auditory, tactile, vestibular, proprioceptive, gustatory, and olfactory) or to characterize behaviors reflective of all modulation challenges.

Scale Research and Development

Expansion of the Scale

The SPS Assessment was expanded to include new activities that elicit SUR and SC behaviors for all seven sensory domains. Consensus on the representativeness of the activities in measuring each sensory domain and the characteristics of each subtype was established from focus groups. Changes based on their recommendations included using high-intensity stimuli to capture SUR and providing continued interaction with test materials beyond activity administration to capture SC behaviors.

Development of the Behavioral Coding Scheme

Three research interns watched videotaped administrations (N = 20) and grouped the most frequently observed behaviors into 17 behavioral categories. These categories were combined, clarified, and simplified to produce descriptors of atypical behaviors and 1 descriptor.
for typical behavior. The final behavioral category labels were typical, concern, avoidance–withdrawal, adverse, unaware–minimal interaction, slow to respond, low energy–weak body, wants more, extra movement–force, noisy–talkative, and distractible.

The scoring scheme was then tested on a new sample of 20 children ages 4–12 yr (10 typically developing children and 10 children with sensory modulation issues). Reliability (>80% agreement) was established among five researchers, and the operational definitions were further refined.

Study Purpose

The purpose of this study was to investigate the psychometric properties of the SPS Assessment Version 2.0. Specifically, we aimed to reduce the number of items in the new version and report its internal reliability and discriminant validity. In addition, we sought to establish a behavioral coding system that accurately reflects behaviors consistent with clinical observation of people with sensory modulation challenges.

Method

The Colorado Multiple Institutional Review Board (IRB) of the University of Colorado, Denver, approved all procedures for this study.

Description of the Scale

The SPS Assessment is composed of domains, which reflect the sensory systems; subtests, which reflect activities or games played; and items, which are individually scored for response during and after game administration. Each domain has multiple activities or games that are administered in a standardized manner to all test takers. Within the subtests, the examiner uses standard criteria to score specific responses that occur either during or after the activity. Activities are designed to closely reflect daily sensory experiences that elicit atypical responses in children with sensory challenges (e.g., adhesive bandages, goo, background noise).

Subtests and Items. Version 2.0 of the SPS Assessment contains 34 subtests across seven sensory domains (Visual, Auditory, Tactile, Vestibular, Proprioceptive, Gustatory, and Olfactory). Items within each subtest are scored to reflect the person’s responses at three time periods: (1) during the activity, (2) after the activity (≤15 s), and (3) during the transition to the next activity.

Behavioral Scoring System. The scoring system uses dichotomous “yes, observed,” and “no, not observed,” behavioral categories. Detailed operational definitions for each behavioral category are included in the manual. After administration of the games in each sensory domain, a rating of global clinical impression (GCI) of SOR, SUR, SC, or typical responsivity is made within each domain.

Procedures

Occupational therapists with expertise in sensory processing and integration (N = 15) were trained in administration and scoring of the SPS Assessment Version 2.0 using videotapes of test administration and examples reflecting each behavioral category. Interrater reliability for the expert panel was calculated using intraclass correlation coefficients (ICCs). Tactile (.77) and Olfactory (.44) domains could not be sufficiently observed on the videotape because of poor camera positioning, but the other domains and the total test had ICCs >.90.

Written consent was obtained from all parents (and assent from children age 7 yr or older) before participation in the study. Test administrations were conducted at a clinic or research laboratory in a quiet, nondistracting room and lasted approximately 1 hr per participant. Parents of child participants completed a demographic form.

Participants

A convenience sample of 128 participants ages 4–19 yr (mean = 9.01, standard deviation = 2.91) participated. Of these, 70% were male. Sixty-three participants were typically developing, and 65 participants had sensory modulation challenges. Data were provided by six expert clinicians from around the country (California, Colorado, Missouri, New York, Ohio, and Virginia) who were trained in administration of the SPS Assessment. All participants were recruited through IRB-approved advertisements. Typical development was verified with a questionnaire. Inclusion criteria included (1) no history of atypical sensory responsivity; (2) no birth risk factors; (3) no neurological, psychological, developmental, or learning disability; and (4) no history of receiving therapeutic services. Sensory modulation difficulties were based on the GCI of referring occupational therapists after a comprehensive evaluation, which included standardized scales, observations in the gym, parent interview, and standardized parent-report measures. All clinical participants were rated as having sensory modulation challenges. Of the sample, 28% of the 65 clinical participants (n = 18) had parent-reported ASD, and 12% (n = 8) had parent-reported attention deficit hyperactivity disorder (ADHD). Because these disorders are
more common in boys, that 70% of the clinical sample was male was not unexpected (Boyle et al., 2011).

