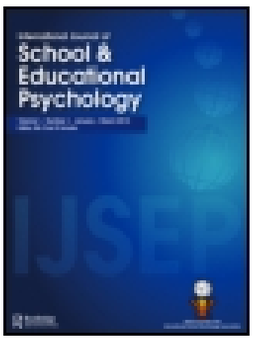
HOKKAIDO UNIVERSITY

Title	Long-term stability of the WISC-IV in children with autism spectrum disorder
Author(s)	Okada, Satoshi; Kawasaki, Yoko; Shinomiya, Mieko et al.
Citation	International Journal of School & Educational Psychology, 10(2), 290-301 https://doi.org/10.1080/21683603.2021.1930307
Issue Date	2022
Doc URL	https://hdl.handle.net/2115/82236
Rights(URL)	https://creativecommons.org/licenses/by-nc-nd/4.0/
Type	journal article
File Information	Long term stability of the WISC in children with autism spectrum disorder.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP



International Journal of School & Educational Psychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/usep20

Long-term stability of the WISC-IV in children with autism spectrum disorder

Satoshi Okada, Yoko Kawasaki, Mieko Shinomiya, Hiroshi Hoshino, Tamiko Ino, Kazuko Sakai, Kimiko Murakami, Rie Ishida, Kaoru Mizuno, Mizuho Takayanagi & Shin-Ichi Niwa

To cite this article: Satoshi Okada, Yoko Kawasaki, Mieko Shinomiya, Hiroshi Hoshino, Tamiko Ino, Kazuko Sakai, Kimiko Murakami, Rie Ishida, Kaoru Mizuno, Mizuho Takayanagi & Shin-Ichi Niwa (2021): Long-term stability of the WISC-IV in children with autism spectrum disorder, International Journal of School & Educational Psychology, DOI: 10.1080/21683603.2021.1930307

To link to this article: https://doi.org/10.1080/21683603.2021.1930307

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.

Published online: 04 Jun 2021.

Submit your article to this journal

Article views: 120

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=usep20

INTERNATIONAL JOURNAL OF SCHOOL & EDUCATIONAL PSYCHOLOGY
https://doi.org/10.1080/21683603.2021.1930307

Routledge
Taylor & Francis Group

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Long-term stability of the WISC-IV in children with autism spectrum disorder

Satoshi Okada[a], Yoko Kawasaki[b], Mieko Shinomiya[b], Hiroshi Hoshino[c], Tamiko Ino[b], Kazuko Sakai[b], Kimiko Murakami[b], Rie Ishida[b], Kaoru Mizuno[b], Mizuho Takayanagi[d], and Shin-Ichi Niwa[e]

[a]Faculty of Education, Hokkaido University, Sapporo, Japan; [b]Department of Child Psychiatry, Musashino Child Development Clinic, Tokyo, Japan; [c]Department of Neuropsychiatry, Fukushima Medical University, Fukushima, Japan; [d]Department of Child and Adolescent Psychiatry, Arisawabashi Hospital, Toyama, Japan; [e]Department of Psychiatry, Aizu Medical Center, Fukushima Medical University, Fukushima, Japan

ABSTRACT

The present study aimed to investigate the test-retest reliability of the Wechsler Intelligence Scale for Children-Fourth edition (WISC-IV) in a sample of 138 children with autism spectrum disorder (ASD) from a child psychiatric clinic in Tokyo, Japan. The stability coefficient of the Full Scale Intelligence Quotient (FSIQ), which is composed of four indices, was very high at .83, while those of the Verbal Comprehension Index (VCI), Perceptual Reasoning Index (PRI), Working Memory Index (WMI), and Processing Speed Index (PSI) individually were moderate to high, ranging from .62 to .79. Comparisons among three age groups revealed that the coefficients for children aged 5 to 7 years tended to be lower than those for children aged 11 years and older. With respect to relative strengths and weaknesses between index scores, approximately half of children did not exhibit the same trend in the second test. These results revealed that the FSIQ and index scores are stable in the long term in children with ASD aged 11 years and older, and that the PSI and discrepancies in index scores are less stable. Thus, practitioners should take into account ecological information and the test-taking behaviors of children when interpreting WISC-IV results for children with ASD.

Abbreviations: ASD: autism spectrum disorder; WISC-IV: wechsler intelligence scale for children – fourth edition; FSIQ: full scale intelligence quotient; VCI: verbal comprehension index; PRI: perceptual reasoning index; WMI: working memory index; PSI: processing speed index

KEYWORDS

Autism spectrum disorder; long-term stability; reliability; WISC-IV

Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder mainly characterized by difficulties in social communication and restricted patterns of behavior. The prevalence of ASD was initially approximately 4.5 persons in a population of 10,000 (Lotter, 1966). However, over the past four decades, prevalence has drastically increased to 86 per 10,000 in Japan (Honda et al., 2005), 168 per 10,000 in the United States (Baio et al., 2014), and 189 per 10,1000 in the UK (Rydzewska et al., 2019). This is primarily due to the increased recognition of ASD in individuals with high IQ scores and an absence of intellectual disabilities (Charman et al., 2011; Fombonne, 2005). In addition, ASD manifests itself in various ways depending on the level of intelligence or language skills (American Psychiatric Association [APA], 2013; Saulnier & Ventola, 2012; Volkmar & Pauls, 2003).

The Wechsler Intelligence Scale for Children (WISC), which has been revised several times, is the most widely utilized tool for assessing cognitive abilities in children. The fourth edition of the WISC (WISC-IV) is the version most frequently used globally, and a Japanese version of the WISC-IV published in 2010 allows the assessment of children between 5 and 16 years of age. It is often applied to evaluate the level of individual educational support needed for children with disabilities. This test is a culmination of in-depth research efforts and produces four main indices to assess the strengths and weaknesses of an individual: the Verbal Comprehension Index (VCI), Perceptual Reasoning Index (PRI), Working Memory Index (WMI), and Processing Speed Index (PSI). Moreover, its reliability and validity have been established according to the classical test theory, and an interpretive system focused on index scores based on statistical evidence has been subsequently developed. When interpreting WISC-IV scores using this system, analyses of the Full Scale Intelligence Quotient (FSIQ); index scores for VCI, PRI, WMI, and PSI; and discrepancies between these index scores are critical.

In previous studies that investigated cognitive profiles in individuals with ASD, some researchers focused on specific conditions such as islets of abilities and language impairment (Lockyer & Rutter, 1970) or theory of mind

CONTACT Satoshi Okada s.okada@edu.hokudai.ac.jp Kita-11, Nishi-7, Kita-ku, Sapporo, Japan

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

and weak central coherence (Happe, 1994; Shah & Frith, 1993). These peaks and troughs were repeatedly reported as specific WISC subtest profiles that reflect dysfunctional phenotypes (Allen et al., 1991; Bartak et al., 1975; Ehlers et al., 1997; Freeman et al., 1985; Lincoln et al., 1988; Ohta, 1987; Siegel et al., 1996). In the 1990s, however, ASD diagnostic criteria were updated (APA, 1994; World Health Organization, 1992), and the prevalence of high-functioning ASD has increased since this change. Along with this trend, the WISC subtest profile for ASD has also changed. Several studies (Mandy et al., 2015; Nader et al., 2015; Oliveras-Rentas et al., 2012; Stack et al., 2017) have reported that children with ASD exhibit relative weakness on subtests for Coding and Comprehension. Furthermore, several research groups have discussed the relationships among low scores in Coding, cognitive flexibility (e.g., attentional shifting), and motor skills (Hedvall et al., 2013; Ozonoff et al., 2000; Sattler & Dumont, 2004).

