**Appendix I. Summary of evidence on sensory-based subtypes in young children (<14 years of age)**

1. *Sensory subtypes identified with cross-sectional data*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Study | Sample | Age | Measure | Subtyping Method | Identified Subtypes (S) (%) | Associations with Other Variables |
| Liss et al., 2006 | 144 autistic children | Mean (SD)=102.4 (50.1) months | Sensory Profile (with other behavioral measures) | Hierarchical agglomerative cluster analysis | S1) High sensory overreactivity and seeking (12%)  S2) Few sensory symptoms (25%)  S3) High sensory underreactivity and seeking (30%)  S4) Moderate sensory overreactivity (33%) | • S1 and S3 had higher autism severity; S3 had the lowest adaptive functioning and more deficits in social communication  • S2 and S4 were less impaired in adaptive behavior and had lower autism severity; children in S2 were younger, while those in S4 were older. |
| Ben-Sasson et al., 2008 | 170 autistic toddlers | 18-33 months | Infant Toddler Sensory Profile (construct scores) | Ward's minimum variance hierarchical cluster analysis | S1) Low frequency of sensory symptoms (26%) S2) High frequency of sensory symptoms (29%) S3) Mixed: high sensory under- and overresponsivity + low sensory seeking (45%) | S2 & S3 scored higher on negative emotionality, depression, and anxiety symptoms than S1. |
| Lane et al., 2010 | 54 autistic children | 33-115 months | Short Sensory Profile  (construct scores) | Model-based cluster analysis | S1) Sensory-based inattentive seeking (44%) S2) Sensory modulation with movement sensitivity (32%) S3) Sensory modulation with taste/smell sensitivity (24%) | • Sensory processing patterns were associated with communication and adaptive behavior. • S3 had significantly greater communication impairment than S2. |
| Lane et al., 2014 | 228 autistic children | 2-10 years | Short Sensory Profile  (construct scores) | Model-based cluster analysis | S1) Sensory adaptive (38%) S2) Taste/Smell sensitive (40%) S3) Postural inattentive (10%) S4) Generalized sensory difference (12%) | • Children in S2 were younger and had lower nonverbal IQ. • No group difference was found in autism severity. |
| Tomchek et al., 2018 | 400 autistic children | 3-6 years | Short Sensory Profile  (construct scores) | Latent profile analysis | S1) Sensorimotor: elevated scores in taste/smell sensitivity, sensory seeking, and hypo-responsivity (51%) S2) Selective-complex: high levels of seeking with hypo-responsivity (15%) S3) Perceptive-adaptable (24%) S4) Vigilant-engaged: elevated sensory sensitivity and seeking behaviors (10%) | • Subtypes differed in age (S1 & S3 younger, S2 & S4 older • S1 had the lowest adaptive functioning while S4 had better developmental performance than the other subtypes. |
| Simpson et al., 2019 | 271 autistic children | 4-11 years | Short Sensory Profile  (construct scores) | Cluster analysis | S1) Uniformly elevated (67%) S2) Raised avoiding and sensitivity (33%): elevated sensitivity/avoiding and typical range on sensory seeking and registration | No difference was found in age and autism characteristics across subtypes. |
| Little et al., 2017 | 1,132 children with/without autism or other diagnoses (population-based sample) | 3-14 years | Sensory Profile (construct scores) | Latent profile analysis | S1) Balanced (79%): low sensory features S2) Interested (7%): elevated sensory seeking S3) Intense (5%): elevated sensory features across all domains S4) Mellow Until… (4%): elevated registration and avoidance S5) Vigilant (5%): elevated avoiding and sensitivity | • Children with different developmental conditions were classified across subtypes. 22% of autistic vs. 2% of typically-developing children were classified to S3. • Children in S2 were significantly younger than the other subtypes. |
| Miller et al., 2017 | 252 children with sensory processing disorder | 4-14 years | Sensory Processing 3-Dimension Inventory  (construct scores) | Ward’s agglomerative hierarchical cluster analysis | S1) High sensory overresponsivity only (46%) S2) High sensory under- and overresponsivity (29%) S3) High sensory craving and overresponsivity (25%) | • There was no significant group difference in adaptive behavior. • S3 showed more challenging behaviors than the other subtypes. |

