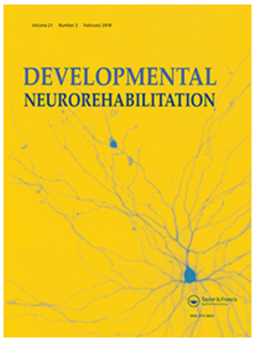
Accuracy of WISC-III and WAIS-IV short forms in patients with neurological disorders



Taylor & Francis
Taylor & Francis Group

Developmental Neurorehabilitation

ISSN: 1751-8423 (Print) 1751-8431 (Online) Journal homepage: https://www.tandfonline.com/loi/ipdr20

Accuracy of WISC-III and WAIS-IV short forms in patients with neurological disorders

Jans S. van Ool, Petra P.M. Hurks, Francesca M. Snoeijen-Schouwenaars, In Y. Tan, Helenius J. Schelhaas, Sylvia Klinkenberg, Albert P. Aldenkamp & Jos G.M. Hendriksen

To cite this article: Jans S. van Ool, Petra P.M. Hurks, Francesca M. Snoeijen-Schouwenaars, In Y. Tan, Helenius J. Schelhaas, Sylvia Klinkenberg, Albert P. Aldenkamp & Jos G.M. Hendriksen (2018) Accuracy of WISC-III and WAIS-IV short forms in patients with neurological disorders, Developmental Neurorehabilitation, 21:2, 101-107, DOI: 10.1080/17518423.2016.1277799

To link to this article: https://doi.org/10.1080/17518423.2016.1277799

Published online: 02 Feb 2017.

DEVELOPMENTAL NEUROREHABILITATION
2018, VOL. 21, NO. 2, 101–107
http://dx.doi.org/10.1080/17518423.2016.1277799

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Accuracy of WISC-III and WAIS-IV short forms in patients with neurological disorders

Jans S. van Ool[a], Petra P.M. Hurks[b], Francesca M. Snoeijen-Schouwenaars[a], In Y. Tan[a], Helenius J. Schelhaas[c], Sylvia Klinkenberg[d,h], Albert P. Aldenkamp[d,e,f,g], and Jos G.M. Hendriksen[d,g,h]

[a]Department of Residential Care, Kempenhaeghe Epilepsy Centre, Heeze, The Netherlands; [b]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands; [c]Department of Neurology, Academic Center for Epileptology Kempenhaeghe, Heeze, The Netherlands; [d]Department of Neurology, Maastricht University Medical Centre, Maastricht, The Netherlands; [e]Faculty of Electrical Engineering, University of Technology, Eindhoven, The Netherlands; [f]Department of Neurology, Ghent University Hospital, Ghent, Belgium; [g]Department of Behavioral Sciences, Kempenhaeghe Epilepsy Centre, Heeze, The Netherlands; [h]Department of Neurological Learning Disabilities, Kempenhaeghe Epilepsy Centre, Heeze, The Netherlands

ABSTRACT

The assessment of intellectual abilities is intensive, time-consuming, and might be considered burdensome for patients. We examined psychometric qualities of short forms (SFs) of the Wechsler Intelligence Scales for Children (WISC-third edition) and for adults (WAIS-fourth edition), in children (n = 986; M_{age} = 10.9) and adults (n = 324; M_{age} = 40.9) with neurological disorders. SF estimates were compared with Full Scale IQ (FSIQ), obtained by a complete administration, for the entire sample and for the subgroups FSIQ < 80 and FSIQ ≥ 80. The FSIQ was correctly identified within ± 7 points in 86% of children and 87% of adults. There were, however, some differences regarding the optimal SF subtest combination between subgroups. Although clinical inferences should not be made, SFs may be useful in research settings to obtain a global estimate of intelligence, and in clinical settings to screen periodically for possible intellectual deterioration.

ARTICLE HISTORY
Received 14 July 2016
Revised 7 November 2016
Accepted 27 December 2016

KEYWORDS
Developmental disability; epilepsy; intelligence; neuropsychological assessment; psychometrics

Introduction

The Wechsler intelligence scales are widely used for assessing the intellectual abilities of children and adults, and are often incorporated as standard components in diagnostic batteries for neuropsychological evaluations.[1] The most recent versions available in the Dutch language are the Wechsler Intelligence Scale for Children-Third Edition, published in 2002 (WISC-III[2,3]), and the Wechsler Adult Intelligence Scale-Fourth Edition, published in 2012 (WAIS-IV[4,5]). The assessment of intellectual abilities is believed to be fundamental for, for example, the diagnosis of intellectual disability (ID), along with a measure of adaptive functioning,[6] or the evaluation of cognitive abilities associated with neurological disorders. Periodic re-evaluation of these abilities is recommended,[7] which is especially relevant for individuals with co-morbid neurological disorders, such as epilepsy, who are at risk of possible deterioration.[8]

Each Wechsler intelligence scale consists of multiple subtests that are used to obtain the summary score full scale IQ (FSIQ) and specific index scores. For instance, the Dutch WISC-III indices include Verbal Comprehension, Perceptual Organization, and Processing Speed, whereas the original WISC-III has an additional index labeled Freedom from Distractibility. The WAIS-IV indices include Verbal Comprehension, Perceptual Reasoning, Processing Speed, and Working Memory. The FSIQ as well as each index score follow a normal distribution with a mean of 100 and standard deviation of 15 IQ points. According to the Wechsler classification system, a FSIQ in the range of 70–79 is labeled as borderline and a FSIQ below 69 as extremely low.[9]