However, in their multiple single-case study, Mandy et al. (2015) reported that a subtest pattern observed at the level of the group mean did not consistently apply to individual participants. Furthermore, since the development of the WISC-IV, there has been a growing argument that the subtest scores are less stable than composite scores such as the total test IQ and index scores. Some have also argued that the analysis of subtest discrepancies to determine specific cognitive characteristics and impairments should be avoided, and that practitioners should focus on composite scores and their differences (Flanagan & Kaufman, 2009).

Research reporting WISC composite scores among ASD subtypes suggests that classical autism is characterized by strengths in nonverbal and visual-spatial composite scores, whereas Asperger's syndrome is characterized by strengths in verbal composite scores (Foley-Nicpon et al., 2012; Mayes & Calhoun, 2004; Nader et al., 2015). However, the diagnostic classifications of Asperger's syndrome and childhood autism have been integrated into one category, and the discrepancy between VCI and PRI in ASD is no longer seen as significant (Li et al., 2017; Mandy et al., 2015; Oliveras-Rentas et al., 2012; Stack et al., 2017). Nevertheless, children with ASD have shown weaknesses in the PSI (Foley-Nicpon et al., 2012; Ishikawa et al., 2013; Li et al., 2017; Mandy et al., 2015; Mayes & Calhoun, 2004; Nader et al., 2015; Oliveras-Rentas et al., 2012; Stack et al., 2017).

Cognitive profiles are clinically assessed at the individual level by analyzing discrepancies across the index scores (Wechsler, 2003). Stack et al. (2017) reported the percentage of participants with index profiles specific to ASD. They examined three different profiles: strong VCI (in comparison with PRI), strong PRI (in comparison with VCI), and weak PSI (in comparison with VCI). Their analysis revealed that 28% of children had strong VCI, 25% had strong PRI, and 41% had weak PSI, and that 53% of participants had either a strong VCI or PRI profile. Approximately half of the participants in their study had individual profiles specific to ASD. This finding demonstrated that individual profiles based on the composite score level (Stack et al., 2017) were relatively more consistent than those based on the subtest score level (Mandy et al., 2015).

Knowledge of the strengths and weaknesses of an individual may help predict his/her future difficulties, while also providing guidance on appropriate treatment. Under the circumstances, it is essential for clinical practices to confirm not only the consistency, but also the stability of the discrepancies across index scores. During its development, researchers investigated the short-term stability of the WISC. For the WISC-IV, stability coefficients of composite scores range from .73 to .89, while those for subtests range from .57 to .80. Moreover, research by Kieng et al. (2017; Non-clinical sample, N = 277, interval: 1.73 years), Watkins and Smith (2013; Special needs sample, N = 344, interval: 2.8 years), and Bartoi et al. (2015; Clinical sample, N = 51, interval: 2.3 years) has confirmed the long-term stability of the WISC-IV (Table 1). These studies have demonstrated that coefficients for the FSIQ, VCI, and PRI are particularly stable, and that children tend to have similar scores in the second test after a long period of time has passed since the first test. Despite this finding, these studies reported that stability coefficients were moderate for the WMI, PSI, and almost all subtests, ranging from approximately .40 to .70, which is relatively lower than the stability coefficients of other indices. Furthermore, they were lower than the short-term stability coefficients (Wechsler, 2003) determined when tests were readministered within 1 month after the first assessment.

Since the PSI and the Coding subtest are influenced by cognitive flexibility and motor skills, both of these scores and stability coefficients are lower in individuals with ASD than in those without (Okada et al., 2010, Table 2). Bartoi et al. (2015) compared the stability coefficients between patients with and without attentional impairments, reporting that stability for the WMI, Digit Span, and Matrix Reasoning scores was lower in the former group than in the latter (Table 2). Thus, it is safe to say that clinical features of ASD and attentional impairments influence the scores and stability of the WMI, PSI, and some subtests. For psychologists, assessing low WMI and PSI scores and the variability in these scores allows them to better identify the symptoms of ASD and attentional impairments.

Table 1. Stability coefficients of WISC-IV index scores and subtest scores.

Articles	Wechsler (2003)	Kieng et al. (2017)	Bartoi et al. (2015)	Watkins and Smith (2013)
Retest-interval	27 days	1.7 year	1.8 year	2.8 year
N	93	248	51	344
Samples	Standardized sample	Non-clinical sample	Clinical sample	Special needs sample
First/second test age		8.9/10.6	11.2/13.1	8.7/11.6
FSIQ	.89	.80	.86	.82
VCI	.88	.80	.81	.72
PRI	.80	.73	.79	.76
WMI	.73	.61	.60	.66
PSI	.79	.62	.58	.65
Similarities	.77	.67	.80	.58
Vocabulary	.80	.71	.81	.69
Comprehension	.80	.59	.49	.48
Block Design	.72	.74	.78	.70
Picture Concepts	.57	.40	.54	.46
Matrix Reasoning	.68	.56	.74	.63
Digit Span	.59	.58	.63	.60
Letter Number Sequencing	.66	42	.35	.48
Coding	.75	.56	.52	.52
Symbol Search	.70	.47	.62	.54

Table 2. Stability coefficients of WISC in children with/without ASD and ADHD.

Articles	Okada et al. (2010)		Okada et al. (2010)		Bartoi et al. (2015)	
Mean of test-retest interval (year)	1.7	1.5	1.5	1.5	1.8[c]	
N	46	91	115	22	21	29
Samples	Clinical sample without PDD[b]	Clinical sample with PDD[b]	Clinical sample without HD[b]	Clinical sample with HD[b]	Low attention problem	High attention problem
VCI	.81	.80	.81	.57	.85	.78
PRI/POI[a]	.81	.82	.79	.89	.86	.67
WMI/FDI[a]	.74	.68	.71	.65	.62	.52
PSI	.69	.49	.59	.51	.72	.47
Similarities	.56	.50	.55	.17	.81	.83
Vocabulary	.65	.66	.62	.67	.87	.78
Comprehension	.57	.61	.62	.30	.52	.36
Block Design	.72	.69	.71	.63	.80	.81
Picture Concepts	-	-	-	-	.65	.39
Matrix Reasoning	-	-	-	-	.88**	.50**
Digit Span	.72	.71	.75*	.42*	.73	.40*
Letter Number Sequencing	-	-	-	-	.35	.39
Coding	.80**	.36**	.56	.64	.65	.43
Symbol Search	.33	.47	.43	.43	.76	.48

[a]POI and FDI are WISC-3 index scores in Okada's study.
[b]PDD: Pervasive Developmental Disorders, HD: Hyperkinetic Disorder in ICD-10.
[c]Test-retest intervals for each subgroup were not provided.
* $p < .05$, ** $p < .01$, significant difference between two stability coefficients (underline) using Z-test.