Data Analysis

Data analyses were conducted at the item, subtest, and domain levels. Each domain consisted of multiple subtests (activities or games), and each subtest consisted of multiple items. We analyzed items using standard α coefficients to evaluate internal consistency and Mann–Whitney U values to determine which items best discriminated between the clinical and the typical samples. Once the pool of items was reduced, we calculated internal consistency reliability and discriminant validity for each sensory domain.

The behavioral scoring categories were analyzed to determine where clarification in behavioral scoring criteria was needed and to reduce overlap between categories. This analysis included (1) Spearman’s ρ correlations among the 10 atypical behavioral categories within each sensory domain, (2) frequencies of agreement between the expert GCI and the rating on the behavioral response category, and (3) factor analyses using principal components analysis with orthogonal varimax rotation of the behavioral scoring categories. SPSS Version 21 (IBM Corp., Armonk, NY) was used for all analyses.

Results

Age and Gender Analyses

We found no significant differences between the typical and clinical groups on age (z = -0.57, p = .57) or gender distribution, χ²(1) = 1.55, p = .21.

Reduction of Items and Subtests

Each subtest had three item responses that were rated: during, after, and transition. We conducted analyses to determine items that could be deleted to increase subtest reliability or validity.

We deleted five response-during items (two Proprioceptive and three Olfactory) because they did not significantly discriminate between the typical and the clinical samples. We also deleted items that correlated <.3 with domain scores because they were the weakest items and the least related to their respective domains (i.e., one Auditory, one Visual, one Gustatory, and two Vestibular items). Behaviors during transitions differentiated typical from clinical groups in 24% of subtests (8 of 34). Therefore, transition scoring was retained for only the end of each domain. We noted the subtests for which response immediately after activity administration showed significant differences between groups so that this psychometric characteristic could be used for final decisions related to subtest retention.

Internal Reliability and Discriminant Validity

We examined total test scores and domain scores to evaluate the internal reliability and discriminant validity of the remaining items and subtests. Overall internal consistency yielded α = .94, and domain reliabilities ranged from .79 to .93 (see Table 1).

Discriminant validities for domains and the total test were all statistically significant (p ≤ .002). Table 2 shows the means and standard deviations for both the typical and the clinical groups. The effect sizes of differences between groups denote large effects for most domains.

Analysis of Behavioral Scoring Categories

We interpreted the SC data cautiously because of the small number in this group (n = 4) except in the Proprioceptive (n = 13) and Vestibular (n = 16) domains.

Association Between Behavioral Categories and Global Clinical Impression.

The behavioral categories hypothesized to measure SOR were “concern,” “avoidance–withdrawal,” and “adverse.” GCIs of SOR endorsed all three behavioral ratings in the Auditory, Tactile, Olfactory, and Gustatory domains. Unanticipated endorsements of “extra movement–force,” “unaware–minimal interaction,” and “noisy–talkative.”

The behavioral categories hypothesized to measure SUR were “unaware–minimal interaction,” “slow to respond,” and “low energy–weak body.” GCIs of SUR endorsed “unaware” in the Visual, Auditory, Tactile, Olfactory, and Gustatory domains and “low energy–weak body” and “extra movement–force” in the Vestibular and Proprioceptive domains.

The behavioral categories hypothesized to relate to SC were “wants more,” “noisy–talkative,” and “extra movement–force.” GCIs of SC endorsed “wants more” in all seven sensory domains and “noisy–talkative” and

<table>
<thead>
<tr>
<th>Domain</th>
<th>α Coefficient</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tactile</td>
<td>.82</td>
<td>1.68</td>
</tr>
<tr>
<td>Auditory</td>
<td>.92</td>
<td>1.66</td>
</tr>
<tr>
<td>Visual</td>
<td>.79</td>
<td>0.94</td>
</tr>
<tr>
<td>Vestibular</td>
<td>.93</td>
<td>1.68</td>
</tr>
<tr>
<td>Proprioceptive</td>
<td>.89</td>
<td>1.51</td>
</tr>
<tr>
<td>Olfactory</td>
<td>.79</td>
<td>0.72</td>
</tr>
<tr>
<td>Gustatory</td>
<td>.81</td>
<td>1.49</td>
</tr>
<tr>
<td>Total</td>
<td>.94</td>
<td>4.88</td>
</tr>
</tbody>
</table>

Note. SEM = standard error of the mean.
“extra movement–force” in the Vestibular and Proprioceptive domains. Unexpectedly, “low energy–weak body” was also endorsed.

“Distractible” correlated only with “noisy–talkative,” “extra movement–force,” and “wants more” in the Vestibular and Proprioceptive domains.

