

4

Development and Standardization

This chapter begins with a summary of Ayres Sensory Integration theory, the conceptual foundation of the SPM-2. It then reviews the development of the SPM-2 from its predecessor instruments, the Sensory Processing Measure (SPM; Parham et al., 2007) and the SPM–Preschool (SPM-P; Ecker et al., 2010). The final section of the chapter describes the research program that supported the standardization and validation of the SPM-2.

Theoretical Background: Ayres Sensory Integration Theory

Ayres Sensory Integration theory (1972, 2005) is the predominant frame of reference for occupational therapy practitioners (OTPs) who provide services for clients with sensory integration and processing problems. This theory also provides the conceptual foundation of the SPM-2.

The term *sensory integration* refers broadly to the central nervous system processes by which incoming sensory data are organized within each sensory system, as well as across different sensory systems (Ayres, 1972, 2005). In contemporary practice, the terms *sensory integration* and *sensory processing* are often used interchangeably. For example, the American Occupational Therapy Association (AOTA, 2018) has published Practice Guidelines that treat “sensory integration” and “sensory processing” as

synonymous terms. Thus, they are also used interchangeably throughout this manual.

One of Ayres’ fundamental ideas is that the early developing, body-centered senses (tactile, proprioceptive, and vestibular) provide a foundation for the development of the later-maturing visual and auditory systems (Ayres, 1972, 2005). In her synthesis of neurobiological and developmental research, Ayres theorized that development and integration of the tactile, proprioceptive, and vestibular systems allow for the formation of body schema, object concepts, and body-centered spatial mapping of the environment. These elementary functions eventually become automatized and serve as a platform for the layering of more complex auditory and visual functions (Ayres, 2005). Based on the premise of

neuroplasticity in the young and developing brain, sensory integration theory holds that many of the child's sensory integration functions improve with individually tailored therapeutic activities that provide opportunities for exploration and challenge (Parham & Mailloux, 2020).

Ayres (1979) defined sensory integration as the “organization of sensation for use” (p. 184). The last two words in that definition, “for use,” are revealing. In contrast to the study of the neural mechanisms of sensory integration in isolation, Ayres' central

concern as an occupational therapist was the relationship between these neural processes and a child's engagement in everyday life occupations, such as playing, eating, and participating in classroom activities. This concern with occupation strongly influenced the evolution of Ayres' ideas about assessment and intervention. Sensory integration–based occupational therapy became a tool for helping children to engage in occupations and, in so doing, to participate as fully as possible in meaningful life activities.

Development of the SPM-2: An Integrated Instrument

The SPM-2 is a revision and expansion of the SPM and the SPM-P, two norm-referenced assessments of sensory processing, praxis, and social participation in children. These instruments, in turn, were developed from the Evaluation of Sensory Processing (ESP; Johnson-Ecker & Parham, 2000), a domain-based sensory integration dysfunction assessment for parents, and the School Assessment of Sensory Integration (SASI; Miller-Kuhaneck et al., 2007), a multi-environment measure of sensory processing and praxis for teachers and school personnel (Glennon et al., 2011; Miller-Kuhaneck & Henry, 2009).

The SPM-2 revision encompasses the age ranges of the SPM (5 to 12 years) and the SPM-P (2 to 5 years), and extends the range downward to a new Infant/Toddler age level (4–30 months) and upward to two new age levels, Adolescent (12 to 21 years) and Adult (21 to 87 years).

The SPM-2 forms were developed, first, by revising the original SPM and SPM-P forms (now called SPM-2 Child and SPM-2 Preschool) to ensure the following:

- The items in each scale continued to represent sensory integration and processing across relevant content facets
- The items represent functioning in the sensory systems (visual, auditory, tactile, olfactory, gustatory, proprioceptive, and vestibular), as well as the more complex functions of praxis and social participation

- The items represent a range of sensory integration vulnerabilities, including over-reactivity, under-reactivity, sensory seeking, and perception difficulties

The forms for the Infant/Toddler, Adolescent, and Adult age levels were created by adapting the item content from the Child and Preschool age levels to the differing sensory-related environments and behaviors associated with each age range. This was accomplished by including the revised SPM-2 Preschool and Child items as appropriate, either verbatim or in an adapted form, and by writing new items that tapped the unique sensory-processing inputs and behaviors for each new age level, reporting format, and environment. The I/T Caregiver Self-Report Form was created by adapting the content of the Adult age-level items to include stimuli associated with the care of infants and toddlers.