There is evidence that attentional impairments are related to the stability of WMI scores and certain subtest scores in the WISC-IV. However, no studies to date have investigated the long-term stability of the WISC-IV in ASD samples. Okada et al. (2010) examined WISC-III results, but 20% of the study samples had a test-retest interval of less than 1 year. Furthermore, no previous studies have investigated the consistency in discrepancies across index scores in ASD samples. In addition, the WISC-IV is administered not only to school aged children but also to younger children who are not familiar with academic tasks. In younger children, WISC-IV scores may be affected primarily by attention and fatigue. Therefore, the stability of WISC-IV scores may differ with age.

Given the abovementioned findings, the present study aimed to confirm the cognitive profiles of WISC-IV index scores at the individual level, as well as the long-term stability of WISC-IV scores in children with ASD based on their test-retest data. We analyzed the cognitive strengths and weaknesses at both the group mean and multiple single-case levels and investigated the stabilities of discrepancies across index scores according to age. In particular, we intended to determine the following:

(1) Are there cognitive profiles specific to ASD as reported in previous studies? What percentage of children have strong VCI or strong PRI? What is the percentage of children with weak PSI?
(2) Are WISC-IV composite scores and their discrepancies stable in children with ASD over a long period of time?

(3) Are there some differences in the stability of composite scores and their discrepancies based on age?
(4) Are the profiles of strong VCI, strong PRI, and weak PSI in an individual consistent over a long period of time?

Methods

Participants

Among children who visited the Musashino Child Development Clinic in Tokyo and were diagnosed with ASD, 145 children completed the WISC-IV twice at intervals of 2 or more years. Among them, seven children with one or more raw subtest scores of 0 were excluded, given that such scores make it impossible to measure cognitive ability. In the end, we analyzed data for 138 children. All participants met the criteria for a clinical diagnosis of ASD in the International Classification of Diseases, Eleventh Revision (ICD-11). A pediatric psychiatrist with more than 40 years of clinical experience made diagnoses using the criteria for ASD outlined in the ICD-11.

Participants in this study consisted of 108 male and 30 female children with a mean age of 9.2 years (standard deviation [SD] = 2.4, minimum: 5.5, maximum: 14.9) at the time of initial WISC-IV administration, and a mean age of 11.8 years (SD = 2.5, minimum: 7.5, maximum: 16.8) at the time of retesting (Table 3). The mean test-retest interval was 2.6 years (SD = 0.7, minimum: 2.0 years, maximum: 5.0 years). Their psychiatric doctors, psychologists, and guardians confirmed that all children had not taken medication and had not had emotional and/or behavioral problems for at least 1 month before taking the WISC-IV.

Procedures and materials

The WISC-IV is composed of 10 core subtests, five ancillary subtests, and five composite scores (FSIQ, four indices). The 10 subtests are organized into four index scores: VCI, composed of Similarities, Vocabulary, and Comprehension; PRI, composed of Block Design, Picture Concepts, and Matrix Reasoning; WMI, composed of Digit Span and Letter Number Sequencing; and PSI, composed of Coding and Symbol Search. The 10 primary subtests of the Japanese version of the WISC-IV were administered in accordance with the standard method by clinical psychologists with 20 to 40 years of experience of working with children with developmental disorders and related conditions. All tests and retests were carried out individually in the laboratory of the clinic where only a desk, two chairs, and test equipment were placed. In every case, the WISC-IV was administered under the supervision of a child psychiatrist or a clinical psychologist, each having more than 40 years of experience.

Prior to administration of the test, informed consent (assent if the child was under 10 years old or with below-average cognitive ability) was obtained from children and their guardians with regard to test administration. Moreover, with respect to the utilization and release of data for research purposes, consent was obtained from guardians provided that personal information would not be disclosed. This study was approved by the Research Ethics Committee, Faculty of Education, Hokkaido University [approval number: 15–35].

Table 3. Distribution of participants by age (N = 138).

	Age											
	5	6	7	8	9	10	11	12	13	14	15	16
First test	16	11	18	24	15	10	29	9	3	3	0	0
Second test	0	0	10	13	9	21	26	8	6	37	2	6

Data analysis

Means, standard deviations, and ranges of scores in the first and second tests were calculated for FSIQ, the four index scores (VCI, PRI, WMI, PSI), and 10 subtests (Similarities, Vocabulary, Comprehension, Block Design, Picture Concepts, Matrix Reasoning, Digit Span, Letter Number Sequencing, Coding, and Symbol Search).

The discrepancies between the VCI and PRI scores were calculated for each child, and if the discrepancy score was significantly higher for the VCI (at a 5% significance level according to the Japanese version of WISC-IV scoring manual (Wechsler, 2010)), the child was classified as having a strong VCI (sVCI) profile. If the discrepancy score was significantly higher for the PRI, the child was classified as having a strong PRI (sPRI) profile. If no significant differences were noted, the child was assessed as having a VCI = PRI (V = P) profile.

Moreover, the discrepancies between the VCI and PSI scores were calculated for each child according to the method utilized by Stack et al. (2017). If the discrepancy score was significantly higher for the PSI at a 5% significance level according to the Japanese version of the WISC-IV scoring manual (Wechsler, 2010), the child was classified as having a strong PSI (sPSI) profile. If the discrepancy score was significantly lower for the PSI, the child was classified as having a weak PSI (wPSI) profile. If no significant differences were noted, the

child was assessed as having a non-significant PSI (nPSI) profile.

We also calculated the long-term stability coefficients of the WISC-IV index scores and of the discrepancies among index scores. In addition, we analyzed these long-term stability coefficients in three different age groups (5–7 years, 8–10 years, and 11 years and older). We also investigated changes in the percentage of children with relative strengths and weaknesses between VCI and PRI, as well as the percentage of those with PSI that were relatively weaker than FSIQ.

SPSS 26.0 (IBM Corp., Armonk, NY, USA) statistical software was used to perform all statistical analyses.

Results

Profiles of WISC-IV index scores in children with ASD

The descriptive statistics of the FSIQ and index scores in the first and second tests for 138 children with ASD are shown in Table 4. In the first test, the FSIQ, VCI, PRI, and WMI were within the average range, whereas index scores differed significantly (F (3, 548) = 12.65, p < .001, partial η^2 = 0.07). A Bonferroni post hoc test revealed no significant difference between the VCI and PRI. The mean difference between the VCI and PRI was 0.14, with a 95% confidence interval ranging from −5.17 to 5.20. Bonferroni post hoc testing also revealed that PSI scores were lower than VCI scores (p < .001), PRI scores (p < .001), and WMI scores (p < .01). The mean difference between the VCI and the PSI was −10.43, with a 95% confidence interval ranging from −15.59 to −5.23. The mean difference between the PRI and the PSI was −10.40, with a 95% confidence interval ranging from −15.58 to −5.22. The mean difference between the WMI and the PSI was −6.16, with a 95% confidence interval ranging from −11.34 to −0.98.