Administration times for the WISC-III and WAIS-IV are approximately 60 minutes and 60–90 minutes, respectively.[10,11] Research on the administration time of the WAIS-III in a clinical sample found an increase of up to 136 minutes, with a mean of 91 minutes.[12] Also, although the discontinue rule applies earlier in individuals who perform more poorly (e.g. individuals with ID), the overall administration times for WISC-III and/or WAIS-IV subtests are not reduced as these individuals often take more time to respond to items and/or need more time in between subtest administrations.[13] In people with ID and/or neurological disorders, the intensive and time-consuming assessment might lead to frustration or fatigue, or might be affected by seizure activity and medication side-effects. Consequently, the validity and reliability of the test outcome might be jeopardized. Briefer methods for estimating global intelligence might, therefore, be desirable, in particular when reassessing patients for screening or rehabilitation purposes, in research settings.[14,15]

To meet the demand for a shortened, but reliable and valid measure of intelligence, the Wechsler Abbreviated Scale of Intelligence was developed for individuals aged 6;0–90;11 (WASI) and later adjusted (WASI-II[15]). The WASI-II consists of four subtests that are similar, although not identical, to their WISC-IV and WAIS-IV counterparts, which measure both verbal and non-verbal abilities. The WASI-II provides accurate and reliable estimates of general intellectual abilities, while reducing administration time to 30 minutes. The WASI-II is, however, a separate instrument, only available in certain languages: English, Norwegian, and Lithuanian.[16] Hence, various shortened versions–hereafter referred to as short forms–have

CONTACT Jans S. van Ool oolJ@kempenhaeghe.nl Department of Residential Care, Kempenhaeghe Epilepsy Centre, P.O. Box 61 5590 AB Heeze, The Netherlands.

been created, mostly based on a specific combination of subtests of Wechsler intelligence scales that are subsequently converted into an estimation of the FSIQ.

Many studies have been conducted to evaluate the accuracy of such short forms and have reported promising results.[17–19] Several studies have, however, indicated possible differences in appropriateness of short forms across populations,[20–22] probably due to population-specific cognitive profiles. Short forms, therefore, need to be validated in specific populations. Furthermore, research on short forms among the population of children and adults with neurological disorders and/or ID has been scarce. As ID and cognitive deterioration are relatively common among people with e.g. epilepsy,[23] gaining insights into the accuracy of short forms in this population would be very useful in order to obtain a valid and reliable screening tool to estimate intelligence.

Van Duijvenbode et al.[24] examined a WASI-based short form, consisting of the four WASI subtests of the WAIS-III, in a sample of adults with mild to borderline ID (n = 117) and concluded that this short form is sufficiently accurate in estimating the FSIQ. Axelrod et al.[25] compared the WASI and the WASI-based short form of the WAIS-III in a sample with neurological disorders (n = 72) and found that the WASI-based short form outperformed the WASI. Hrabok et al.[26] investigated the accuracy of multiple short forms, varying from two to seven subtests, extracted from the more recent WISC-IV (which has not been adapted to Dutch language) in an epilepsy sample (n = 104). There were some differences in short form accuracy between children with FSIQ ≤ 80 and FSIQ > 80, with higher classification rates obtained for subjects with FSIQ ≤ 80 than with FSIQ > 80. This pattern requires validation in an extensive sample of children and should also be investigated in adults, as results are rarely compared in subsamples divided by IQ performance. Also, there is a paucity of index-based short forms which are known to have greater construct validity and provide more clinically relevant information.[27]

The aims of the present study were (1) to investigate the psychometric qualities of the Dutch version of the WISC-III and WAIS-IV short forms in patients with a neurological disorder and (2) to determine which specific short form had the best psychometric qualities for the entire sample as well as for specific subgroups based on FSIQ performance, e.g. FSIQ < 80 (extremely low to borderline[9]) and FSIQ ≥ 80.

Methods

Subjects

The subjects were patients assessed in Kempenhaeghe, a tertiary referral center for epilepsy and other neurological disorders in The Netherlands. We used archival data of clinical referrals to the Department of Behavioral Sciences of Kempenhaeghe to whom the core subtests of the Dutch version of the WISC-III or WAIS-IV had been administered. Data were available for 986 children, collected between May 2009 and December 2014, and for 324 adults, collected between January 2013 and December 2014. The subjects included heterogeneous samples of patients with a variety of epilepsy classifications and congenital and acquired neurological disorders, of which neurofibromatosis, traumatic brain injury, prematurity, and chromosomal abnormalities were most common.

The mean age of subjects who completed the WISC-III was 10.9 years (SD = 2.8, range 6–16) and comprised 57.5% males. Mean FSIQ scores were 89.1 (SD = 17.1, range 45–137). The subjects were divided into two FSIQ-categories, resulting in the following distribution: extremely low to borderline FSIQ (FSIQ < 80, n = 274) and average to high FSIQ (FSIQ ≥ 80, n = 712). The mean age of subjects who completed the core subtests of the WAIS-IV was 40.9 (SD = 16.3, range 16–78) and comprised 49.4% males. They had a mean FSIQ of 83.0 (SD = 16.1, range 46–129). Again, the subjects were divided into three FSIQ-categories: 145 subjects had an extremely low to borderline FSIQ (FSIQ < 80), and 179 had average to high FSIQ (FSIQ ≥ 80).

Procedure

Subjects completed the age-appropriate Wechsler Intelligence Scale as part of a more extensive neuropsychological evaluation of regular medical care. The scales were administered and scored by qualified diagnostic staff members of the Department of Behavioral Sciences according to the standardized procedures outlined in the manuals,[8,9] resulting in age-corrected FSIQ and scaled subtest scores. Prior to this study, the data from patient files were de-identified in order to ascertain anonymous data analyses, and the use for research purposes was approved by the local ethical committee of Kempenhaeghe.