The resulting SPM-2 item pools were subjected to an iterative process of development, analysis, and selection leading to the final sets of items represented in the published forms.

New age levels The forms for the new Infant/Toddler, Adolescent, and Adult age levels were piloted on approximately 150 participants to analyze and select those items with the best psychometric characteristics for inclusion in the standardization research forms. These procedures included 1) psychometric analyses based on classical test theory (e.g., item-scale correlations); 2) item response theory

(e.g., developing a Rasch scale of measurement and examining the distributions of item and person thetas); and 3) judgments by experts in sensory modulation and perception, to exclude or modify underperforming items.

Existing age levels The forms for the Preschool and Child age levels, which have substantial prior validation, were not piloted but were evaluated through expert judgment for a small number of item deletions, light editing to update terminology and harmonize phrasing across age levels, and the addition of new candidate items prior to standardization.

For all forms, Rasch-based differential item functions were used to eliminate items that were biased for gender, ethnicity, and socioeconomic status.

The SPM-2 standardization research forms each included between 108 and 123 items (12 to 16 items per scale) drawn from the item pools. These forms

were administered to the standardization sample and then subjected to an additional round of psychometric analysis and expert review, after which the 80 (10 per scale) best-performing items were selected for inclusion in the published forms. Chapter 5 gives additional information about the content validity and facet structure of the SPM-2 forms.

Like the SPM and the SPM-P, the SPM-2 uses a 4-point Likert-type scale that refers to frequency of behavior: *Never* = 1, *Occasionally* = 2, *Frequently* = 3, *Always* = 4, with items that indicate positive behaviors or traits, such as appropriate social participation, reverse-scored, so that *Never* = 4, *Occasionally* = 3, *Frequently* = 2, *Always* = 1. This response scale, investigated during the development of the SPM and the SPM-P, meets criteria for validity based on item response theory and the Rasch model, and comports with accepted standards for clinical test item response processes (e.g., Frick et al., 2010).

Standardization Study

Standardization Sample Characteristics

The SPM-2 standardization sample consisted of 3,850 participants, ranging in age from 4 months to 87 years. All participants were assessed with at least one SPM-2 Research Form. One hundred and nine SPM-2 study coordinators from across the United States contributed data to the standardization study. Normative samples were collected from typically developing children and adults. For example, children were included in the normative sample if they attended school without the assistance of an aide and were not enrolled in special needs or special education programs. However, no attempt was made to exclude participants with mild academic, occupational, psychological, or behavioral difficulties. To be considered demographically representative, the standardization sample should reflect the base rate of such mild problems in the general population.

Tables 4.1 through 4.5 present the demographic characteristics of the standardization sample, along with corresponding percentages from the U.S. Census (U.S. Census Bureau, 2017). For each SPM-2 age level (Infant/Toddler, Preschool, Child, Adolescent, and Adult), separate census-based demographic targets were developed based on the characteristics of the U.S. population within those age ranges. In general, an adequate sample population percentage falls within 5% of its corresponding census percentage (Andersson, 2005). Across the five samples, after random desampling from oversampled demographic groups, all but one target characteristic met this standard: at the Adolescent age level, participants in the Northeast region of the U.S. were slightly underrepresented.

Table 4.1. Demographic Characteristics of the SPM-2 Standardization Sample:
Infant/Toddler Age Level Forms (Ages 4–30 Months)

Characteristic ^a	<i>n</i>	% of sample	U.S. Census % ^b
Gender			
Male	348	49.9	51.1
Female	349	50.1	48.9
Race/Ethnicity^c			
Asian	21	3.0	4.5
Black/African American	89	12.8	13.4
Hispanic Origin	182	26.1	25.7
Native Hawaiian/Pacific Islander	1	0.1	0.2
American Indian/Alaska Native	2	0.3	0.7
White	340	48.8	50.1
Other	62	8.9	5.4
Parents' educational level			
Less than high school diploma	55	7.9	11.8
High school graduate	172	24.7	24.2
Some college	237	34.0	31.0
4-year college degree or more	233	33.4	33.1
U.S. geographic region			
Northeast	101	14.5	16.1
South	235	33.7	38.2
Midwest	176	25.3	21.1
West	185	26.5	24.6

Note. *N* = 697. Due to rounding, total percentages may not equal 100.0%.

^aDemographic characteristics are given for the infant and toddler participants (4–30 months); caregiver characteristics were incidental to the infant and toddler recruitment and were similar to the Adult age level forms sample.