1. *Sensory subtypes identified with longitudinal data*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Study | Sample | Age | Measure | Subtyping Method | Identified Subtypes (%) | Associations with Other Variables |
| Ausderau et al., 2014, 2016 | Autistic children (n=1,294 at T1, n=884 at T2) | 2-12 years | Sensory Experiences Questionnaire v3.0 (construct scores) | Latent profile transition analysis | 91% of the sample remained in the same latent profile across time: S1) Mild (29%): low sensory features S2) Sensitive-distressed (27%): low HYPO and SIRS + high HYPER and enhanced perception S3) Attenuated-preoccupied (17%): high HYPO and SIRS + low HYPER and enhanced perception S4) Extreme-mixed (17%): high across all sensory features | • S3 was associated with younger age and IQ on average and showed lowest levels of adaptive behavior. • S4 was associated with higher autism symptom severity and lower socioeconomic status, as well as highest levels of maladaptive behavior and parenting stress. |
| Dwyer et al., 2020 | Autistic children and typically-developing control (n=245 at T1, n=132 at T2) | 2-5 years at T1, 4-10 years at T2 | Short Sensory Profile  (total score) | Growth mixture modeling | Trajectory subgroups:  S1) Stable Mild (55%) S2) Stable Intense (43%) S3) Increasingly Intense (2%) | • 64% of the autistic children were classified to S2 and 3% to S3; 2% of typically-developing children were in S2, none in S3. • Children in S2 showed elevated anxiety; S3 tended to have higher cognitive ability. |
| The current study (Chen et al.) | 1,517 children with/without autism or other diagnoses (population-based sample) | 6-19 months at T1, 3-4 years at T2, 6-7 years at T3 | Sensory Experiences Questionnaire v2.1 and First Years Inventory v3.1 (construct scores harmonized across measures) | Latent class growth modeling | Trajectory subgroups:  S1) Adaptive – All improving (35%)  S2) Moderate – HYPO worsening (11%)  S3) Moderate SIRS – HYPER improving (21%)  S4) Mild – SIRS improving (30%)  S5) Elevated – All worsening (3%) | • Children in S2 and S5 were more likely to have an autism diagnosis/ elevated autistic traits and/or an ADHD diagnosis/concerns by school age.  • S5 showed lower adaptive functioning and more challenging behaviors at school age compared to all the other subtypes.  • Children in S2 and S5 were more likely to have parents with lower education; those in S5 were more likely to be boys; those in S3 were more likely to be non-White races. |

Note. Constructs in the Sensory Experiences Questionnaire: HYPER=sensory hyper-responsiveness; HYPO=sensory hypo-responsiveness; SIRS=sensory interests, repetitions and seeking behaviors

**Appendix II. Establishing item-response-theory (IRT) based trait scores of sensory features**

The 14 sensory items extracted from the FYIv3.1 (measured at T1) and SEQv2.1 (measured at T2 & T3) across the three sensory constructs are listed below:

|  |  |  |
| --- | --- | --- |
| **HYPER (5 items)** | **HYPO (4 items)** | **SIRS (5 items)** |
| * Refuse certain food texturesab | * Ignore when name calledab | * Look at objects in unusual ways\*ab |
| * Avoid looking at faceab | * Tune-out sounds/noises\*ab | * Fascinated with texturesa |
| * Distressed when touchedab | * Slow to notice new objectsa | * Fascinated with lightsb |
| * Distressed at loud sounds\*ab | * Slow to react to painb | * Flap arms/handsa |
| * Distressed during grooming\*a |  | * Mouth non-food objectsb |

\*Anchor items; abcommon items across FYI-A/B form; aitems specific to FYI-A form; bitems specific to FYI-B form (Note: All the 14 items are available in the SEQ); HYPER=sensory hyper-responsiveness; HYPO=sensory hypo-responsiveness; SIRS=sensory interests, repetitions and seeking behaviors