All short form estimates were derived by applying the formulae described by Tellegen and Briggs,[28] using tables B-10 and B-11 from Sattler and Ryan.[29] This method is based on a linear scaling method to yield deviation quotients that maintain the FSIQ's distribution (i.e. mean = 100, standard deviation = 15), and has been used in recent studies on short form accuracy.[26,30] The lower bound of the original WISC-III and WAIS-IV, i.e. FSIQ = 45, was applied if a short form yielded an estimate below 45.

Selection of short forms

For both the WISC-III and WAIS-IV, short form versions were created consisting of a unique combination of subtests. See Table 1 for an overview of the short forms, in which abbreviations are explained. As multiple studies have shown that the accuracy of Wechsler short forms decreases when fewer subtests are included,[19,26] two-subtest combinations were not considered in this study. Taking into account both accuracy and administration time, three- to five-subtest combinations were used. The subtest combinations were a priori selected, primarily based on psychometric results from earlier studies on clinical samples and the WASI-II. This resulted in Short Form 1 (SF1) to SF4 with regard to the WISC-III[18,31–35] and in SF6–SF8 for the WAIS-IV.[15,26,36]

For SF5, SF9 and SF10, the selection of subtests was index-based. As the short forms described above do not perfectly match the index structure of the Dutch WISC-III and WAIS-IV, i.e. one subtest per index, short forms were added by selecting a subtest for

Table 1. Overview of short forms of the WISC-III and WAIS-IV.

Name	Subtests included	Formula[a]	Specification
WISC-III			
SF1_PcInBdVc	Picture Completion, Information, Block Design, Vocabulary	$\Sigma_{subtests}$*1.7 + 32	[31]
SF2_PcSiBdAm	Picture Completion, Similarities, Block Design, Arithmetic	$\Sigma_{subtests}$*1.7 + 32	[18]
SF3_PcSiBdVc	Picture Completion, Similarities, Block Design, Vocabulary	$\Sigma_{subtests}$*1.6 + 36	[32]
SF4_PcInBdVcCo	Picture Completion, Information, Block Design, Vocabulary, Coding	$\Sigma_{subtests}$*1.5 + 25	[33–35]
SF5_BdVcCo	Block Design, Vocabulary, Coding	$\Sigma_{subtests}$*2.2 + 34	One subtest per WISC-III-NL index[b]
WAIS-IV			
SF6_BdMrVcSi	Block Design, Matrix Reasoning, Vocabulary, Similarities	$\Sigma_{subtests}$*1.6 + 36	Based on WASI-II[15]
SF7_MrSsVcSi	Matrix Reasoning, Symbol Search, Vocabulary, Similarities	$\Sigma_{subtests}$*1.7 + 32	[26,36]
SF8_MrSsVcCo	Matrix Reasoning, Symbol Search, Vocabulary, Coding	$\Sigma_{subtests}$*1.7 + 32	[26,36]
SF9_MrSsVcAm	Matrix Reasoning, Symbol Search, Vocabulary, Arithmetic	$\Sigma_{subtests}$*1.7 + 32	One subtest per WAIS-IV index[c]
SF10_VpSsVcAm	Visual Puzzles, Symbol Search, Vocabulary, Arithmetic	$\Sigma_{subtests}$*1.7 + 32	One subtest per WAIS-IV index[d]
SF11_BdVcCo	Block Design, Vocabulary, Coding	$\Sigma_{subtests}$*1.7 + 32	Subtest combination similar to SF5

Note. SF = Short form version; $\Sigma_{subtests}$ = sum of subtest scaled scores.
[a]The formulae used to derive short form IQ estimates were based on a linear equating method described by Tellegen and Briggs (1967).
[b]Selection of subtests was based on highest correlation with the corresponding indices and the FSIQ.
[c]Selection was based on highest correlation with the FSIQ.
[d]Selection was based on highest correlation with the corresponding indices.

each index based on the highest correlation between the subtest and the FSIQ (corrected r = 0.50–0.69) or its index (corrected r = 0.55–0.79). This resulted in SF5, SF9 and SF10. In this way, each subtest itself would still provide information about one's abilities on the corresponding index, which increases the clinical relevance of the short forms by providing more detailed diagnostic information and increases the construct validity.[27,30] In order to compare short form accuracy between the WISC-III and WAIS-IV, a WAIS-IV short form (SF11) was added that included the same subtests as SF5.

Analyses

Psychometric properties of the short forms were evaluated separately for the entire WISC-III and WAIS-IV samples in the study, and were repeated, per Wechsler scale, for the subgroups with a FSIQ < 80 and FSIQ ≥ 80. The results were interpreted according to the minimal criteria set by Donders and Axelrod:[37] (1) a reliability coefficient of the short form ≥ 0.90 (calculated using the methods described by Sattler, 1992); (2) a correlation coefficient between short form and FSIQ ≥ 0.82, corrected for redundancy;[38] and (3) a proportion of short form estimates that fell within the 90% confidence interval of the FSIQ (i.e., within ± 7 IQ points) ≥ 0.81. Additionally, the agreement between short form estimates and FSIQ was calculated using intra-class correlations (ICC) and the differences in mean scores were compared by means of paired t-tests. All analyses were performed using IBM SPSS Statistics 21.

Results

The means, standard deviations and ranges of the FSIQ and the estimated short form IQ scores are presented in Table 2. The IQ scores of all separate IQ categories (e.g. < 80 and ≥ 80) were normally distributed (WISC-IIII: Skewness = –0.75–0.65; Kurtosis = –0.85–0.04; WAIS-IV: Skewness = –0.65–0.52; Kurtosis = –0.88–0.09). The psychometric properties of each short form are summarized in Table 2 for all subjects and in Table 3 for the subgroups.