In the second test, the FSIQ, VCI, PRI, and WMI were within the average range, whereas index scores differed significantly (F (3, 548) = 21.28, p < .001, partial η^2 = 0.10). A Bonferroni post hoc test revealed no significant difference between the VCI and the PRI. The mean difference between the VCI and the PRI was 1.36 points, with a 95% confidence interval ranging from −3.83 to 6.56. Bonferroni post hoc testing revealed that PSI scores were lower than VCI scores (p < .001) and PRI scores (p < .001). The mean difference between the VCI and the PSI was −13.48 points, with a 95% confidence interval ranging from −18.68 to −8.28. The mean difference between the PRI and the PSI was −12.12 points, with a 95% confidence interval ranging from −17.31 to −6.92. The mean difference between the WMI and the PSI was −4.39 points, with a 95% confidence interval ranging from −9.59 to 0.81.

Table 4. Descriptive statistics of the WISC-IV scores (N = 138).

	First test		Second test	
	Mean	SD	Mean	SD
FSIQ	95.2	15.4	98.6	15.9
VCI	99.1	17.1	103.7	17.8
PRI	99.1	15.8	102.4	16.1
WMI	94.8	17.6	94.6	16.3
PSI	88.7	14.3	90.3	14.9
Similarities	10.4	3.5	11.3	3.5
Vocabulary	10.4	3.4	10.9	3.3
Comprehension	9.1	3.5	10.0	3.3
Block Design	10.1	3.2	10.4	3.4
Picture Concepts	9.5	2.8	10.3	3.1
Matrix Reasoning	9.9	3.5	10.4	3.3
Digit Span	9.2	3.2	9.1	3.1
Letter Number Sequencing	9.1	3.5	9.1	3.2
Coding	7.7	3.2	7.9	3.3
Symbol Search	8.4	2.9	8.7	3.1

Comparison of index scores for the first and second tests

When comparing scores on the first and second tests, the mean FSIQ and VCI scores increased by 3.4 and 4.6 points in the second test, and the effect sizes were medium (respectively, r = 0.36; r = 0.35). The mean PRI score increased by 3.3 points, and the effect size was small (r = 0.25). The WMI and PSI scores changed little between the first and second tests.

WISC-IV subtest score profile in children with ASD

The descriptive statistics of the WISC-IV subtest scores in the first and second tests are shown in Table 4. In both the first and second tests, children had the lowest score in Coding, a subtest of the PSI. As for subtests of the VCI, Comprehension scores were low, and there was a 1.3-point difference between Similarities and Comprehension scores in both the first and second tests.

Long-term stability of index scores and subtest scores

The stability coefficients of the WISC-IV FSIQ and index scores are shown in Table 5. Although stability coefficients for the four indices were moderate to high, ranging from .620 to .794, the stability coefficient of the FSIQ composed of the four indices was .83, indicating very high stability.

Table 5. Stability coefficients of the WISC-IV scores (N = 138).

	Total	Aged 5–7	Aged 8–10	Aged 11–14	Significant difference
FSIQ	.834	.752	.796	.899	
VCI	.743	.762	.779	.711	
PRI	.688	.592	.499	.834	5–7 < 11–14*, 8–10 < 11–14**
WMI	.794	.624	.831	.883	5–7 < 8–10*, 5–7 < 11–14**
PSI	.620	.458	.570	.772	5–7 < 11–14*
Similarities	.629	.535	.687	.695	
Vocabulary	.654	.660	.637	.641	
Comprehension	.626	.634	.622	.590	
Block Design	.737	.701	.723	.774	
Picture Concepts	.419	.303	.267	.590	
Matrix Reasoning	.626	.649	.381	.808	8–10 < 11–14*
Digit Span	.709	.559	.785	.784	5–7 < 8–10*
Letter Number Sequencing	.669	.576	.663	.778	
Coding	.577	.430	.523	.746	5–7 < 11–14*
Symbol Search	.475	.322	.415	.639	
VCI-PRI	.560	.596	.590	.596	
VCI-WMI	.601	.347	.703	.660	5–7 < 8–10*
VCI-PSI	.567	.662	.611	.494	
PRI-WMI	.582	.362	.640	.638	
PRI-PSI	.526	.496	.451	.622	
WMI-PSI	.518	.373	.528	.615	

* $p < .05$, ** $p < .01$.
significant differences in correlations between two age groups using Z-test (two tailed).
5–7 < 8–10: the group aged 5–7 showed significantly lower stability than the group aged 8–10.
5–7 < 11–14: the group aged 5–7 showed significantly lower stability than the group aged 11–14.
8–10 < 11–14: the group aged 8–10 showed significantly lower than the group aged 11–14.

Comparison of long-term stability according to age group

When comparing the three age groups (5–7 years, 8–10 years, 11 years and older), scores were less stable overall in children aged 5 to 7 years, whereas children ≥11 years old had index scores consistently above .700. A comparison of stability coefficients revealed that, in terms of PRI scores, significant differences were observed between children aged 5–7 and those aged 11–14, and between those aged 8–10 and those aged 11–14 ($z = -2.32$, $p = .024$, $z = -3.03$, $p = .002$). In terms of WMI scores, significant differences were noted between children aged 5–7 and those aged 8–10, and between those aged 5–7 and those aged 11–14 ($z = -2.16$, $p = .031$, $z = -2.96$, $p = .003$). In terms of PSI scores, a significant difference was noted between those aged 5–7 and those aged 11–14 ($z = -2.46$, $p = .014$), and a marginally significant difference was noted between those aged 8–10 and those aged 11–14 ($z = -1.72$, $p = .085$). While VCI results were consistent across different age groups, consistent results for PRI and PSI scores were only obtained if children were aged 11 years or older. Similarly, consistent results for WMI scores were only obtained if children were aged 8 years or older.

The stability coefficients for Picture Concepts, Coding, and Symbol Search scores were below .600, consistent with the findings of previous studies (Bartoi et al., 2015; Kieng et al., 2017; Watkins & Smith, 2013). The stability coefficient for the Letter Number Sequence was .669, which was higher than that reported in a previous study (Bartoi et al., 2015; Kieng et al., 2017; Watkins & Smith, 2013).

When comparing results across age groups, the stability coefficient was under .400 for Picture Concepts and Symbol Search scores in children aged 5–7, and for Picture Concepts and Matrix Reasoning scores in children aged 8–10. Overall, stability coefficients for subtests were higher in children aged 11–14, ranging from .590 to .808.

A pair-wise comparison of the difference in stability coefficients between age groups was carried out using the z-test. This analysis revealed a significant difference in stability coefficients for Matrix Reasoning scores between children aged 8–10 and those aged 11–14 ($z = -3.26$, $p = .001$). Furthermore, significant differences in Digit Span scores were observed between children aged 5–7 and those aged 11–14, and between those aged 8–10 and those aged 11–14 ($z = -2.15$, $p = .044$, $z = -1.96$, $p = .050$). In terms of Coding scores, a significant difference was observed between children aged 5–7 and those aged 11–14 ($z = -2.23$, $p = .025$).