Factor Analyses of Behavioral Response Categories. We used principal-components analyses to explore the underlying structure of the behavioral categories. A summary variable was calculated for each domain by summing the dichotomous item scores for each behavioral response category. Results partially support the hypothesized grouping of behavioral categories as described for SOR, SUR, and SC. Table 3 displays the factor loadings. These analyses replicated our findings that the behavioral response categories “noisy–talkative,” “extra movement–force,” and “slow to respond” might be problematic.

Discussion
This study investigated the psychometric properties of the SPS Assessment. On the basis of item analyses, the number of subtests and items in Version 1.0 was reduced, and a new version was constructed and evaluated for psychometric integrity. The SPS Assessment Version 2.0 was documented to be reliable and valid in characterizing sensory modulation challenges across seven sensory domains. The behavioral coding system is accurate in reflecting GCI of expert occupational therapy examiners. In addition, the subtests depict the proposed characteristics of sensory modulation dysfunction.

Further study using this new performance measure is needed to investigate the phenotypic variation in sensory modulation dysfunction and to define sensory-related symptoms across a wide variety of diagnostic groups and ages. In addition, studies are recommended to determine whether the SPS Assessment is appropriate for measuring change after intervention.

Reliability and Validity
Strong internal consistency reliability was shown for children ages 4–18 yr, suggesting that the subtests are homogeneous and measure the same construct. Construct validity was supported by differentiation between children with sensory modulation challenges and typically developing control children. SOR, SUR, and SC behaviors were compared between the two groups, and because the clinical group had more symptoms of these behaviors, the result can be considered to provide evidence of the test’s construct validity. Group differences on all sensory domain scores were significant, with large effect sizes.

Behavioral Scoring Categories
This study also produced a successful categorization of behaviors reflective of sensory modulation dysfunction. On the basis of examiner feedback, categories required tightened definitions so as to reduce ambiguity and misinterpretation and produce discrete, nonoverlapping descriptions. The correlations within proposed subtypes

<table>
<thead>
<tr>
<th>Behaviors</th>
<th>SOR</th>
<th>SUR</th>
<th>Distractible</th>
<th>SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concern</td>
<td>.83</td>
<td>.73</td>
<td>.57</td>
<td>.52</td>
</tr>
<tr>
<td>Avoidance–withdrawal</td>
<td>.69</td>
<td>.69</td>
<td>.66</td>
<td>.50</td>
</tr>
<tr>
<td>Noisy–talkative</td>
<td>.69</td>
<td>.69</td>
<td>.69</td>
<td>.50</td>
</tr>
<tr>
<td>Low energy–weak body</td>
<td>.69</td>
<td>.69</td>
<td>.69</td>
<td>.50</td>
</tr>
<tr>
<td>Extra movement–force</td>
<td>.69</td>
<td>.69</td>
<td>.69</td>
<td>.50</td>
</tr>
<tr>
<td>Unaware–minimal interaction</td>
<td>.69</td>
<td>.69</td>
<td>.69</td>
<td>.50</td>
</tr>
<tr>
<td>Slow to respond</td>
<td>.69</td>
<td>.69</td>
<td>.69</td>
<td>.50</td>
</tr>
</tbody>
</table>

Note. SC = sensory craving; SOR = sensory overresponsivity; SUR = sensory underresponsivity.
and the factor analysis suggest that the behavioral categories reflect different sensory modulation response patterns.

On the basis of these analyses and examiner feedback, the 10 atypical behavioral categories were reduced to 6, and clarification and revisions were made to the operational definitions. Specifically, we clarified “unaware–minimal interaction” and “slow to respond” as not describing a coping strategy for SOR. We redefined “noisy–talkative” to reflect SC rather than a stress reaction to an adverse sensory experience. Although “low energy–weak body” factored with other SUR categories, the consensus among examiners was that it was more related to postural abilities than to sensory modulation because the description included such behaviors as “appears weak,” “leans or slouches,” and “cannot keep body upright.” Thus, it was deleted. Feedback also suggested that “extra movement–force” was interpreted as a proprioceptive discrimination problem resulting from underresponsivity. Therefore, it was not unexpected that “extra movement–force” correlated with “low energy” and factored with the SUR categories. Examiners recommended deleting this category because it seemed unrelated to sensory modulation. In the correlation analysis by domain, “distractible” was associated only with craving behaviors in the Proprioceptive and Vestibular domains. Additionally, in the factor analysis, “distractible” appeared to be a separate dimension. Therefore, it was deleted from the sensory modulation scoring system. The two behavioral categories “avoidance–withdrawal” and “concern” were highly correlated, so they were combined.

**Description of Version 2.0**

The SPS Assessment Version 2.0 was finalized, reducing the subtests from 34 to 27 and the items from 131 to 72. Transition responses are included at the end of each sensory domain. Specific materials are left on the table during each domain administration to elicit SC behaviors. The scoring system was simplified to include six behavioral categories: “concern,” “adverse,” “unaware,” “slow to respond,” “wants more,” and “talkative.”