Table 2. Descriptive statistics and results of the WISC-III and WAIS-IV short forms for all subjects.

	All subjects (n = 986)								
WISC-III	M	SD	Range	r_{sf}	r'	$p(\pm7)$	ICC	95% CI	M_{diff}
FSIQ	89.13	17.10	45–137						
SF1_PcInBdVc	**90.23**	**18.09**	**45–139**	**0.91**	**0.86**	**0.81**	**0.94**	**0.92-0.96**	**1.11*****
SF2_PcSiBdAm	90.93	17.97	45–136	0.92	0.86	0.80	0.93	0.85–0.96	1.80***
SF3_PcSiBdVc	92.22	16.93	45–137	0.91	0.86	0.77	0.96	0.96–0.97	3.09***
SF4_PcInBdVcCo	**88.82**	**18.90**	**45–135**	**0.92**	**0.88**	**0.86**	**0.92**	**0.92–0.93**	**0.31**
SF5_BdVcCo	89.37	17.75	45–135	0.94	0.87	0.74	0.95	0.94–0.95	0.24
	All subjects (n = 324)								
WAIS-IV	M	SD	Range	r_{sf}	r'	$p(\pm7)$	ICC	95% CI	M_{diff}
FSIQ	83.04	16.21	46–129						
SF6_BdMrVcSi	**83.86**	**15.29**	**47–122**	**0.94**	**0.88**	**0.82**	**0.80**	**0.73–0.86**	**0.81***
SF7_MrSsVcSi	**84.64**	**17.39**	**45–127**	**0.93**	**0.87**	**0.86**	**0.74**	**0.45–0.86**	**1.60*****
SF8_MrSsVcCo	82.86	18.24	45–126	0.93	0.85	0.77	0.74	0.16–0.89	0.18
SF9_MrSsVcAm	**83.66**	**16.97**	**45–129**	**0.93**	**0.88**	**0.87**	**0.77**	**0.69–0.83**	**0.62***
SF10_VpSsVcAm	**82.62**	**16.12**	**47–132**	**0.93**	**0.88**	**0.87**	**0.81**	**0.74–0.86**	**-0.42**
SF11_BdVcCo	83.61	16.13	47–122	0.92	0.86	0.80	0.80	0.73–0.86	0.57

Note. FSIQ = Full Scale Intelligence Quotient; SF = Short form version; r_{sf} = reliability of short form (Sattler, 2002); r′ = correlation corrected for redundancy (Levy, 1967); ICC = intra-class correlation; CI = confidence interval; $p(\pm7)$ = proportion of estimated scores that fell within ± 7 points of the FSIQ; M_{diff} = mean difference from FSIQ; Am = Arithmetic; Bd = Block Design; Co = Coding; In = Information; Mr = Matrix Reasoning; Pc = Picture Completion; Si = Similarities; Ss = Symbol Search; Vc = Vocabulary; Vp = Visual Puzzles. Short forms that meet the minimal criteria set by Donders and Axelrod (2002) are displayed in bold. * p < 0.05, ** p < 0.01, *** p < 0.001.

Table 3. Descriptive statistics and results of the WISC-III and WAIS-IV short forms for the subgroups FSIQ < 80 and FSIQ ≥ 80.

	FSIQ < 80 (*n* = 274)								FSIQ ≥ 80 (*n* = 712)							
WISC-III	M	SD	Range	*r'*	*p*(±7)	ICC	95% CI	M_{diff}	M	SD	Range	*r'*	*p*(±7)	ICC	95% CI	M_{diff}
FSIQ	67.86	8.39	45–79						97.31	11.70	80–137					
SF1_PcInBdVc	69.40	11.22	45–92	0.78	0.84	0.86	0.82–0.88	0.65*	98.59	12.38	71–139	0.80	0.81	0.89	0.87–0.91	1.28***
SF2_PcSiBdAm	69.33	11.28	45–95	0.77	0.84	0.84	0.79–0.87	1.47***	99.24	12.24	71–136	0.80	0.79	0.87	0.83–0.90	1.93***
SF3_PcSiBdVc	72.09	10.99	45–89	0.77	0.73	0.78	0.42–0.89	4.23***	99.96	11.60	71–137	0.80	0.78	0.86	0.77–0.91	2.65***
SF4_PcInBdVcCo	65.21	11.29	45–88	0.81	0.82	0.85	0.71–0.91	-2.65***	97.91	12.24	72–135	0.84	0.88	0.93	0.91–0.94	0.60**
SF5_BdVcCo	67.78	11.29	45–91	0.76	0.78	0.80	0.76–0.84	-0.08	97.68	11.79	65–135	0.76	0.73	0.82	0.80–0.85	0.37
	FSIQ < 80 (*n* = 145)								FSIQ ≥ 80 (*n* = 179)							
WAIS-IV	M	SD	Range	*r'*	*p* (±7)	ICC	95% CI	M_{diff}	M	SD	Range	*r'*	*p* (±7)	ICC	95% CI	M_{diff}
FSIQ	68.41	8.11	46–79						94.90	10.37	80–129					
SF6_BdMrVcSi	70.79	8.98	47–89	0.77	0.83	0.65	-0.03–0.86	2.38***	94.44	10.35	66–129	0.77	0.78	0.78	058–0.87	-0.46
SF7_MrSsVcSi	69.65	10.64	45–88	0.77	0.83	0.82	0.61–0.90	1.24**	96.79	11.21	73–127	0.77	0.78	0.83	0.76–0.88	1.89***
SF8_MrSsVcCo	67.05	12.11	45–92	0.78	0.75	0.84	0.78–0.88	-1.35**	95.67	10.78	69–126	0.77	0.72	0.80	0.74–0.85	0.77
SF9_MrSsVcAm	69.09	10.12	45–90	0.80	0.88	0.85	0.69–0.91	0.68	95.46	11.26	69–129	0.82	0.87	0.90	0.86–0.92	0.56
SF10_VpSsVcAm	68.85	9.16	47–90	0.78	0.90	0.83	0.69–0.90	0.44	93.78	11.14	69–132	0.82	0.87	0.89	0.85–0.92	-1.12**
SF11_BdVcCo	69.79	9.78	47–100	0.73	0.84	0.80	0.73–0.86	1.37**	94.81	10.60	69–122	0.75	0.76	0.83	0.78–0.87	-0.09