Only the stability coefficient for Matrix Reasoning scores decreased in those aged 8–10 years. In contrast, stability increased with age for Picture Concepts, Digit Span, Letter Number Sequence, Coding, Symbol Search, and other items.

Long-term stability of discrepancies between index scores

The stability coefficients for discrepancies between index scores are shown in Table 5. Among all 138 participants, the correlation coefficients were moderate, ranging from .51 to .60. In children aged 5–7 years, correlation coefficients for discrepancies between the WMI and other indices ranged from .34 to .37 and were statistically significant at a significance level of 5%, indicating little stability.

Changes in ASD-specific profiles from the first to second test

In the first test, 35 children had an sVCI profile (25%), 68 children had a V = P profile (49%), and 35 children

Table 6. Changes in the proportion of those who had relative strengths and weaknesses between index scores (N = 138).

	Second test					Second test			
	sVCI	V = P	sPRI	Total		sPSI	nPSI	wPSI	Total
First test						First test			
sVCI	18	16	1	35	sPSI	3	6	0	9
	51%	46%	3%	(25%)		33%	67%	0%	(7%)
V = P	16	40	12	68	nPSI	8	44	21	73
	24%	59%	17%	(49%)		11	60	29	(53%)
sPRI	2	19	14	35	wPSI	0	14	42	56
	6%	54%	40%	(25%)		0%	25%	75%	(41%)
Total	36	75	27	138	Total	11	64	63	138
	(26%)	(54%)	(20%)			(8%)	(46%)	(46%)	

The percentages were computed by dividing the count for a cell by the total sample size for that row.
The percentages of total, shown in parentheses, were computed by dividing the count for a cell by the total sample (N = 138).
The VCI-PRI discrepancy analysis was performed and the participants showing a difference at a 5% significance level were classified as having strong VCI (sVCI) or strong PRI (sPRI), and those showing no significant difference were classified as having a V = P profile.
The VCI-PSI discrepancy analysis was performed and the participants showing a difference at a 5% significance level were classified as having strong PSI (sPSI) or weak PSI (wPSI); those showing no significant difference were classified as having a non-significant PSI (nPSI) profile.

had an sPRI profile (25%). In the second test, 36 (26%), 75 (54%), and 27 (20%) children had sVCI, V = P, and sPRI profiles, respectively, and the proportions of children with each profile were similar to those observed for the first test (Table 6). The number of children classified under the same profile in both the first and second tests was 18 (51%) for sVCI, 40 (58%) for V = P, and 14 (40%) for sPRI. The relative strengths and weaknesses between the VCI and PRI scores were consistent in nearly half the participants. However, one child (3%) changed from an sVCI to sPRI profile, and two children (6%) changed from an sPRI to sVCI profile.

There were nine (7%), 73 (53%), and 56 (41%) children in the sPSI, nPSI, and wPSI categories, respectively, after the first test. After the second test, there were 11 (8%), 64 (46%), and 63 (46%) children in the sPSI, nPSI, and wPSI categories, respectively. In both the first and second tests, approximately 40% of children were in the weak PSI category. Of the nine children in the sPSI category after the first test, three children (33%) remained in the sPSI category. Of the 73 children in the nPSI category after the first test, 44 (60%) remained in the nPSI category after the second test. Of the 56 children in the wPSI category after the first test, 42 (75%) remained in the wPSI category after the second test.

Discussion

The purposes of this study were to confirm the cognitive profiles of index scores at an individual level, and to examine the long-term stability of WISC-IV scores in children with ASD based on their test-retest data.

WISC-IV profiles in children with ASD

In this study, we collected the WISC-IV test data of children with ASD assessed at intervals of 2 or more years. An analysis of variance (ANOVA) demonstrated that, in both the first and second tests, the participants had a profile characterized by relatively low PSI and high VCI and PRI scores at the group mean level. These results are consistent with the findings of past WISC profile studies for ASD (Ishikawa et al., 2013; Nader et al., 2015; Oliveras-Rentas et al., 2012; Stack et al., 2017; Wechsler, 2003). Such studies indicate that the profile specific to ASD may be common across cultures.

On an individual level, approximately half of the participants had a strong VCI (25% on the first test and 26% on the second test) or a strong PRI profile (25% on the first test and 20% on the second test), as well as a weak PSI profile (41% on the first test and 46% on the second test). These results are in accordance with those of previous research (Stack et al., 2017) and indicate that approximately half of children with ASD have index score profiles specific to ASD. The results also suggest that discrepancies between VCI and PRI, and between PSI and other indices characterizes the cognitive feature of ASD, rather than a single index score such as PSI.

Conversely, half of our participants exhibited a profile in which there was no difference between VCI and PRI. In addition, approximately half of participants did not have a weak PSI. These results indicate that approximately half of the children with ASD in our study had a flat profile of index scores. There are two possible explanations for this finding: Either the intra-individual differences among cognitive abilities in many children with ASD are latently small, or they are not likely to be well represented by WISC-IV composite scores. Irrespective of the explanation, the finding implies that psychologists must make careful clinical decisions without being restricted to the cognitive profiles determined at a group mean level.

Long-term stability of composite scores

The stability coefficients for two tests separated by a 2-year interval, which is long enough to reduce the practice effect, were high for the FSIQ, VCI, and WMI. Thus, the results of the present study verify the long-term stability of the FSIQ, VCI, and WMI in children with ASD. In ASD, intellectual and verbal functions during childhood are known to be strong predictors of outcomes in adulthood (Volkmar & Pauls, 2003). FSIQ is composed of primary index scores (primary subtests) and has been reported to be most stable during a short

interval (Wechsler, 2003, 2010, 2014). The fact that the FSIQ is stable in children with ASD after long intervals has important clinical implications.

The stability coefficients of the PRI and PSI were slightly lower than those of other indices (under .70). The low stability of the PRI (r = .688) seems related to the low stability of the Picture Concepts subtest (r = .419). Indeed, previous studies (Bartoi et al., 2015; Kieng et al., 2017; Watkins & Smith, 2013) have reported that stability coefficients are lower for this subtest than for other PRI subtests.

The stability of the PSI (r = .620) was similar to that observed in the clinical groups in previous studies (Table 1). The instability of the PSI may not be limited to ASD. The PSI is influenced by attention, motivation, and fatigue (Sattler & Dumont, 2004; Wechsler, 2003); therefore, it may be less stable than other indices. It is therefore essential that, when interpreting PSI scores, practitioners observe the child's psychological and physical response during the subtest sessions to analyze influencing factors.