^bU.S. Census Bureau (2017). Race/Ethnicity, gender, and region based on ages 0–3 years; parents' educational level based on ages 25–44 years (those most likely to have children ages 0–3 years).

^cIndividuals of Hispanic origin are included in the race/ethnicity category under Hispanic Origin; the remaining race/ethnicity categories include only individuals of non-Hispanic origin. Individuals of two or more races (*n* = 56) are included in the Other category.

Table 4.2. Demographic Characteristics of the SPM-2 Standardization Sample:
Preschool Age Level Forms (Ages 2–5 Years)

Characteristic ^a	<i>n</i>	% of sample	U.S. Census % ^b
Gender			
Male	308	50.8	51.1
Female	298	49.2	48.9
Race/Ethnicity^c			
Asian	24	4.0	4.5
Black/African American	92	15.2	13.7
Hispanic Origin	153	25.2	25.8
Native Hawaiian/Pacific Islander	3	0.5	0.2
American Indian/Alaska Native	1	0.2	0.8
White	292	48.2	49.9
Other	41	6.8	5.2
Parents' educational level			
Less than high school diploma	68	11.2	11.8
High school graduate	145	23.9	24.2
Some college	187	30.9	31.0
4-year college degree or more	206	34.0	33.1
U.S. geographic region			
Northeast	99	16.3	16.0
South	231	38.1	38.5
Midwest	125	20.6	21.1
West	151	24.9	24.5

Note. *N* = 606. Due to rounding, total percentages may not equal 100.0%.

^aDemographic characteristics are given for the child participants (2–5 years).

^bU.S. Census Bureau (2017). Race/Ethnicity, gender, and region based on ages 2–5 years; parents' educational level based on ages 25–44 years (those most likely to have children ages 2–5 years).

^cIndividuals of Hispanic origin are included in the race/ethnicity category under Hispanic Origin; the remaining race/ethnicity categories include only individuals of non-Hispanic origin. Individuals of two or more races (*n* = 32) are included in the Other category.

Table 4.3. Demographic Characteristics of the SPM-2 Standardization Sample:
Child Age Level Forms (Ages 5–12 Years)

Characteristic ^a	<i>n</i>	% of sample	U.S. Census % ^b
Gender			
Male	421	50.1	51.1
Female	419	49.9	48.9
Race/Ethnicity^c			
Asian	20	2.4	4.6
Black/African American	113	13.5	13.5
Hispanic Origin	204	24.3	24.3
Native Hawaiian/Pacific Islander	2	0.2	0.2
American Indian/Alaska Native	5	0.6	0.8
White	436	51.9	52.0
Other	60	7.1	4.6
Parents' educational level			
Less than high school diploma	63	7.5	11.9
High school graduate	239	28.5	25.7
Some college	276	32.9	30.6
4-year college degree or more	262	31.2	31.7
U.S. geographic region			
Northeast	106	12.6	16.2
South	350	41.7	38.1
Midwest	200	23.8	21.5
West	184	21.9	24.2

Note. *N* = 840. Due to rounding, total percentages may not equal 100.0%.

^aDemographic characteristics are given for the child participants (5–12 years).

^bU.S. Census Bureau (2017). Race/Ethnicity, gender, and region based on ages 5–12 years; parents' educational level based on ages 25–54 years (those most likely to have children ages 5–12 years).

^cIndividuals of Hispanic origin are included in the race/ethnicity category under Hispanic Origin; the remaining race/ethnicity categories include only individuals of non-Hispanic origin. Individuals of two or more races (*n* = 51) are included in the Other category.

Table 4.4. Demographic Characteristics of the SPM-2 Standardization Sample:
Adolescent Age Level Forms (Ages 12–21 Years)

Characteristic ^a	<i>n</i>	% of sample	U.S. Census % ^b
Gender			
Male	516	52.4	51.2
Female	469	47.6	48.8
Race/Ethnicity^c			
Asian	20	2.0	4.6
Black/African American	144	14.6	14.3
Hispanic Origin	241	24.5	21.9
Native Hawaiian/Pacific Islander	1	0.1	0.2
American Indian/Alaska Native	9	0.9	0.8
White	504	51.2	54.6
Other	66	6.7	3.6
Parents' educational level			
Less than high school diploma	97	9.8	11.9
High school graduate	282	28.6	26.4
Some college	316	32.1	30.5
4-year college degree or more	290	29.4	31.2
U.S. geographic region			
Northeast	116	11.8	17.2
South	365	37.1	37.5
Midwest	254	25.8	21.5
West	250	25.4	23.7

Note. *N* = 985. Due to rounding, total percentages may not equal 100.0%.