Note. FSIQ = Full Scale Intelligence Quotient; SF = Short form version; r_{sf} = reliability of short form (Sattler, 2002); r′ = correlation corrected for redundancy (Levy, 1967); ICC = intra-class correlation; CI = confidence interval; *p* (±7) = proportion of estimated scores that fell within ± 7 points of the FSIQ; M_{diff} = mean difference from FSIQ; Am = Arithmetic; Bd = Block Design; Co = Coding; In = Information; Mr = Matrix Reasoning; Pc = Picture Completion; Si = Similarities; Ss = Symbol Search; Vc = Vocabulary; Vp = Visual Puzzles. Short forms that meet the minimal criteria set by Donders and Axelrod (2002) are displayed in bold. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

WISC-III

Entire sample

All short forms met the reliability criterion of $r_{sf} \geq 0.90$ with coefficients varying between 0.91 and 0.94 (see Table 2). The corrected correlations between the short form and FSIQ met the criterion of *r'* ≥ 0.82 for each short form (*r'* = 0.86–0.88). Only SF1 (Picture Completion, Information, Block Design and Vocabulary) and SF4 (Picture Completion, Information, Block Design, Vocabulary and Coding) met the criterion of at least 81% of the short form estimates falling within the 90% confidence interval of the FSIQ. SF1 met this criterion with 81% and SF4 met this criterion with 86%.

The ICC coefficients indicated very strong agreement between the FSIQ and each short form (ICC ≥ 0.92; $p < 0.001$). Paired t-tests suggested significant overestimations of the FSIQ by SF1, SF2 and SF3 (SF1: M_{diff} = 1.11, t (985) = 6.37, $p < 0.001$; SF2: M_{diff} = 1.80, t (985) = 9.90, $p < 0.001$; SF3: M_{diff} = 3.09, t (985) = 17.53, $p < 0.001$). SF4 is the only short form that met the minimal criteria set and whose estimates did not significantly differ from the FSIQ.

Samples FSIQ < 80 and FSIQ ≥ 80

Psychometric properties for the subgroups FSIQ < 80 and FSIQ ≥ 80 are presented in Table 3. The second criterion is only met by SF4, subgroup FSIQ ≥ 80 (*r'* = 0.84). The corrected correlations for SF4, subgroup FSIQ < 80, and SF1, SF2 and SF3, subgroup FSIQ ≥ 80, fell just short (*r'* = 0.80–0.81), whereas the correlation coefficients for all other groups were insufficient (*r'* = 0.76–0.78). For both SF1 and SF4, the third criterion was met in each subgroup (percentage of short form estimates that fell within ± 7 IQ points = 81–88%), leaving SF4 being the only short form that met the minimal criteria set in subgroup FSIQ ≥ 80 and having values that were closest toward meeting the minimal criteria in subgroup FSIQ < 80.

ICC coefficients indicated strong agreement between short form estimates and the FSIQ for all subgroups (ICC = 0.78–0.93). Except for SF5, paired t-test revealed a significant overestimation or underestimation of the FSIQ for all subgroups and short forms (M_{diff} = –2.65–2.65; *ps* = 0.044–< 0.001).

WAIS-IV

Entire sample

All short forms met the reliability criterion of $r_{sf} \geq 0.90$ with coefficients varying between 0.92 and 0.94 (see Table 2). The corrected correlations between the short form and FSIQ met the criterion of *r'* ≥ 0.82 for each short form (*r'* = 0.85–0.88). SF6, SF7, SF9, and SF10 also met the criterion of at least 81% of the short form estimates falling within the 90% confidence interval of the FSIQ (percentage of estimated scores within ± 7 IQ points). The percentages varied between 82% and 87%.

The ICC coefficient indicated a very strong agreement between the FSIQ and SF6, SF10 and SF11 (ICC = 0.80–0.81; $p < 0.001$) and a strong agreement between FSIQ and SF7, SF8 and SF9 (ICC = 0.74–0.77; $p < 0.001$). Paired t-test suggested significant overestimations of the FSIQ by SF6, SF7, and SF9 (SF6: M_{diff} = 0.81, t (323) = 2.50, $p = 0.013$; SF7: M_{diff} = 1.60, t (323) = 5.09, $p < 0.001$; SF9: M_{diff} = 0.62, t (323) = 2.30, $p = 0.022$). SF10 (Visual Puzzles, Symbol Search, Vocabulary, and Arithmetic) is the only short form that met the minimal criteria set and whose estimates did not significantly differ from the FSIQ.