Long-term stability of composite scores according to age

In the analysis of long-term stability across age groups, the correlation coefficients of the PRI, WMI, and PSI were significantly lower in younger children (5–7 years) than in other age groups. PSI, WMI, and PRI values were under .700 in children aged 5–7 years, while PRI and PSI values were under .600 in children aged 8–10 years. Moderate correlations were observed in all cases. Since young children are unfamiliar with academic tasks, they may have difficulty in dealing well with numbers, symbols and a pencil. As a result, the scores obtained in the first test depend on the academic experiences in preschool stage, and hence, these scores may vary between the first and second tests. The fourth edition of Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV) for very young children implements new game-like subtests, which offer animal and bug pictures and use of an ink dauber to indicate responses, thereby minimizing the demands of pencil-based tasks (Wechsler, 2014). The latest revision of the WPPSI may be meaningful for young children. When psychologists apply the WISC to young children, they should focus on whether a child is familiar with academic tasks and pencil-based activities.

Long-term stability of subtest scores according to age

Notably, in children who completed the initial WISC assessment at 8 to 10 years of age and the second WISC at 10 to 12 years of age or older, the stability coefficients for the PRI (r = .499), Matrix Reasoning (r = .381), and Picture Concepts (r = .267) subtests were small to moderate. There may be two reasons for the stability gaps in these subtests when analyzed by age. First, the stage of cognitive processing may differ, depending on the individual's age and ASD characteristics. Matrix Reasoning and Picture Concepts subtests include fluid reasoning components, and the Block Design test – which exhibited a relatively high stability coefficient – is considered to represent visual-spatial processing ability (Wechsler, 2003). At around age 11, a child enters a stage of formal operations and is able to perform abstract and logical reasoning (Harris & Westermann, 2014). During this period, children with ASD are split into those who do develop this capacity for operational thinking without relying solely on visual processing and those who cannot develop thinking skills beyond visual or concrete thinking. This might have led to large gaps between the first and second test results. Since this remains speculative, future studies must elaborate on this point by assessing task-solving strategies in individuals completing the WISC-IV and by analyzing correct and incorrect answers in each subtest.

Second, stability gaps may be related to whether each subtest requires attentional and inhibitory control. Performance on Matrix Reasoning and Picture Concepts subtests may be influenced by inattention and impulsivity, given that children are required to carefully consider response options before making their choice (Bartoi et al., 2015; Sattler & Dumont, 2004). In a previous study (Bartoi et al., 2015), the stability of Matrix Reasoning scores was lower in children with more severe attentional impairments than in those with less severe attentional impairments. In the present study, the stability of Matrix Reasoning scores was moderate in children aged 5–7 years (r = .649) but low in children aged 8–10 years (r = .381). Moreover, the stability of Picture Concepts results was small in both the 5–7 and 8–10 age groups (r = .303 and .267, respectively). Further research is required to confirm the relationships among age, attentional control, and subtest stability.

Conversely, the stability coefficient of the Block Design test exceeded .700. This was the only subtest with a large stability coefficient in the 5–7 age group. In this subtest, children may handle blocks without careful thinking, thereby allowing them to solve tasks based on trial and error without inhibiting impulsivity. In addition, many preschoolers usually like to play with blocks. As the Block Design task is familiar to them, they may be able to perform such tasks in a stable manner.

Long-term stability of discrepancies between index scores and consistency of index score profiles

Among our participants with ASD, we observed moderate long-term stability coefficients ranging from .518 to .601 for discrepancies between index scores. In addition, the numbers of children classified under the same profile after both the first and second tests were 18 (51%) for sVCI, 14 (40%) for sPRI, and 42 (75%) for wPSI.

These findings indicate that discrepancies between index scores tend to be more variable than each index score itself. Thus, when identifying whether a child has a strength in visual ability or verbal ability, and a weakness in processing speed, it is not sufficient for psychologists to rely solely on the statistical evaluation of the discrepancies between index scores.

Moreover, in children aged 5–7 years, the discrepancies between WMI and other indices were all low (stability coefficients ranging from .347 to .375). The WMI is impacted by attention, concentration, and executive functions (Wechsler, 2003). Young children such as preschoolers may be more inattentive than school-aged children. Therefore, they are likely to respond without thinking during tasks, or to be distracted by unnecessary stimuli. This may result in unstable discrepancies between WMI and other indices; it also indicates the importance of psychological assessment focusing on the child's development of attention and executive skills. In a different light, it indicates that analysis of influencing factors may lead to valuable evaluation of a child's executive function development.

Implications for practice

Our findings indicated that the stability coefficients of discrepancies between index scores and the consistency of index score profiles specific to ASD were not sufficiently high for predicting children's abilities over time. To compensate for this limitation, we must better understand what strategies are used when solving tasks and what factors affect children's scores by carefully observing their behavior during each subtest. This qualitative behavioral analysis is termed a process approach (McCloskey 2009) or an Assessment of Test Behaviors (Oakland et al., 2005) and is used to assess neuropsychological symptoms such as executive dysfunctions. When interpreting the PRI score, PSI score, and discrepancies related to the WMI score in children with ASD, it is necessary to link these results with the behavioral assessments in a comprehensive manner to specifically address each child's individual characteristics. Moreover, when interpreting test results in a way that is not described in the test manual or in a manner that has not been fully validated, it is necessary to link these results to a child's daily functioning and difficulties (e.g., previous school records, histories, and ratings) to confirm the ecological validity of interpretation (Hale et al., 2011).

Limitations

There are some limitations to this study. First, we could not collect data pertaining to comorbid symptoms such as inattention and motor deficits. Furthermore, we did not assess influencing factors using external variables related to executive functions, such as cognitive flexibility. It is presumed that these deficits may influence the index and discrepancy scores. Future studies should investigate the relationships between scores and influencing factors using additional scales such as the Behavior Rating Inventory of Executive Function (BRIEF; Gioia et al., 2002), which includes flexibility and inhibition subscales; the Attention Deficit Hyperactive Disorder Rating Scale (ADHD-RS; DuPaul et al., 1998), which includes an inattention subscale; and the Developmental Coordination Disorder Questionnaire (DCDQ; Wilson et al., 2009). Furthermore, the findings of this study must be confirmed in a comparative study involving other clinical groups of children with learning disabilities or attention-deficit/hyperactive disorders.

Second, the stabilities of the WMI and PSI scores were low in children aged 5–7 years, as was the stability of discrepancies in scores. Thus, application of the WISC-IV to young children should be reconsidered. The Wechsler scales are available in three different formats according to the age of test takers: the WPPSI, WISC, and Wechsler Adult Intelligence Scale (WAIS). The Japanese versions of the WISC-IV and WPPSI-III overlap at ages 5 to 7 years and 3 months. The WPSSI-III is recommended for younger children with low intellectual levels (Sattler & Dumont, 2004). This recommendation is based on the FSIQ ceiling, floor, and breadth of coverage in the WISC-IV and WPPSI-III. Our results suggest that use of the WPPSI should be considered for young children in the overlapping age bracket. To this end, it is important to confirm the long-term stability of the WPPSI, as well as changes in scores and the long-term stability of outcomes when shifting from using the WPPSI to the WISC.