^aDemographic characteristics are given for the adolescent participants (12–21 years).

^bU.S. Census Bureau (2017). Race/Ethnicity, gender, and region based on ages 12–21 years; parents' educational level based on ages 25–64 years (those most likely to have children ages 12–21 years).

^cIndividuals of Hispanic origin are included in the race/ethnicity category under Hispanic Origin; the remaining race/ethnicity categories include only individuals of non-Hispanic origin. Individuals of two or more races (*n* = 54) are included in the Other category.

Table 4.5. Demographic Characteristics of the SPM-2 Standardization Sample: Adult Age Level Forms (Ages 21–87 Years)

Characteristic ^a	<i>n</i>	% of sample	U.S. Census % ^b
Gender			
Male	347	48.1	48.8
Female	375	51.9	51.2
Race/Ethnicity^c			
Asian	35	4.8	5.3
Black/African American	95	13.2	11.9
Hispanic Origin	116	16.1	15.1
Native Hawaiian/Pacific Islander	3	0.4	0.2
American Indian/Alaska Native	3	0.4	0.6
White	448	62.0	65.2
Other	22	3.0	1.7
Parents' educational level			
Less than high school diploma	66	9.1	13.4
High school graduate	212	29.4	28.0
Some college	254	35.2	31.4
4-year college degree or more	190	26.3	27.2
U.S. geographic region			
Northeast	129	17.9	18.1
South	285	39.5	37.3
Midwest	127	17.6	21.3
West	181	25.1	23.3

Note. *N* = 722. Due to rounding, total percentages may not equal 100.0%.

^aDemographic characteristics are given for the adult participants (21–87 years).

^bU.S. Census Bureau (2017). Demographic characteristics are based on the general adult population.

^cIndividuals of Hispanic origin are included in the race/ethnicity category under Hispanic Origin; the remaining race/ethnicity categories include only individuals of non-Hispanic origin. Individuals of two or more races (*n* = 16) are included in the Other category.

Final Item Composition

Data from the standardization studies were used to determine the final item composition of the SPM-2 forms at every age level. Tables 4.6 through 4.10 present descriptive statistics for each form in the standardization sample. See Chapter 5 for a description of the development of each form in detail.

Environment Forms

A subsample of children who were rated by their main classroom teacher in the SPM-2 Child standardization sample were also rated by one or more additional school staff members using the School

Environment Forms. Similarly, a subsample of adolescent and adult drivers from the standardization sample were rated by themselves or others using the Driving Environment Self-Report and Rater Report Forms. The descriptive statistics for the Environment forms are listed in Table 4.11.

Moderator Variables

When demographic groups (e.g., those defined by age, gender, ethnicity, and socioeconomic status) differ meaningfully in test performance, it can be difficult to interpret the results without separate norms for each group. Such differences are said to reflect the *moderating effects* of the demographic factors.

An *effect size* metric was used to determine whether the demographic moderator variables were associated with any clinically meaningful differences between groups in the SPM-2 raw scale scores. The effect size was calculated as the difference between the mean raw scale score of the demographic group in question and the grand mean of the standardization sample, divided by the pooled standard deviation. Using this method, an effect size of 0.2 is considered small, 0.5 is considered medium, and 0.8 is considered large (see Cohen, 1992). In general, moderating effects are considered clinically meaningful only if they are associated with medium-to-large effect sizes and if the findings are consistent with other knowledge of the groups in question.

Age Preliminary analyses suggested a statistically significant effect of age within two age levels, Infant/Toddler and Preschool. To investigate whether this was a clinically meaningful effect, each standardization sample was split into younger and older age groups, and effect sizes were calculated for the scales.

A pattern of medium-to-large effect sizes was seen on both the Toddler Form and the Preschool School Form. On the Toddler Form, the scores of 21- to 30-month-olds were lower than those of 10- to 20-month-olds by a clinically meaningful margin; likewise, on the Preschool School Form, the scores of 5-year-olds were lower than those of 2- to 4-year-olds by a clinically meaningful margin. These findings pointed to the need for separate norms for both age groups. In fact, the differences observed in the standardization sample probably reflect a typical developmental progression in behavioral regulation skills of younger and older children in both groups.