Samples FSIQ < 80 and FSIQ ≥ 80

With regard to the subgroups FSIQ < 80 and FSIQ ≥ 80, the second criterion is only met by SF9 and SF10, subgroup FSIQ ≥ 80 (*r'* = 0.82; see Table 3). The corrected correlation in subgroup FSIQ < 80 fell just short for SF9 (*r'* = 0.80) and was insufficient for all other groups (*r'* = 0.73–0.78). The third criterion was met for SF9 and SF10, for both subgroups (percentage of short form estimates that fell within ± 7 IQ points = 87–90%). For SF6, SF7, and SF11, the cut-off was only met for subgroup FSIQ < 80 (83–84%) and not for subgroup FSIQ ≥ 80. SF9 and SF10 both met the minimal criteria set in subgroup FSIQ ≥ 80 and SF9 had values that were closest toward meeting the minimal criteria in subgroup FSIQ < 80.

ICC coefficients indicated a very strong agreement between FSIQ and SF7, SF8, SF9, SF10 and SF11 (ICC = 0.82–0.90, both subgroups). With respect to SF6, the ICC was moderate for subgroup FSIQ < 80 (ICC = 0.65; $p < 0.001$) and strong for subgroup FSIQ ≥ 80 (ICC = 0.78; $p < 0.001$). For subgroup FSIQ < 80, paired t-tests revealed significant overestimation or underestimations of the FSIQ by SF6, SF7, SF8, and SF11 (M_{diff} = −1.35–2.38; ps = 0.009–< 0.001). For subgroup FSIQ ≥ 80, paired t-tests revealed significant differences between the FSIQ and SF7 and SF10 (SF7: $M_{diff} = 1.89$, $t\ (178) = 4.24$, $p < 0.001$; SF10: $M_{diff} = -1.12$, $t\ (178) = -2.95$, $p = 0.004$).

Comparison of short forms across samples

Only two short forms included an identical subtest combination (BdVcCo) for both the WISC-III and WAIS-IV, i.e. SF5 and SF11. Although statistically driven comparisons could not be made, it appeared that the short forms' results followed a slightly similar pattern overall (see Tables 2 and 3). Both short forms did not meet the third criterion of at least 81% of the short form estimates falling within the 90% confidence interval of the FSIQ for the entire sample. Regarding the subgroups, ICC and corrected correlation coefficients were similar for both subgroups. Although paired t-tests yielded no significant difference between the short forms and the FSIQ in subgroup FSIQ ≥ 80, SF11 significantly overestimated the FSIQ in subgroup FSIQ < 80.

Discussion

In the present study, we examined the psychometric qualities of multiple WISC-III and WAIS-IV short forms in heterogeneous clinical samples of children and adults with neurological disorders, with a specific focus on subsamples with FSIQ < 80 (extremely low to borderline) and FSIQ ≥ 80.

Regarding the WISC-III, only two out of five short forms met the minimal criteria set:[37] SF1 and SF4. With respect to the full sample, SF4 was superior to SF1 as SF1 significantly overestimated the FSIQ with approximately 2 IQ points. Based on results per subgroup, SF4 is still recommended as this short form met the minimal criteria set for subgroup FSIQ < 80 and is very close to meeting the criteria for subgroup FSIQ ≥ 80 ($r' = 0.81$, instead of ≥ 0.82). As described by Evers, Sijtsma, Lucassen and Meijer,[39] a reliability coefficient of ≥ 0.70 of a test is considered sufficient when less important decisions are to be made, for example in research settings. We only encourage the use of short forms in such "less important" situations.

When comparing our results to Hrabok et al.,[26] who have used a similar methodology with the WISC-IV in a sample of children with epilepsy, it appears that mean differences between short form estimates and the FSIQ are in general smaller in our sample and subgroups. Also, the classification accuracy of four-subtest short forms across subgroups appears higher in our subsamples (percentage ± 7 points: 73–86% versus percentage ± 5 points: 33–71%).

Regarding the WAIS-IV, four out of five short forms met the minimal criteria set: SF6, SF7, SF9, and SF10. With respect to the full sample, SF10 seems superior to all others, as the short form estimates of the other short forms significantly differed from the FSIQ. The differences for SF6 and SF9, however, were rather small (0.8 and 0.6 IQ points, respectively). When taking a closer look to the psychometric qualities of these short forms in each subgroup, some differences are noted. Both SF9 and SF10 outperformed SF6 and SF7 in the subgroup FSIQ < 80, and are thus recommended for adults who are expected to have poor intellectual abilities. Although both SF9 and SF10 met the minimal criteria in the subgroup FSIQ ≥ 80, SF10 significantly underestimated the FSIQ with approximately 1 IQ point. SF9 is therefore recommended for adults with average intellectual abilities.

When comparing our results to Girard et al.,[30] who have created two-subtest WAIS-IV short forms based on the same methodology in a sample of clinical referrals, it appears that most of the psychometric properties in our sample are higher. The corrected correlations are slightly higher in our sample (r' = 0.85–0.88 versus 0.76–0.86) as well ass the classification accuracy (percentage ± 7 points: 72–87% versus percentage ± 10 points: 66–84%), whereas ICC coefficients are generally slightly higher in the sample of Girard et al. (ICC = 0.73–0.87 versus 0.74–0.81).

The differences in psychometric qualities across subgroups are especially relevant when re-evaluating a patient of whom previous results are already known, as it can guide the choice for most optimal short form in individual cases. Such differences might be due to variation in reliability coefficients between subtests[10,40] or to an inconsistent subtest profile within subjects. An inconsistent subtest profile is characterized by a significantly better or poorer performance on one or several subtests relative to the other subtests, possibly indicative of strengths and weaknesses in the individual. In our study population, there are several factors known to affect cognitive aspects, such as type of neurological condition, epilepsy characteristics (e.g. seizure type and seizure frequency), use of antiepileptic drugs and co-morbid deficits.[41–43] One might speculate that these factors result in a different subtest profile in individuals with both a neurological condition and ID, compared to subjects without ID, which might explain why some subtests are more representative of the FSIQ than others.

