

## Psychometric properties of the Serbian version of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV)

Ljiljana B. Lazarević<sup>1,3</sup>, Goran Knežević<sup>2,3</sup>,  
Marija Mitić<sup>3</sup>, and Dragana Đurić Jočić<sup>4,5</sup>

<sup>1</sup>*Institute of Psychology, Faculty of Philosophy, Belgrade*

<sup>2</sup>*Department of Psychology, Faculty of Philosophy, Belgrade*

<sup>3</sup>*The Laboratory for the study of individual differences, Faculty of Philosophy, Belgrade*

<sup>4</sup>*Faculty of media and communication, Singidunum University*

<sup>5</sup>*Institute of Psychiatry, Clinical Center of Serbia*

The Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) is an individually administered clinical instrument designed for assessment of intellectual abilities of adolescents and adults. The WAIS-IV consists of 15 subtests (10 core and 5 supplemental) reflecting the efficacy of cognitive functioning in four domains (verbal comprehension-VCI, perceptual reasoning-PRI, working memory-WMI, and processing speed-PSI), and general intellectual ability (Full-Scale IQ-FSIQ). The WAIS-IV was administered to a sample of 262 respondents – specifically, 104 respondents from sample representative for the wider Belgrade area, 62 schizophrenic, 63 depressive, and 33 patients with intellectual disability. Psychometric properties of WAIS-IV subtests were analysed within the frameworks Item Response Theory (IRT) and Classical Test

---

Corresponding author: [ljiljana.lazarevic@f.bg.ac.rs](mailto:ljiljana.lazarevic@f.bg.ac.rs)

*Note.* Standardisation of the WAIS-IV scale in Serbia is conducted by Sinapsa Edicije, Belgrade. Authors of this report are editors of the Serbian version of the WAIS-IV Administration and Scoring Manual that will be published by Sinapsa Edicije, Belgrade. Serbian norms will be available after the standardisation of the WAIS-IV is completed. Results presented in the paper were obtained from the data collected in a preliminary study, conducted as part of the project number 179018 funded by the Ministry of Education, Science, and Technological Development, Serbia. Respondents from the general population of the larger Belgrade area that were tested in this study are not included in the standardisation sample, i.e., standardization sample is completely independent from the sample used in this study.

*Acknowledgement.* We would particularly like to thank research assistants and colleagues who helped in data collection: Ivana Anđelković, Jasmina Barišić, Jovana Bjekić, Teodora Đokić, Milica Erić, Zorana Filipović, Vitomir Jovanović, Dušanka Vučinić Latas, Ivana Leposavić, Milan Marković, Ana Mirković, Olivera Novaković, Ivana Papulić, Velimir Popović, Danka Purić, Nataša Radojević, Sunčica Stamenković, Elena Stojimirović, Biljana Šaula, Predrag Teovanović, Olivera Todorović, Maja Todosovski, Mirjana Zebić, Marko Živanović.

\* The first and the second author equally contributed to this paper.

Theory (CTT). Results show that IRT reliability of subtests ranges from .86 (Comprehension and Picture Completion) to .95 (Cancellation), while the IRT reliability of the FSIQ is .94. Overall, all subtests discriminate participants well along the whole continuum of intellectual abilities. The WAIS-IV enables highly reliable assessment of intellectual abilities on Serbian population.

**Key words:** Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV), IRT, psychometric properties

#### Highlights:

- The latest WAIS-IV scale is based on Cattell-Horn-Carroll theory of intelligence.
- The WAIS-IV scale consists of 15 subtests, i.e., 10 core, and 5 supplemental.
- Psychometric properties are highly similar to those on US standardisation sample.
- The WAIS-IV enables reliable assessment of the full span of intellectual abilities.

The Wechsler Intelligence Scales are the most frequently used measures of intelligence worldwide (Lichtenberger & Kaufman, 2009). The Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV; Wechsler, 2008a) was introduced in 2008 where the greatest modification was the elimination of the dual IQ, i.e., verbal and performance IQ scales. This dichotomy was replaced by four composite indices representing functioning in specific cognitive domains, i.e., Verbal Comprehension Index – VCI, Perceptual Reasoning Index – PRI, Working Memory Index – WMI, and Processing Speed Index – PSI<sup>1</sup>. As four specific domains are correlated, they can be used for calculation of the higher order Full-Scale score – FSIQ (Wechsler, 2008b). This structure of the WAIS-IV is representing Wechsler’s view of intelligence which is conceptualised as a global capacity relevant for functioning in everyday life with several specific cognitive domains (Wechsler, 2008b).