To keep the scoring systems parallel within each age level, age-stratified norms were extended to other

forms in each group as well. At the Infant/Toddler age level, the Infant Form was also stratified so that younger (4–6 months) and older (7–9 months) infants would be compared to their own developmental group. At the Preschool age level, the age-stratified norms were extended from the School to the Home Form.

The age-stratified normative structure of the SPM-2 is documented in Appendix A, which includes standard score conversion tables for every form.

Other moderator variables The age group comparisons stand in contrast to the analyses of all other variables. In terms of gender, none of the differences reached even the level of a small effect size, indicating that any differences were not clinically meaningful. In analyzing the moderating effects of ethnicity, only the Black and Hispanic ethnic groups had sample sizes large enough to allow for comparison. The effect sizes for these two groups were small and the direction of these effects was not consistent, indicating that these differences were not clinically meaningful. Effect sizes based on socioeconomic status (SES), represented in the standardization sample by education level of either the participant or the participant's parent, depending on the age level, were likewise small and showed no clear directional effect for most SES groups. As with the other moderator variables, therefore, the effect sizes associated with SES were not clinically meaningful.

In summary, the analyses of demographic moderators support the use of age-stratified norms when interpreting the raw scores for the SPM-2 Infant/Toddler and Preschool age levels. In contrast, differences in SPM-2 scores due to gender, ethnicity, and SES were small and not likely to have any meaningful impact on the interpretation of test results.

Table 4.6. Descriptive Statistics for SPM-2 Raw Scores in the Standardization Sample:
Infant/Toddler Age Level Forms

Scale	No. of items	Mean	SD
Infant Form			
Vision (VIS)	10	20.3	4.5
Hearing (HEA)	10	14.4	3.0
Touch (TOU)	10	14.6	2.9
Taste and Smell (T&S)	10	11.9	2.6
Body Awareness (BOD)	10	14.7	3.2
Balance and Motion (BAL)	10	14.2	3.2
Sensory Total (ST)	60	90.0	13.5
Planning and Ideas (PLN)	10	16.6	6.0
Social Participation (SOC)	10	15.2	4.5
Toddler Form			
Vision (VIS)	10	15.9	3.7
Hearing (HEA)	10	15.4	3.6
Touch (TOU)	10	14.6	3.2
Taste and Smell (T&S)	10	13.2	3.1
Body Awareness (BOD)	10	15.4	3.6
Balance and Motion (BAL)	10	14.0	3.3
Sensory Total (ST)	60	88.4	15.7
Planning and Ideas (PLN)	10	13.4	4.6
Social Participation (SOC)	10	16.3	4.5
Caregiver Self-Report Form			
Vision (VIS)	10	14.7	3.4
Hearing (HEA)	10	14.8	3.9
Touch (TOU)	10	12.8	2.5
Taste and Smell (T&S)	10	15.0	3.8
Body Awareness (BOD)	10	11.9	2.3
Balance and Motion (BAL)	10	13.5	3.0
Sensory Total (ST)	60	82.6	14.1
Planning and Ideas (PLN)	10	15.4	3.8
Social Participation (SOC)	10	16.9	3.8

Note. Infant Form $n = 266$; Toddler Form $n = 422$; Caregiver Self-Report Form $n = 870$. The ST scale is a composite of the VIS, HEA, TOU, T&S, BOD, and BAL scales.

Table 4.7. Descriptive Statistics for SPM-2 Raw Scores in the Standardization Sample:
Preschool Age Level Forms

Scale	No. of items	Mean	SD
Home Form			
Vision (VIS)	10	15.9	4.7
Hearing (HEA)	10	15.5	5.1
Touch (TOU)	10	14.5	4.3
Taste and Smell (T&S)	10	14.1	4.0
Body Awareness (BOD)	10	15.9	5.0
Balance and Motion (BAL)	10	13.0	3.4
Sensory Total (ST)	60	88.8	21.9
Planning and Ideas (PLN)	10	15.2	4.6
Social Participation (SOC)	10	19.5	5.5
School Form			
Vision (VIS)	10	13.7	3.9
Hearing (HEA)	10	13.7	4.1
Touch (TOU)	10	12.9	3.6
Taste and Smell (T&S)	10	11.9	3.1
Body Awareness (BOD)	10	12.9	4.0
Balance and Motion (BAL)	10	12.2	3.2
Sensory Total (ST)	60	77.3	19.0
Planning and Ideas (PLN)	10	14.6	4.7
Social Participation (SOC)	10	18.6	5.9

Note. Home Form $n = 520$; School Form $n = 539$. The ST scale is a composite of the VIS, HEA, TOU, T&S, BOD, and BAL scales.