When comparing results from post-hoc paired t-tests for the subgroups FSIQ < 80 and FSIQ ≥ 80, with Bonferroni correction applied, it appeared that with regard to the WISC-III, only subjects with FSIQ < 80 performed significantly lower on Block Design and higher on Similarities (M_{diff} = −0.6 and 1.4, respectively, $p < 0.001$) and only subjects with FSIQ ≥ 80 performed significantly higher on Information, Vocabulary, and Similarities (M_{diff} = 0.2–1.2, p = 0.004 –< 0.001). Regarding the WAIS-IV, subjects with FSIQ < 80 appeared to perform significantly lower on Symbol Search and Coding (M_{diff} = −1.1 and −0.8, respectively, $p < 0.001$) and higher on Vocabulary ($M_{diff} = 0.9$, $p < 0.001$) and subjects with FSIQ ≥ 80 performed significantly higher on Matrix Reasoning and Similarities (M_{diff} = 0.9 and 0.6, $p < 0.001$) and significantly lower on Symbol Search ($M_{diff} = -1.1$, $p < 0.001$).

An interesting finding of this study is that the index-based WAIS-IV short forms (SF9, and SF10) yielded stronger psychometric qualities than the short forms based on previous literature (SF6–SF8, including the WASI-based short form). In SF9 and SF10, the four-index structure has been maintained, providing more clinically relevant information and increasing the construct validity of the short form.[27,30] Against our hypothesis, the WASI-based short form overestimated the FSIQ in subgroups FSIQ < 80 with 2.4 IQ points. The overestimation by a WASI-based short form compared to FSIQ was also found in previous studies conducted in a clinical sample of adults with neurological disorders[25] and a sample with mild to borderline ID,[24] indicating that these subtests might overshadow a poorer performance on other cognitive aspects in these populations.

Our study is limited by several factors. The subjects comprised convenience samples of children and adults with heterogeneous neurological conditions. As the overall mean FSIQ was approximately 89 for WISC-III and 83 for WAIS-IV, the results should not be generalized to children and adults with above-average intellectual functioning and only to a limited extent for average intellectual functioning. Furthermore, the short form estimates were derived from a complete scale administration; it is uncertain what results would have been in a stand-alone version. It could be hypothesized that because of the shorter attention span required, short forms may lead to better performance than would be obtained with the complete scale administration. Results may also be influenced by individual characteristics such as (less) fatigue, (less) frustration or (better) motivation. Future research should re-examine psychometric qualities of short forms administered independently of all FSIQ subtests of the WISC-III or WAIS-IV. Finally, as the present study included the currently available Dutch WISC-III, the results should be validated for the updated Dutch version of the WISC-V.

To conclude, the results of the present study support the use of WISC-III and WAIS-IV short forms in children and adults with neurological conditions in specific situations. We recommend that, where possible, professionals select the short form that yielded highest psychometric results for the specific intelligence category, e.g. if ID is suspected or has previously been determined, one should select a different short form than if one expects average intellectual functioning. The short form IQ obtained should be treated very cautiously and be reported as a global estimation of the FSIQ in clinical situations. Professionals should bear in mind the critique of using short forms,[44,45] especially when evaluating a single individual rather than a group performance. A short form should not be used to make clinical inferences, but can be useful in research settings or in clinical settings to screen for possible deterioration in patients with epilepsy or other neurological disorders.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Lezak MD, Howieson DB, Loring DW. Neuropsychological assessment. New York: Oxford University Press; 2004.
2. Wechsler D. Wechsler intelligence scale for children, 3rd ed. San Antonio, TX: The Psychological Corporation; 1991.
3. Kort W, Compaan EL, Bleichrodt N, Resing WCM, Schittekatte M, Bosmans M. WISC-III [Dutch translation of the Wechsler intelligence scales for children]. Amsterdam: NIP; 2002.
4. Wechsler D. Wechsler adult intelligence scale, 4th ed. San Antonio, TX: Pearson; 2008.
5. Wechsler, D. Wechsler adult intelligence scale, 4th ed. (WAIS-IV-NL) Dutch version. Amsterdam: Pearson; 2012.
6. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.
7. Harris JC. Intellectual disability: a guide for families and professionals. New York, NY: Oxford University Press; 2010.
8. Breuer LEM, Boone P, Bergmans JWM, Mess WH, Besseling RMH, de Louw AJ, et al. Cognitive deterioration in adult epilepsy: does accelerated cognitive ageing exist? Neuroscience and Biobehavioral Reviews 2016;64:1–11.
9. Sattler JM. Assessment of children: cognitive foundations. San Diego, CA: JM Sattler; 2008.
10. Wechsler D. Manual for the Wechsler intelligence scale for children, 3rd ed. San Antonio, TX: The Psychological Corporation; 1991.
11. Wechsler D. WAIS-IV administration and scoring manual. San Antonio, TX: Pearson; 2008.
12. Ryan JJ, Lopez SJ, Werth TR. Administration time estimates for WAIS-III subtests, scales, and short forms in a clinical sample. Journal of Psychoeducational Assessment 1998;16:315–325.
13. Axelrod BN. Administration duration for the Wechsler adult intelligence scale-III and Wechsler memory scale-III. Archives of Clinical Neuropsychology 2001;16:293–301.
14. Sattler JM, Ryan JJ. Assessment of children: revised and updated third edition. WAIS-III supplement. San Diego, CA: Sattler; 1998.
15. Wechsler D. Wechsler abbreviated scale of intelligence, 2nd ed. manual. Bloomington, MN: Pearson; 2011.
16. Content and Translations [Internet]. Pearson clinical assessment. [cited 2016 March 7]. Available from: http://www.pearsonclinical.com/pharma-licensing/content-and-translations.html.
17. Umfleet LG, Ryan JJ, Gontkovsky ST, Morris J. Estimating WAIS-IV indexes: proration versus linear scaling in a clinical sample. Journal of Clinical Psychology 2012;68(4):390–396.
18. Kaufman AS, Kaufman JC, Balgopal R, McLean JE. Comparison of three WISC-III short forms: weighing psychometric, clinical, and practical factors. Journal of Clinical Child Psychology 1996;25 (1):97–105.
19. Hurks P, Hendriksen J, Dek J, Kooij A. Accuracy of short forms of the Dutch Wechsler preschool and primary scale of intelligence, 3rd ed. Assessment 2016;23(2):240–249.
20. Brooks BL, Weaver LE. Concurrent validity of WAIS-III short forms in a geriatic sample with suspected dementia: verbal, performance and full scale IQ scores. Achives of Clinical Neuropsycholy 2005;20:1043–1051.
21. Reid-Arndt SA, Allen BJ, Schopp L. Validation of WAIS-III four-subtest short forms in patients with traumatic brain injury. Applied Neuropsychology 2011;18:291–297.
22. Connery S, Katz D, Kaufman AS, Kaufman NL. Correlations between two short cognitive tests and a WISC-III short form using a sample of adolescent inpatients. Psychological Reports 1996;78(3):1373–1378.
23. Morgan CL, Baxter H, Kerr MP. Prevalence of epilepsy and associated health service utilization and mortality among patients with intellectual disability. American Journal of Mental Retardation 2003;108(5):293–300.
24. Van Duijvenbode N, Didden R, Van den Hazel T, Engels RCME. Psychometric qualities of a tetrad WAIS-III short form for use in individuals with mild to borderline intellectual disability. Developmental Neurorehabilitation 2014;19(1):26–30.