To ensure the comparability of construct scores across measures and timepoints (i.e., to meet the assumption of measurement invariance) before conducting latent growth modeling, we first tested whether longitudinal invariance held at the configural level over time, followed by metric and scalar invariance tests for each of the three sensory constructs1 with full-information maximum likelihood (FIML) estimation. Differences in fit indices between models were evaluated to determine whether invariance held at different levels. A decrease in comparative fit index (CFI) or Tucker-Lewis Index (TLI) >.01, or an increase in root-mean-square error of approximation (RMSEA) >.01 indicates measurement non-invariance.2 The purpose of invariance testing in the current study was to ensure that at least configural invariance was met before constructing IRT trait scores that adjust for differential item functioning (DIF). Longitudinal invariance testing on each of the constructs demonstrated invariance at least at the configural level (see below for the model fits), indicating that the constructs to be measured by the selected items held constant across timepoints.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | HYPER | | | HYPO | | | SIRS | | |
|  | CFI | TLI | RMSEA | CFI | TLI | RMSEA | CFI | TLI | RMSEA |
| Configural | .955 | .934 | .039 | 1.00 | 1.00 | <.001 | .982 | .973 | .021 |
| Metric | .951 | .935 | .038 | .981 | .971 | .016 | .978 | .971 | .022 |
| Scalar | .719 | .662 | .087 | .571 | .420 | .072 | .767 | .722 | .067 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | N | Mean | SD | Min | Max | Skewness | Kurtosis |
| HYPER *T1* | 1,515 | .10 | .66 | -1.39 | 2.06 | .23 | 2.54 |
| *T2* | 1,507 | -.01 | .75 | -1.35 | 3.03 | .54 | 3.45 |
| *T3* | 1,508 | -.10 | .80 | -1.34 | 3.17 | .58 | 3.48 |
| HYPO *T1* | 1,517 | -.03 | .59 | -.90 | 1.93 | .41 | 2.48 |
| *T2* | 1,508 | .03 | .70 | -.94 | 3.01 | .37 | 2.96 |
| *T3* | 1,508 | .06 | .71 | -.86 | 2.79 | .42 | 2.85 |
| SIRS *T1* | 1,502 | .11 | .81 | -1.49 | 2.81 | .25 | 2.46 |
| *T2* | 1,507 | -.02 | .88 | -1.16 | 2.99 | .46 | 2.60 |
| *T3* | 1,508 | -.12 | .85 | -1.03 | 3.22 | .67 | 2.68 |

Next, DIF was evaluated to determine which non-DIF items could be used as anchor items for scale equating, using a test-characteristic-curve equating procedure.3 A relatively conservative criterion (McFadden’s pseudo-R2 change ≥ 0.02 between nested logistic regression DIF models) was used to detect meaningful DIF.4 It has been recommended to have at least one anchor item for every four non-common items to avoid construct drift.5 As a result, we identified one to two anchor (non-DIF) items for each sensory construct with McFadden’s pseudo-R2 change ranging from .003 to .013. By recalibrating group-specific item parameter estimates (i.e., estimates specific to each timepoint) for the DIF items, trait scores of HYPER, HYPO, and SIRS that accounted for DIF across timepoints were generated. The DIF detection and trait score derivation were implemented with R package lordif.6 The descriptive statistics of the trait scores for the full sample are as below:

References:

1. Millsap, R. E. (2012). *Statistical Approaches to Measurement Invariance*. Routledge.
2. Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255. doi: 10.1207/s15328007sem0902\_5.
3. Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, *7*, 201-210. doi: 10.1177/014662168300700208
4. Paz, S. H., Spritzer, K. L., Morales, L. S., & Hays, R. D. (2013). Evaluation of the patient-reported outcomes information system (PROMIS®) Spanish-language physical functioning items. *Quality of Life Research*, *22*, 1819-1830. doi: 10.1007/s11136-012-0292-6.
5. Kolen, M. J., Brennan, R. L., & Kolen, M. J. (2004). *Test Equating, Scaling, and Linking: Methods and Practices* (pp. 177-180). New York: Springer. doi: 10.1007/978-1-4757-4310-4
6. Choi, S. W., Gibbons, L. E., & Crane, P. K. (2016). Lordif: Logistic ordinal regression differential item functioning using IRT. https://CRAN.R-project.org/package=lordif, R package version 0.3-3.