Table 4.8. Descriptive Statistics for SPM-2 Raw Scores in the Standardization Sample:
Child Age Level Forms

Scale	No. of items	Mean	SD
Home Form			
Vision (VIS)	10	13.2	3.4
Hearing (HEA)	10	13.0	4.0
Touch (TOU)	10	13.6	3.4
Taste and Smell (T&S)	10	14.1	3.9
Body Awareness (BOD)	10	13.8	3.9
Balance and Motion (BAL)	10	12.1	2.8
Sensory Total (ST)	60	79.8	17.3
Planning and Ideas (PLN)	10	14.1	4.3
Social Participation (SOC)	10	17.2	5.5
School Form			
Vision (VIS)	10	13.6	3.6
Hearing (HEA)	10	13.1	3.6
Touch (TOU)	10	11.4	2.5
Taste and Smell (T&S)	10	11.2	2.5
Body Awareness (BOD)	10	11.9	3.3
Balance and Motion (BAL)	10	12.3	3.2
Sensory Total (ST)	60	73.5	15.9
Planning and Ideas (PLN)	10	14.4	5.1
Social Participation (SOC)	10	15.4	5.4

Note. Home Form $n = 875$; School Form $n = 555$. The ST scale is a composite of the VIS, HEA, TOU, T&S, BOD, and BAL scales.

Table 4.9. Descriptive Statistics for SPM-2 Raw Scores in the Standardization Sample:
Adolescent Age Level Forms

Scale	No. of items	Mean	SD
Home Form			
Vision (VIS)	10	13.2	4.1
Hearing (HEA)	10	13.2	4.2
Touch (TOU)	10	12.9	4.0
Taste and Smell (T&S)	10	14.6	4.5
Body Awareness (BOD)	10	13.0	4.1
Balance and Motion (BAL)	10	12.1	3.8
Sensory Total (ST)	60	78.9	21.1
Planning and Ideas (PLN)	10	14.4	5.3
Social Participation (SOC)	10	17.4	4.9
School Form			
Vision (VIS)	10	11.6	3.6
Hearing (HEA)	10	12.0	3.9
Touch (TOU)	10	11.2	3.1
Taste and Smell (T&S)	10	11.5	3.5
Body Awareness (BOD)	10	11.4	3.4
Balance and Motion (BAL)	10	11.4	3.5
Sensory Total (ST)	60	69.1	19.3
Planning and Ideas (PLN)	10	14.0	5.0
Social Participation (SOC)	10	17.5	5.6
Self-Report Form			
Vision (VIS)	10	16.6	5.1
Hearing (HEA)	10	16.9	5.2
Touch (TOU)	10	15.4	4.2
Taste and Smell (T&S)	10	17.2	5.1
Body Awareness (BOD)	10	15.1	4.6
Balance and Motion (BAL)	10	14.1	4.3
Sensory Total (ST)	60	95.2	24.0
Planning and Ideas (PLN)	10	16.2	5.0
Social Participation (SOC)	10	17.5	4.3

Note. Home Form $n = 1,014$; School Form $n = 645$; Self-Report Form $n = 957$. The ST scale is a composite of the VIS, HEA, TOU, T&S, BOD, and BAL scales.

Table 4.10. Descriptive Statistics for SPM-2 Raw Scores in the Standardization Sample:
Adult Age Level Forms

Scale	No. of items	Mean	SD
Self-Report Form			
Vision (VIS)	10	16.0	4.4
Hearing (HEA)	10	18.6	5.3
Touch (TOU)	10	17.4	4.4
Taste and Smell (T&S)	10	17.4	4.5
Body Awareness (BOD)	10	14.0	4.0
Balance and Motion (BAL)	10	16.4	4.8
Sensory Total (ST)	60	99.8	21.9
Planning and Ideas (PLN)	10	16.6	4.9
Social Participation (SOC)	10	18.2	4.4
Rater Report Form			
Vision (VIS)	10	13.8	4.1
Hearing (HEA)	10	15.0	4.5
Touch (TOU)	10	14.4	3.8
Taste and Smell (T&S)	10	15.1	4.7
Body Awareness (BOD)	10	12.9	3.6
Balance and Motion (BAL)	10	13.1	4.0
Sensory Total (ST)	60	84.4	20.5
Planning and Ideas (PLN)	10	14.6	4.5
Social Participation (SOC)	10	16.7	4.6

Note. Self-Report Form $n = 604$; Rater Report Form $n = 711$. The ST scale is a composite of the VIS, HEA, TOU, T&S, BOD, and BAL scales.