Third, although retests were administered after 2 years, which is long enough to reduce the practice effect, the VCI and FSIQ scores increased significantly with medium effect sizes in the second test. The FSIQ also increased in the second test because it is composed of the VCI, PRI, WMI, and PSI. Although the standard deviations of the VCI were large (SD = 17.1 for the first

test, SD = 17.8 for the second test) and the increase in scores must be interpreted with caution, the second VCI score may be influenced by social and educational experiences. The VCI not only measures verbal ability but also crystalized ability (Flanagan & Kaufman, 2009). Almost all participants had participated in some type of educational treatment (e.g., special educational support in the classroom, professional treatment in a resource room, etc.). We could not collect data regarding the quality and quantity of educational treatment, the family's nursing skills, or socioeconomic status. Further studies are required to understand the factors influencing WISC scores and their stability.

Acknowledgments

We would like to express our deep gratitude to the children who took these tests and to their parents for willingly consenting to provide the test results for this study.

Author contributions

SO and YK designed the study. YK, MS, TI, KS, KM, RI, and KM contributed with data collection and test administration. SO wrote the initial draft of the manuscript. SO, YK, HH, MT, and SN contributed to the discussion regarding the interpretation of the collected papers and revised the manuscript. All authors approved the final version of the manuscript.

Data availability

Due to the nature of this research, participants of this study did not agree to share their data publicly; hence, supporting data is not available.

Disclosure statement

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

Funding

This work was supported in part by JSPS KAKENHI Grant Number JP18H01087.

Notes on contributors

Dr. Satoshi Okada is a clinical psychologist and associate professor at the Hokkaido University. His research interests are psychological assessments and interventions for children with social and emotional difficulties. He has published books and articles related to cognitive and behavioral assessment and social skills training for developmental disabilities.

Dr. Yoko Kawasaki is a child psychiatrist and the Director of Musashino Child Development Clinic. She and her colleagues have been engaged in the clinical treatment and community support service for people with Autism Spectrum Disorder more than 40 years.

Mieko Shinomiya is a medical technologist at Musashino Child Development Clinic. She and her colleagues have been engaged in the clinical treatment and community support service for people with Autism Spectrum Disorder more than 40 years.

Hiroshi Hoshino is a researcher at the Department of Neuropsychiatry, School of Medicine, Fukushima Medical University, Fukushima, Japan. His research focuses on cognitive neurophysiology and general psychiatry.

Tamiko Ino is a clinical psychologist at Musashino Child Develoment Clinic. She and her colleagues have been engaged in the clinical treatment and community support service for people with Autism Spectrum Disorder more than 40 years.

Kazuko Sakai is a clinical psychologist at Musashino Child Develoment Clinic. She and her colleagues have been engaged in the clinical treatment and community support service for people with Autism Spectrum Disorder more than 40 years.

Kimiko Murakami is a clinical psychologist at Musashino Child Develoment Clinic. She has been engaged with her colleagues in the clinical treatment and community support service for people with Autism Spectrum Disorder more than 40 years.

Rie Ishida is a clinical psychologist at Musashino Child Develoment Clinic. She has been engaged with her colleagues in the clinical treatment and community support service for people with Autism Spectrum Disorder more than 20 years.

Kaoru Mizuno is a clinical psychologist at Musashino Child Develoment Clinic. She and her colleagues have been engaged in the clinical treatment and community support service for people with Autism Spectrum Disorder more than 40 years.

Mizuho Takayanagi is a child psychiatrist at Arisawabashi Hospital. She has been engaged with her colleagues in the clinical treatment and community support service for people with Autism Spectrum Disorder more than 20 years.

Shin-Ichi Niwa is a psychiatrist at Aizu Medical Center, Fukushima Medical University. His research interests mainly exist in cognition and cognitive dysfunction of psychiatric diseases.

References

Allen, M. H., Lincoln, A. J., & Kaufman, A. S. (1991). Sequential and simultaneous processing abilities of high-functioning autistic and language-impaired children. *Journal of Autism and Developmental Disorders*, *21*(4), 483–502. https://doi.org/10.1007/BF02206872

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Rosenberg, C. R., White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L.-C., Harrington, R.,

Lopez, M., Fitzgerald, R. T., Hewitt, A., & Dowling, N. F. (2014). Prevalence of autism spectrum disorder among children aged 8 years. *MMWR Surveill Summaries*, *67*(6), 1–23. https://doi.org/10.15585/mmwr.ss6706a1

Bartak, L., Rutter, M., & Cox, A. (1975). A comparative study of infantile autism and specific development receptive language disorder. I. The children. *British Journal of Psychiatry*, *126*(2), 127–145. https://doi.org/10.1192/bjp.126.2.127

Bartoi, M. G., er, J. B., Hetterscheidt, L., January, A. M., Kuentzel, J. G., & Barnett, D. (2015). Attention problems and stability of WISC-IV scores among clinically referred children. *Applied Psychology: Child*, *4*(3), 133–140. https:doi.org/10.1080/21622965.2013811075

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: Data from the Special Needs and Autism Project (SNAP). *Psychological Medicine*, *41*(3), 619–627. https://doi.org/10.1017/S0033291710000991

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD rating scale-IV: Checklists, norms, and clinical interpretation*. The Guilford Press.

Ehlers, S., Nyden, A., Gillberg, C., Sanberg, A. D., Dahlgren, S.-O., Hjelmquist, E., & Oden, A. (1997). Asperger syndrome, autism and attention disorders: A comparative study of the cognitive profiles of 120 children. *Journal of Child Psychology and Psychiatry*, *38*(2), 207–217. https://doi.org/10.1111/j.1469-7610.1997.tb01855.x

Flanagan, D. P., & Kaufman, A. S. (2009). *Essentials of WISC-IV assessment* (2nd ed.). Wiley.

Foley-Nicpon, M., Assouline, S. G., & Stinson, D. D. (2012). Cognitive and academic distinctions between gifted students with autism and Asperger syndrome. *Gifted Child Quartery, 65*, 77–89. https://doi.org/10.1177/0016986211433199

Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *Journal of Clinical Psychiatry*, *66*(10), 3–8.

Freeman, B. J., Lucas, J. C., Forness, S. R., & Ritvo, E. R. (1985). Cognitive processing of high-functioning autistic children: Comparing the K–ABC and the WISC–R. *Journal of Psychoeducational Assessment*, *3*(4), 357–362. https://doi.org/10.1177/073428298500300408

Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & Espy, K. A. (2002). Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. *Child Neuropsychology*, *8*(4), 249–257. https://doi.org/10.1076/chin.8.4.249.13513

Hale, J. B., Wycoff, K. L., & Fiorello, C. A. (2011). RTI and cognitive hypothesis testing for identification and intervention of specific learning disabilities. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 198). John Wiley & Sons.

Happe, F. G. (1994). Wechsler IQ profile and theory of mind in autism: A research note. *Journal of Child Psychology and Psychiatry*, *35*(8), 1461–1471. https://doi.org/10.1111/j.1469-7610.1994.tb01287.x

Harris, M., & Westermann, G. (2014). *A student's guide to developmental psychology*. Psychological Press.