**Table S1. Demographic characteristics of respondents and non-respondents at T3**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Respondents (R)  (N=1,519) | Non-respondents (NR)†  (N=717) | R vs. NR | |
| Odds ratio | *p*-value |
| Sex (male) | 743 (48.9%) | 368 (51.3%) | .91 | .29 |
| Race† | | |  |  |
| *White* | 1,317 (86.7%) | 564 (78.7%) | 1.77 | <.001 |
| *Black* | 65 (4.3%) | 74 (10.3%) | .39 | <.001 |
| *Asian* | 16 (1.1%) | 20 (2.8%) | .37 | .003 |
| *American Indian/Hawaiian* | 11 (.7%) | 2 (.3%) | 2.61 | .21 |
| *Multi-racial* | 108 (7.1%) | 53 (7.4%) | .96 | .81 |
| *Other* | 2 (.1%) | 4 (.5%) | .24 | .10 |
| Parent Education (5% missing) | | |  |  |
| *Both parents had a college degree (or beyond)* | 898 (59.1%) | 322 (44.9%) | 1.82 | <.001 |
| *One of the parents had a college degree (or beyond)* | 327 (21.5%) | 160 (22.3%) | 1.03 | .81 |
| *None of the parents had a college degree (or beyond)* | 210 (13.8%) | 190 (26.5%) | .48 | <.001 |

†Non-respondents at T3 were excluded in the current analysis.

**Table S2. Descriptive statistics of school-age outcome measures (n=389)**

|  |  |
| --- | --- |
|  | Mean (SD) |
| SRS-2 Total (T-score) | 52.95 (11.20) |
| VABS-3 Adaptive Domain (standard score) | |
| *Communication* | 99.66 (15.85) |
| *Socialization* | 97.57 (13.59) |
| *Daily Living Skills* | 96.59 (13.43) |
| *Motor Skills* | 99.89 (12.80) |
| VABS-3 Maladaptive Domain (v-scale score) | |
| *Internalizing Behavior* | 16.22 (2.99) |
| *Externalizing Behavior* | 16.02 (3.13) |

SRS-2= Social Responsiveness Scale, 2nd Edition; VABS-3= Vineland Adaptive Behavior Scales, 3rd Edition

**Table S3. Parallel process latent class growth model parameter estimates [mean (SE)]**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Sensory constructs | Growth parameters | Class 1  (n=537) | Class 2  (n=171) | Class 3  (n=316) | Class 4  (n=449) | Class 5  (n=44) | †Overall differences χ2 (df=4) |
| HYPER | Intercept | -.22 (.04)\*\*\* | .41 (.13)\*\* | .11 (.06) | .30 (.06)\*\*\* | .58 (.18)\*\* | 568.85\*\*\* |
|  | Slope | -.21 (.02)\*\*\* | .09 (.05) | -.22 (.05)\*\*\* | -.03 (.05) | .72 (.17)\*\*\* | 373.36\*\*\* |
| HYPO | Intercept | -.24 (.04)\*\*\* | .32 (.10)\*\* | -.14 (.04)\*\* | .16 (.04)\*\*\* | .29 (.16) | 712.56\*\*\* |
|  | Slope | -.05 (.02)\* | .25 (.05)\*\*\* | .03 (.03) | .04 (.04) | .59 (.14)\*\*\* | 380.72\*\*\* |
| SIRS | Intercept | -.32 (.05)\*\*\* | .68 (.11)\*\*\* | .64 (.07)\*\*\* | -.02 (.06) | .37 (.22) | 838.82\*\*\* |
|  | Slope | -.17 (.02)\*\*\* | .05 (.07) | -.06 (.06) | -.19 (.04)\*\*\* | .47 (.19)\* | 354.66\*\*\* |