25. Axelrod BN. Validity of the Wechsler abbreviated scale of intelligence and orther very short forms of estimating intellectual functioning. Assessment 2002;9:17–23.
26. Hrabok M, Brooks BL, Fay-McClymont TB, Sherman EMS. Wechsler intelligence scale for children-fourth edition (WISC-IV) short form validity: a comparison study in pediatric epilepsy. Child Neuropsychology: A Journal of Normal and Abnormal Development in Childhood and Adolescence 2014;20(1):49–59.
27. Crawford JR, Anderson V, Rankin PM, MacDonald J. An index-based short-form of the WISC-IV with accompanying analysis of the reliability and abnormality of differences. British Journal of Clinical Psychology 2010;49:235–258.
28. Tellegen A, Briggs PF. Old wine in new skins: grouping wechsler subtests into new scales. Journal of Consulting Psychology 1967;31:499–506.
29. Sattler JM, Ryan JJ. Assessment with the WAIS-IV. San Diego, CA: J. M. Sattler Publishing Company; 2009.
30. Girard TA, Axelrod BN, Patel R, Crawford JR. Wechsler adult intelligence scale-IV dyads for estimating global intelligence. Assessment 2015;22(4):441–448.
31. Thompson A, Browne J, Schmidt F, Boer M. Validitiy of the Kaufman brief intelligence test and a four-subtest WISC-III short form with adolescent offenders. Assessment 1997;4(4):385–394.
32. Franzen M, Smith-Seemiller L. Comparison among seven different short forms of the WISC-III. Oxford, England: Pergamon-Elsevier Science; (pp. 825–825), 2001.
33. Comninel ME, Bordieri JE. Estimating WISC-III scores for special education students using the Dumont-Faro short form. Psychology in the Schools 2001;38(1):11–16.
34. Dumont R, Faro C. A WISC-III short form for learning-disabled students. Psychology in the Schools 1993;30(3):2012–2019.
35. Reiter BA. The application of a WISC-III short form for screening gifted elementary students in Canada. Canadian Journal of School Psychology 2004;19(1–2):191–203.
36. Sattler JM, Dumont R. Assessment of children: WISC-IV and WPPSI-III supplement. San Diego, CA: Jerome M. Sattler; 2004.
37. Donders J, Axelrod BN. Two-subtest estimations of WAIS-III factor index scores. Psychological Assessment 2002;14:360–364.
38. Levy P. The correction for spurious correlation in the evaluation of short-form tests. Journal of Clinical Psychology 1967;23:84–86.
39. Evers A, Sijtsma K, Lucassen W, Meijer RR. The Dutch review process for evaluating the quality of psychological tests: history, procedure, and results. International Journal of testing 2010;10:295–317.
40. Wechsler D. Wechsler adult intelligence scale-fourth edition: technical and interpretive manual. San Antonio, TX: Pearson; 2008.
41. Aldenkamp AP, Bodde N. Behaviour, cognition and epilepsy. Acta Neurology Scandinavica 2005;112(suppl 182):19–25.
42. Ijff DM, Aldenkamp AP. Cognitive side-effects of antiepileptic drugs in children. Handbook of Clinical Neurology 2013;111:707–718.
43. Jakobson A, Kikas E. Cognitive functioning in children with and without Attention-Deficti/Hyperactivty disorder with and without comorbid learning disabilities. Journal of Learning Disabilities 2007;40(3):194–202.
44. Kaufman AS, Kaufman JC, Balgopal R, McLean JE. Comparison of three WISC-III short forms: weighing psychometric, clinical, and practical factors. Journal of Clinical Child Psychology 1996;25 (1):97–105.
45. Donders J. Using a short form of the WISC-III: sinful or smart? Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence 2001;7(2):99–103.