The current administration takes approximately 1 hr to complete. Plans include expansion of the SPS Assessment to include items for postural control, praxis, and discrimination as well as additional studies of interrater and test–retest reliability.

Unlike informant-based scales of sensory processing, the SPS Assessment provides structured opportunities and specific scoring criteria on which to base interpretations of typical versus atypical behavioral responses to sensory input. A benefit of the activities is that they closely resemble sensory experiences in daily life (e.g., putting on an adhesive bandage, playing with a gooey substance, trying to work when there is background noise).

**Use in Characterizing Clinical Populations**

The new *DSM–5* criteria for ASD underscore the need for performance-based measures of sensory modulation to supplement existing parent- and teacher-report questionnaires. The sensory features highlighted in the *DSM–5* parallel those tested on the SPS Assessment: “hyper or hypo reactivity to sensory input” (American Psychiatric Association, 2013, p. 50), reflecting the SOR and SUR subtypes, and “unusual interest in sensory aspects of environment,” reflecting SC. Although the SPA (Baranek et al., 2007, 2013) contributes to our understanding of developmentally related changes in sensory modulation, the impact of context on sensory modulation, and the ways sensory modulation relates to the acquisition of broader developmental skills and abilities (Baranek et al., 2013), the SPA is limited by the age range, sensory domains, and subtypes it purports to measure.

The SPS Assessment will be useful in describing the phenotypic variation within sensory modulation dysfunction. Other clinical groups are suspected of having sensory modulation problems (Dar et al., 2012; Reynolds & Lane, 2009), such as ADHD (Cheung & Siu, 2009; Dunn & Bennett, 2002; Lane et al., 2010; Mangeot et al., 2001; Parush et al., 2007; Yochman, Parush, & Ornoy, 2004); however, the research is inconclusive, in part because of limitations in standard performance measures (Ghanizadeh, 2011). The SPS Assessment has the potential to help characterize sensory modulation challenges across a wide range of clinical disorders and to differentiate patterns of sensory impairments among clinical groups. It will be a useful tool for specifying symptom characteristics indicative of clinical impairment. Thus, further study of sensory modulation in ADHD and other clinical disorders will benefit from this examiner-administered measure.

**Implications for Occupational Therapy Practice**

In future studies, the SPS Assessment can be used to explore the concurrence between respondent-based information and performance-based information as well as...
the extent to which combining respondent and performance data adds predictive accuracy to the diagnosis of sensory modulation challenges. Few standardized methods exist for identifying and confirming sensory modulation challenges. A performance-based measure such as the SPS Assessment provides objective information aiding the clinician’s understanding of children’s sensory challenges.

In conclusion, the major implications for practice include the potential to
- Corroborate findings from informant-based tools;
- Provide a comprehensive assessment of the child;
- Explore the relation between sensory symptoms and participation in daily life;
- Help formulate appropriate interventions; and
- Measure outcomes related to sensory modulation.

**Limitations and Future Research**

This is the first in a series of studies designed to expand and refine the assessment of sensory modulation. Limitations include relatively small sample sizes that prevent analyses by modulation subtype, age, and gender. This study used a mixed clinical sample. Thus, future studies should determine condition-specific manifestations of sensory modulation symptoms. Future studies should compare parent reports of sensory symptoms with an examiner-administered scale. Moreover, although sensory processing challenges have been linked to difficulties in multiple aspects of daily life (Bar-Shalita et al., 2008; Cohn et al., 2000; Cosbey et al., 2010), studies are needed to validate the relationship between the SPS Assessment and broader social, emotional, and developmental outcomes, as well as measures of adaptive functioning and behavior. Last, studies using the SPS Assessment as an outcome measure would be helpful in determining its potential usefulness in effectiveness research.

**Conclusion**

This study makes an important contribution to the study of sensory modulation challenges. It proposes a standard method of assessment and examines functioning across all sensory domains. Preliminary evidence supports the psychometrics of the scale. The SPS Assessment can be used to expand knowledge of patterns of sensory modulation dysfunction, identify homogeneous groups for research, facilitate improved treatment planning, and evaluate questions about treatment effectiveness. Further research and standardization of this scale are warranted to achieve these goals. ▲

**Acknowledgments**

We thank the children and families who participated in this study and the testers located at the STAR Center; Virginia Commonwealth University; The Rose F. Kennedy Center; University of California, San Francisco; Ohio State University; and Washington University in St. Louis. We also thank our research assistant, Shannon Hampton, for her tireless efforts on this article. Primary funding for this work was provided by the Wallace Research Foundation. Additional support was provided by National Institutes of Health Grant R21 HD41614-01 to Lucy Jane Miller and by National Institutes of Health Grant T32MH15442 to the University of Colorado, Denver, in support of Sarah A. Schoen.

**References**