Table 4.11. Descriptive Statistics and Cutoffs for SPM-2 Raw Scores in the Standardization Sample:
Child School Environment and Adolescent and Adult Driving Environment Forms

Environment form	Subsample size	No. of items	Mean	SD	Cutoff score	% \geq cutoff
Child Age Level						
School Environment						
Art (ART)	171	15	19.8	6.3	30	9.4
School Bus (BUS)	106	15	19.8	6.3	27	10.4
Cafeteria (CAF)	173	15	18.2	4.1	22	9.2
Music (MUS)	179	15	19.0	5.9	25	10.1
Physical Education (PHY)	208	15	20.9	6.2	30	10.1
Recess/Playground (REC)	170	15	19.5	5.6	25	9.4
Adolescent Age Level						
Driving Environment						
Adolescent Driving Self-Report	292	18	24.4	6.5	34	9.6
Adolescent Driving Rater Report	777	18	23.3	6.1	32	9.5
Adult Age Level						
Driving Environment						
Adult Driving Self-Report	915	18	22.7	5.5	29	10.3
Adult Driving Rater Report	679	18	21.5	4.6	27	9.7

Clinical Validity Study

A sample of 639 clinic-referred participants was included to evaluate the validity and clinical utility of the SPM-2 forms across all age groups. Participants were eligible for this study if they were being treated by an occupational therapist at the time of the assessment. Table 4.12 provides frequency counts of the types of disorders present in the SPM-2 clinical sample. Many of the participants had multiple disorders or diagnoses; however, the cases were classified by the disorder that was the primary source of problems for the individual.

Table 4.13 presents the demographic characteristics of the clinical sample. Because of the goals of the study and method of recruitment, there was no expectation that the demographic composition of the clinical sample would correspond closely to that of the general population. The sample is diverse, with 35% of participants self-identified as ethnic minorities. Males outnumber females by about 2 to 1, mirroring the preponderance of boys in actual clinical settings. The sample is distributed over the intended age range of the SPM-2.

Table 4.12. Types of Disorders in the SPM-2 Clinical Sample

Disorder group	<i>n</i>	% of sample
Attention-deficit/hyperactivity disorder	80	12.5
Autism spectrum disorder	103	16.1
Developmental disorder	18	2.8
Fetal alcohol syndrome disorder	14	2.2
Hearing impairment	48	7.5
Intellectual disability	38	5.9
Learning disability	42	6.6
Low birth weight	11	1.7
Mental health disorder ^a	71	11.1
Motor disorder ^b	37	5.8
Sensory processing dysfunction	42	6.6
Speech–language pathology	77	12.1
Traumatic brain injury	9	1.4
Visual impairment	11	1.7
Other	38	5.9

Note. *N* = 639. Due to rounding, total percentages may not equal 100.0%.

^aMental health disorders were 50% anxiety disorders, 40% mood disorders, 10% other. ^bMotor disorders were 50% cerebral palsy, 30% developmental coordination disorder, 20% other.

Table 4.13. Demographic Characteristics of the SPM-2 Clinical Sample

Characteristic	<i>n</i>	% of sample
Gender		
Male	412	64.5
Female	227	35.5
Age level		
Infant/Toddler	54	8.5
Preschool	108	16.9
Child	227	35.5
Adolescent	127	19.9
Adult	123	19.2
Race/Ethnicity		
Asian	14	2.2
Black/African American	81	12.7
Hispanic Origin	105	16.4
Native Hawaiian/Pacific Islander	1	0.2
American Indian/Alaska Native	4	0.6
White	405	63.4
Multiracial	20	3.1
Other	9	1.4
Parents' educational level^a		
Less than high school diploma	23	3.6
High school graduate	115	18.0
Some college	176	27.5
4-year college degree or more	296	46.3
U.S. geographic region		
Northeast	86	13.5
South	372	58.2
Midwest	75	11.7
West	106	16.6

Note. *N* = 639, ages 4 months–67 years. Sample size differs from total clinical sample due to the removal of outliers prior to main analyses.

^aEducational level shown for the adult participants and the parents of minor participants (*n* = 610 due to missing data).

