Please use the following citation when referencing this work:

McGill, R. J. (in press). Test review of Sensory Processing Measure, Second Edition. In J. F. Carlson, K. F. Geisinger, J. L. Jonson, & N. A. Anderson (Eds.), *The twenty-second mental measurements yearbook* (pp. xx-xx). Buros Center for Testing.

Review of the Sensory Processing Measure, Second Edition by Ryan J. McGill, Associate

Professor of School Psychology, William & Mary School of Education, Williamsburg, VA:

Description

The Sensory Processing Measure, Second Edition (SPM-2; Parham et al., 2021a) is a revision and expansion of the Sensory Processing Measure (SPM; Parham et al., 2007) and the Sensory Processing Measure-Preschool (SPM-P; Ecker et al., 2010). The SPM-2 is designed to theoretically cohere with the Ayres Sensory Integration® theory (Ayers, 1972). According to the Technical Manual (Parham et al., 2021b), "this theory postulates that how the central nervous system integrates and responds to sensations, from outside and within one's own body, influences physiological, regulatory, emotional, motor, and adaptive functioning (p. 1). The SPM-2 is designed to align with the previously aforementioned theory's principles of assessment which are integrated into the structure of the instrument. Specifically, the assessment of sensory systems, sensory integration vulnerabilities, and assessment across multiple environments and raters.

The measure is best considered as a suite of different rating forms that can be used according to the needs of the user. To wit, at the Infant\Toddler age level (4 to 30 Months) there is an Infant Form, a Toddler Form, and a Caregiver Self-Report Form. At the Child-Age level (5 to 12 years) there are Home and School report forms as well as several briefer environment specific forms that can be completed by appropriate professionals who encounter the examinee in those particular areas (e.g., Art, School Bus, Cafeteria, Music, Physical Education, and Recess). At the Adolescent age level (12 to 21 years), a Self-Report Form is available in addition

to the Home and School Forms as well as brief adolescent Driving Form. Finally, at the Adult

age level (21 to 87 years), a Rater report is provided for caregivers and significant others as well

as a Self-Report and Adult Driving Environmental Form.

Each of the main forms contain 80 items that measure eight clinical scales (Vision,

Hearing, Touch, Taste and Smell, Body Awareness, Balance and Motion, Planning and Ideas

[PLN], and Social Participation [SOC] as well as an omnibus total score. However, the PLN and

SOC scales do not contribute to the measurement of the global composite score. Conversely, the

environment-specific short forms contain between 15-18 items and do not yield conventional

standardized scores; but instead, use prescribed cutoff scores based on the sum of total items

evaluated.

Development

The goals for the revision that are highlighted in the Manual include expanding the age

range across the lifespan, introduce new and revised Infant/Toddler Forms, conduct additional

clinical validity studies, and update the norms based on a new nationally representative sample.

It is noted in the Manual that the Preschool and Child age forms were not piloted but were

evaluated through expert judgement which resulted in a small number of item deletions and

editing. For all forms, Rasch modeling was used to identify potentially problematic items during

the development of the instrument. For the Adolescent and Adult scales, research forms were

initially developed containing between 108 and 123 items that were drawn from item pools and

evaluated for potential inclusion in the final rating forms which resulted in 80 items per form (10

per scale).

SPM-2 forms are written at approximately an eighth-grade reading level and users are

permitted to read items to raters who are unable to read at that level though, all items and

directions must be read verbatim as they are described in the forms. Rating forms can be

administered via paper and pencil versions as well as online and are accompanied with scoring

summary sheets and a Manual which contains multiple case examples to aide users with clinical

interpretation. Users can also supplement their interpretation of the instrument through the use of

a Quick Tips™ guide which was not made available as part of this review; though, the case

examples in the Manual illustrate the use of some of those features such as the application of

prescribed intervention recommendations based on the strengths and weaknesses that are

observed among the SPM-2 scores.

Technical

Standardization

The SPM-2 standardization sample consisted of 3,850 participants ranging in age from

four months to 87 years. All participants were assessed via, at least, one of the previously

mentioned forms associated with their age range. Inspection of the demographic tables reported

in the Manual reveal appropriate matching on variables such as gender, race/ethnicity, parent

educational level as a proxy for socioeconomic status, and geographic region in accord with

target estimates furnished from the 2017 U. S. Census data.

All scale and composite scores are expressed as T scores ($M = 50$, $SD = 10$, range $= 40$ to

80) with higher scores indicating increasingly higher levels of perceived difficulty from the rater.