Hedvall, Å., Fernell, E., Holm, A., Johnels, J. Å., Gillberg, C., & Billstedt, E. (2013). Autism, processing speed, and adaptive functioning in preschool children. *Scientific World Journal*, *2013*, ID158263. https://doi.org/10.1155/2013/158263

Honda, H., Shimizu, Y., & Rutter, M. (2005). No effect of MMR withdrawal on the incidence of autism: A total population study. *Journal of Child Psychology and Psychiatry*, *46*(6), 572–579. https://doi.org/10.1111/j.1469-7610.2005.01425.x

Ishikawa, N., Kawanura, Y., & Ogasawara, A. (2013). WISC-IV profiles of 42 Japanese children with high functioning pervasive developmental disorders. *Psychiatria Et Neurologia Paediatrica Japonica*, *53*(1), 33–39.

Kieng, S., Rossier, J., Faves, N., & Lecerf, T. (2017). Long-term stability of the WISC-IV: Standard and CHC index scores. *European Review of Applied Psychology*, *67*(1), 51–60. https://doi.org/10.1016/j.erap.2016.10.001

Li, G., Du, Y., Luan, F., Li, M., & Ousley, O. (2017) IQ profiles and clinical symptoms of Chinese school-aged boys with autism spectrum disorder. *The European Journal of Psychiatry*,*31*(2), 59–65. https://doi.org/10.1016/j.ejpsy.2017.04.001

Lincoln, A. J., Courehesne, E., Kilman, B. A., Elmasian, R., & Allen, M. (1988). A study of intellectual abilities in high-functioning people with autism. *Journal of Autism and Developmental Disorders*, *18*(4), 505–524. https://doi.org/10.1007/BF02211870

Lockyer, L., & Rutter, M. (1970). A five- to fifteen-year follow-up study of infantile psychosis. IV. Patterns of cognitive ability. *British Journal of Social and Clinical Psychology*, *9*(2), 152–163. https://doi.org/10.1111/j.2044-8260.1970.tb00654.x

Lotter, V. (1966). Epidemiology of autistic conditions in young children. *Social Psychiatry*, *1*(3), 124–135. https://doi.org/10.1007/BF00584048

Mandy, W., Murin, M., & Skuse, D. (2015). The cognitive profile in autism spectrum disorders. In M. Leboyer & P. Chaste (Eds.), *Autism spectrum disorders: Phenotypes, mechanisms and treatments* (pp. 34–45). Karger.

Mayes, S. D., & Calhoun, S. L. (2004). Similarities and differences in Wechsler Intelligence Scale for Children, Third Edition (WISC-III) profiles: Support for subtest analysis in clinical referrals. *Clinical Neuropsychology*,*18*, 559–572. https://doi.org/10.1080/13854040490888530

McCloskey, G., & Maerlender, A. (2009). WISC-IV integrated. In D.P. Flanagan, A.S. Kaufman (Eds.), *Essentials of WISC-IV Assessment. Second Edition* (pp. 310–467). Wiley

Nader, A. M., Jelenic, P., & Soulieres, I. (2015). Discrepancy between WISC-III and WISC-IV cognitive profile in autism spectrum: What does it reveal about autistic cognition? *PLoS One*, *10*(12), e0144645. https://doi.org/10.1371/journal.pone.0144645

Oakland, T., Glutting, J., & Watkins, M. W. (2005). Assessment of test behaviors with the WISC-IV. In A. Prifitira, D. Saklofske, & L. Weiss (Eds.), *WISC-IV clinical use and interpretation. scientist-practitioner perspective* (pp. 435-463). Academic Press.

Ohta, M. (1987). Cognitive disorders of infantile autism: A study employing the WISC, spatial relationship conceptualization, and gesture imitations. *Journal of Autism and Developmental Disorders*, *17*(1), 45–62. https://doi.org/10.1007/BF01487259

Okada, S., Mizuno, K., Yokota, K., & Kawasaki, Y. (2010). Effects of test-retest interval on the Japanese version of WISC-III among students with developmental disabilities: Practice effect and stability. *Japanese Journal of Child and Adolescent Psychiatry*, *51*(1), 31–43.

Oliveras-Rentas, R. E., Kenworthy, L., Roberson, R. B., 3rd, Martin, A., & Wallace, G. L. (2012). WISC-IV profile in high-functioning autism spectrum disorders: Impaired processing speed is associated with increased autism communication symptoms and decreased adaptive communication abilities. *Journal of Autism and Developmental Disorders*, *42*(5), 655–664. https://doi.org/10.1007/s10803-011-1289-7

Ozonoff, S., South, M., & Miller, J. N. (2000). DSM-IV-defined Asperger syndrome: Cognitive, behavioral and early history differentiation from high-functioning autism. *Autism*, *4*(1), 29–46. https://doi.org/10.1177/1362361300041003

Rydzewska, E., Hughes-mccormack, L. A., Gillberg, C., Henderson, A., MacIntyre, C., Rintoul, J., & Cooper, S. (2019). Age at identification, prevalence and general health of children with autism: Observational study of a whole country population. *BMJ Open*, *9*(7), e025904. https://doi.org/10.1136/bmjopen-2018-025904 .

Sattler, J. M., & Dumont, R. (2004). *Assessment of children WISC-IV and WPPSI-III supplement*. Sattler Publisher.

Saulnier, C. A., & Ventola, P. E. (2012). *Essentials of autism spectrum disorders evaluation and assessment*. John Wiley & Sons.

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry*, *34*(8), 1351–1364. https://doi.org/10.1111/j.1469-7610.1993.tb02095.x

Siegel, D. J., Minshew, N. J., & Goldstein, G. (1996). Wechsler IQ profiles in diagnosis of high-functioning autism. *Journal of Autism and Developmental Disorders*, *26*(4), 389–406. https://doi.org/10.1007/BF02172825

Stack, K., Murphy, R., Prendeville, P., & O'Halloran, M. (2017). WISC-IV UK profiles of children with autism spectrum disorder in a specialist autism service. *Educational and Child Psychology*, *34*(2), 40–53.

Volkmar, F. R., & Pauls, D. (2003). Autism. *Lancet*, *362*(9390), 1133–1141. https://doi.org/10.1016/S0140-6736(03)14471-6

Watkins, M. W., & Smith, L. G. (2013). Long-term stability of the Wechsler Intelligence Scale for children-fourth edition. *Psychological Assessment*, *25*(2), 477–483. https://doi.org/10.1037/a0031653

Wechsler, D. (2003). *WISC-IV technical and interpretive manual*. Psychological Corporation.

Wechsler, D. (2010). *Japanese version of WISC-IV technical and interpretive manual*. Nihonbunkakagakusya.

Wechsler, D. (2014). *Wechsler primary and preschool scale of intelligence* (4th ed.). Pearson.

Wilson, B. N., Crawford, S. G., Green, D., Roberts, G., Aylott, A., & Kaplan, B. J. (2009). Psychometric properties of the revised developmental coordination disorder questionnaire. *Physical & Occupational Therapy in Pediatrics*, *29*(2), 184–204. https://doi.org/10.1080/01942630902784761

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. WHO.

World Health Organization. (2018). *ICD-11, the 11th revision of the international classification of diseases*. WHO. https://www.who.int/classifications/icd/en/