Derivation of Standard Scores

T-Scores

To interpret the results of the SPM-2 main forms, the evaluator converts raw scale scores into normalized *T*-scores, using the procedure described in Chapter 2. To construct the normalized *T*-scores, each standardization sample was first stratified into multiple age groups, and the original distributions of SPM-2 raw scores were transformed so that each approximated a normal distribution. The normalized raw scores were transformed into *z*-scores, which were then converted to *T*-scores, which have a mean of 50 and a standard deviation of 10. The use of normalized *T*-scores means that a given *T*-score value corresponds to the same percentile rank for all SPM-2 scales.

In all SPM-2 raw-score distributions, there was a strong floor effect (i.e., the mean of the distribution was close to the lowest possible score). This type of distribution is commonly seen when clinical scales are studied in samples of typically developing individuals. Because of the floor effect, the lower bound of the raw-score-to-*T*-score-conversion tables was truncated at $T = 40$.

DIF Scores

The DIF score interpretive ranges were developed to classify DIF scores by how frequently they appeared in the SPM-2 standardization samples. Because the DIF score is the difference between two normalized

T-scores, it is also normally distributed. The interpretive logic is that larger DIF scores (positive or negative) are found on the tails of the normal distribution and are thus less frequent and more likely to represent clinically significant differences in sensory functioning between two forms. The interpretive ranges were chosen to separate the scores as follows:

- **No Difference range:** Relatively frequent DIF scores, which are the least likely to represent a clinically significant difference
- **Probable Difference range:** Less frequent DIF scores, which are more likely to represent a clinically significant difference
- **Definite Difference range:** Least frequent DIF scores, which are the most likely to represent a clinically significant difference

Environment Cutoff Scores

A percentile-based cutoff score is used to interpret the results from the Child School Environment and Adolescent and Adult Driving Environment Forms. To develop the cutoffs, the raw-score distributions were examined for each form. A cutoff score was chosen so that raw scores at or above the cutoff indicated that the person being assessed was displaying a higher level of sensory-based problems than 90% of the research sample for that form.

Spanish and English Matched Sample

After the SPM-2 item sets and form instructions were finalized (see Chapter 5), Spanish-language versions of each form were developed. The items and instructions were translated into Spanish by a bilingual clinical psychologist with extensive translation experience and expertise in test and rating scale development. The Spanish versions were then independently back-translated by another bilingual professional into English, and the back-translations were reviewed by the authors. The authors provided

feedback to the translator, who adjusted the translations accordingly. Finally, an in-house panel of bilingual research staff reviewed the translation to achieve 100% agreement on the final word choices for each item to ensure that they would be readily understood by a variety of Spanish speakers, regardless of dialect or region. This iterative process was designed to yield Spanish translations that would be understood by the widest range of Spanish speakers.

A matched-sample study ($n = 210$) across all five age levels examined the extent to which SPM-2 Spanish and English forms showed similar results between Spanish-speaking and English-speaking respondents. Spanish-speaking participants were administered the Spanish forms, and then matched, on the variables of age, gender, education level, and ethnicity, with a randomly selected subsample from the standardization sample. The matched sampling was accomplished by using the propensity score approach (nearest neighbor method), as implemented in the R package “MatchIt” (Ho et al., 2007, 2011). Pearson correlation coefficients were calculated between the Spanish and control-group respondents’ *T*-scores on each form (except School Forms, which were not included in the study).

As expected, across all age levels, the median correlations across forms were moderate to strong for all scales (.58), including the Sensory Total scores (.71). Unlike an equivalence study in which the same participants receive each form, a matched sample is based on pairs of participants who are matched on

demographic characteristics. Different participants, even if they are similar demographically, are expected to produce different SPM-2 score profiles because there is no reason to expect their SPM-2 ratings to be at all similar, and this within-pairs variance tends to attenuate the correlations between the Spanish- and English-language scores. For example, one member of the pair could have typical sensory functioning and the other could have over-reactivity to light and sound, yielding very different score profiles. The fact that these correlations explain significant portions of the variance between the Spanish and English *T*-scores, combined with a lack of clinically meaningful effect sizes across the two groups, indicates that separate raw-to-*T*-score conversions for Spanish speakers is not warranted. Based on these results, the Spanish-speaking subsamples ($n = 210$) were included in the overall standardization sample. Taken together, these factors suggest that in clinical applications, the Spanish and English forms can be used interchangeably with the English-language norms.