According to the Manual, T scores between 60-69 are indicative of *moderate difficulties* and

scores between 70-80 represent *severe difficulties*. In terms of interpretation, clinicians are

encouraged to utilize SPM-2 scores to aide in the identification of both normative and

idiographic sensory processing strengths and weaknesses. A unique feature of the SPM-2 is that

the summary score sheet at each age level provide users a means to evaluate the consistency of

scores across raters (expressed as conventional deviation scores). However, interpretive

guidelines follow conventional statistical interpretation of T scores (i.e., a difference score

equaling or exceeding 10 is regarded as a *probable* to *definite* difference among different raters)

though no base rates associated the standardization sample are for those rules of thumb are

provided. Given the historically low correlations between raters on multi-informant scales (e.g.,

Achenbach et al., 1987), this information would better permit users to evaluate the potential

utility of those proposed interpretive guidelines.

Reliability

Multiple methods for estimating reliability of SPM-2 scores are reported in the Manual

including but not limited to internal consistency, test-retest stability, and interscorer agreement.

Inspection of the internal consistency tables revealed estimates that varied from weak (.68) to

strong (.98) depending on age, form, and the type of score (i.e., scale versus composite).

Whereas, at all age ranges, the Sensory Total score can be confidentially interpreted, the

coefficients for the subscales vary across the age range of the measure though most of the adult

scores exceed recommended levels for confidant clinical interpretation. Average test-retest

correlations ranged from .71 to .94 which indicate relatively strong consistency. Though, it

should be noted that the retest interval was limited to two-weeks which may not be sufficient for

capturing temporal change in target behaviors. Median interrater reliability correlations ranged

from .63 to .78, which is consistent with contemporary rating scale assessment.

Validity

Consistent with the Standards for Educational and Psychological Testing (American

Educational Research Association, American Psychological Association, & National Council on

Measurement in Education, 2014), evidence for validity was structured around the areas of test

content, internal structure, and relations with other variables. Confirmatory factor analysis was

conducted to ascertain the structure of the SPM-2. In the Manual, it is reported that analyses

were limited to examining the improvement in the fit of an eight-factor (consistent with the eight

SPM-2 scales) as compared to a one-factor model and only the fit statistics associated with the former are reported which does not permit inspection of local fit for potential identification issues. Nevertheless, the fit statistics in Table 5.13 generally support the final model that was retained with some exceptions (e.g., Infant/Toddler, and the Adult self-report form). Special groups and matched controls were compared to test for specific group differences. Mean score differences between the groups were typically what one would expect with individuals diagnosed with various conditions (e.g., ADHD, sensory processing disorder) at the Infant, Child, and Adolescent age ranges; however, a decline effect was observed during adulthood where the differences seemed to evaporate. To support these findings, additional diagnostic efficiency statistics were provided to establish optimal diagnostic thresholds for accurately diagnosing sensory processing disorder via the Sensory Total Score.

Commentary

As with any test the SPM-2 has strengths and weaknesses. Most notably, the structural validity analyses do not fully disclose information necessary for determining whether the preferred eight-factor, model cohering with Ayres' theory, is viable. In particular, it remains unknown the degree to which the eight-factor model, that was estimated, is consistent with the *actual* interpretive structure of the test as a higher-order factor of sensory processing is implied by the provision of a total composite score. In terms of diagnostic efficiency, the Manual notes that "Disorders of sensory integration and processing represent the clinical conditions that are most frequently targeted by the SPM-2 sensory systems scales" (p. 133). As such, it appears in Table 5.35 that the most optimal cut-point for demarcating those with and without such disorders is optimally placed at the *moderate difficulties* range as opposed to the *severe difficulties* range where specificity was high (.98) but sensitivity to detect a true effect was relatively low (.23) and trended lower at higher cut-points. As noted by Swets (1966), every diagnostic test features a

different cut-point at which there is an optimal balance between sensitivity and specificity. Users are encouraged to consider the results reported in Table 5.35 (p. 133) accordingly.

Summary

Overall, the SPM-2 has many strengths. The test materials are easy to access and navigate and the rating forms are relatively undemanding to score using the conventional AutoScore™ forms. The Manual provides users with novel case examples that illustrate the clinical versatility that is possible via the scores afforded by the instrument across raters and environments.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232. https://doi.org/10.1037/0033-2909.101.2.213

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Ayers, A. J. (1972). *Sensory integration and learning disorders*. Western Psychological Services.

Ecker, C., Parham, L.D., Miller-Kuhaneck, H., Henry, D. A., & Glennon, T. J. (2010). *Sensory Processing Measure, Preschool*. Western Psychological Services.

Parham, L. D., Ecker, C. L., Kuhaneck, H., Henry, D. A., & Glennon, T. J. (2021a). *Sensory Processing Measure* (2nd ed.). Western Psychological Services.

Parham, L. D., Ecker, C. L., Kuhaneck, H., Henry, D. A., & Glennon, T. J. (2021b). *Sensory Processing Measure* (2nd ed.) [Manual]. Western Psychological Services.

Parham, L. D., Ecker, C. L., Miller-Kuhanek, H., Henry, D. A., & Glennon, T. J. (2007). *Sensory Processing Measure*. Western Psychological Services.

Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnosis*. Lawrence Erlbaum Associates