*\*p*< .05; \*\**p*< .01; \*\*\**p*< .001 (significantly different from 0); †Kruskal-Wallis rank-sum tests for comparing growth parameter estimates across latent classes. HYPER=sensory hyper-responsiveness; HYPO=sensory hypo-responsiveness; SIRS=sensory interests, repetitions and seeking behaviors

**Table S4. Equality tests of means across latent classes using the three-step BCH procedure**

|  |  |  |
| --- | --- | --- |
| Distal outcome domains | Overall difference χ2 (df=4) | Pairwise comparisons χ2  (Excluding comparisons with Class 1 as shown in Table 3) |
| SRS-2 Total T-score | 199.82\*\*\* | C2>C3 (25.93\*\*\*), C2>C4 (16.50\*\*\*), C5>C2 (33.04\*\*\*), C5>C3 (94.43\*\*\*), C5>C4 (79.99\*\*\*), C3=C4 |
| VABS-3 Subdomain Standard Score | | |
| *Communication* | 95.60\*\*\* | C2>C5 (34.87\*\*\*), C3>C2 (4.71\*), C3>C5 (52.99\*\*\*), C4>C5 (58.79\*\*\*), C2=C4, C3=C4 |
| *Socialization* | 128.27\*\*\* | C2>C5 (34.71\*\*\*), C3>C2 (13.26\*\*\*), C3>C4 (5.31\*), C3>C5 (67.63\*\*\*), C4>C2 (4.8\*), C4>C5 (72.35\*\*\*) |
| *Daily Living Skills* | 61.81\*\*\* | C2>C5 (33.14\*\*\*), C3>C5 (33.07\*\*\*), C4>C5 (29.61\*\*\*), C2=C3=C4 |
| *Motor Skills* | 53.08\*\*\* | C2>C5 (19.74\*\*\*), C3>C4 (5.02\*), C3>C5 (31.19\*\*\*), C4>C5 (20.24\*\*\*), C2=C3, C2=C4 |
| *Internalizing Behavior* | 142.14\*\*\* | C2>C3 (18.73\*\*\*), C2>C4 (4.40\*), C2>C5 (14.20\*\*\*), C4>C3 (6.95\*\*), C5>C3 (54.67\*\*\*), C5>C4 (33.33\*\*\*) |
| *Externalizing Behavior* | 104.99\*\*\* | C2>C3 (21.45\*\*\*), C2>C4 (14.71\*\*\*), C5>C3 (33.07\*\*\*), C5>C4 (25.14\*\*\*), C2=C5, C3=C4 |

*\*p*< .05; \*\**p*< .01; \*\*\**p*< .001 (two-tailed)

**Table S5. Pairwise comparisons across latent classes for parent-reported clinical outcome status (odds ratio [95% confidence interval])**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | C2 vs. C3 | C2 vs. C4 | C4 vs. C3 | C5 vs. C2 | C5 vs. C3 | C5 vs. C4 |
| AUT | 7.29\*\*\*  [2.67, 19.96] | 2.93\*\* [1.52, 5.62] | 2.49  [.89, 7.02] | 8.06\*\*\*  [3.02, 21.49] | 58.80\*\*\*  [16.55, 208.87] | 23.58\*\*\*  [8.59, 64.72] |
| ADHD | 3.66  [1.00, 13.40] | 1.60  [.68, 3.75] | 2.29  [.62, 8.52] | 4.15\*\*  [1.68, 10.21] | 15.17\*\*\*  [3.96, 58.15] | 6.62\*\*\*  [2.63, 16.66] |

*\*p*< .05; \*\**p*< .01; \*\*\**p*< .001 (two-tailed); Comparisons with Class 1 (as shown in Table 3) were excluded. AUT: presence of a parent-reported autism diagnosis via DCQ and/or with SRS-2 total T-score ≥60; ADHD: presence of a parent-reported ADHD diagnosis and/or related concerns via DCQ.