Theoretical grounding for the WAIS-IV is the Cattell-Horn-Carroll (CHC) theory (Alfonso, Flanagan, & Radwan, 2005). The CHC theory of cognitive abilities is a combination of Horn and Cattell’s theory of fluid and crystallized intelligence (*Gf-Gc* theory), and Carroll’s theory of cognitive abilities proposing three strata, i.e., general intelligence (*g*), 10 broad cognitive abilities, and more than 100 narrow abilities (McGrew, 2009). The broad cognitive abilities include *Fluid Reasoning (Gf)*, *Comprehension-Knowledge (Gc)*, *Short-term Memory (Gsm)*, *Visual Processing (Gv)*, *Auditory Processing (Ga)*, *Long-term Retrieval*

---

1 Previous version of the scale – WAIS-III, provided scores for Verbal IQ, Performance IQ, and Full-Scale IQ. Important to say is that, in addition to VIQ and PIQ, four secondary indices (Verbal Comprehension, Working Memory, Perceptual Organization, and Processing Speed) were already introduced in WAIS-III. However, WAIS-IV edition explicitly discarded dual IQ.

(*Glr*), *Processing Speed (Gs)*, *Reading and Writing (Grw)*, *Quantitative knowledge (Gq)*, and *Decision/Reaction Time or Speed (Gt)* (Evans, Floyd, McGrew, & Leforgee, 2002). Empirical evidence goes in favour of this viewpoint and indicates that intelligence is composed of specific abilities clustering into higher-order cognitive ability domains (Carroll, 1993; Keith, 1990; Keith & Reynolds, 2010). The CHC theory is considered to be the most validated model of cognitive abilities (Evans et al., 2002; Flanagan, 2000). Lately, the CHC theory has been the most used theoretical framework for the development of intelligence tests and is considered as very important in defining and interpreting cognitive abilities constructs (Alfonso et al., 2005; Newton & McGrew, 2010).

The theoretical model behind WAIS scale accommodates both theoretical cognitive constructs and empirical findings. Up to date, no single measurement covered all CHC abilities (Alfonso et al., 2005; McGrew, 1997). In the WAIS-IV, five broad cognitive abilities from the CHC theory were included, i.e., crystallised knowledge (*Gc*), fluid intelligence (*Gf*), short-term memory (*Gsm*), visual processing (*Gv*), and processing speed (*Gs*) (Alfonso et al., 2005). In developing this revised version of the scale of special importance were concepts of Fluid Intelligence (*Gf*), Working memory (*WM*), and Processing Speed (*PS*) since they are considered as key aspects of intellectual functioning (Lichtenberger & Kaufman, 2009).

The WAIS-IV scale consists of 15 subtests (10 core and 5 supplemental), providing information on four composite indices confirmed by factor analyses (Wechsler, 2008a). The core WAIS-IV subtests<sup>2</sup> are Block Design, Similarities, Digit Span, Matrix Reasoning, Vocabulary, Arithmetic, Symbol Search, Visual Puzzles, Information, and Coding. The supplemental subtests are Letter-Number Sequencing, Figure Weights, Comprehension, Cancellation, and Picture Completion (Table 1). On average, for the administration of the core subtests testing time between 60 and 90 minutes is needed (Wechsler, 2008a, 2008b), while for the administration of all subtests on average 2 hours are needed.

### Enhancements made with the WAIS-IV scale

There were several enhancements made with the latest, WAIS-IV scale, i.e., enhancement of theoretical underpinnings, improvement of psychometric qualities, clinical utility, and user-friendliness (Wechsler, 2008b). In WAIS-IV, nine of the original 11 subtests from VITI (Vekslerov Individualni Test Inteligencije<sup>3</sup>, Berger, Marković, & Mitić, 1991, which is a Serbian adaptation of WAIS-R, Wechsler, 1955) were retained but they were assigned to different groupings (Table 1). Two subtests from VITI that were still used in WAIS-III, but excluded from WAIS-IV were Picture Arrangement and Object Assembly.

2 The subtests are listed according to the order of the administration of the subtests. All core subtests are administered first, and then supplemental. This rule is followed always, unless specific rules provided in the Manual should be applied (like age of the respondent, specific condition of the respondent, etc.).

3 Important to mention is that WAIS-III was not adapted for Serbian population, and that VITI is still in clinical use.

Table 1  
*WAIS-R (VITI – Serbian version of WAIS-R), WAIS-III and WAIS-IV Subtests, and corresponding index scores*

Subtest	WAIS-R (VITI)	WAIS-III <sup>d</sup>	VIQ	PIQ	WAIS-IV	VCI	PRI <sup>a</sup>	WMI	PSI	Dominant CHC Ability
Information	+	+	√		+	√				Gc
Vocabulary	+	+	√		+	√				Gc
Similarities	+	+	√		+	√				Gc
Comprehension <sup>c</sup>	+	+	√		+	√				Gc
Block Design	+	+		√	+		√			Gv
Matrix Reasoning	/	+		√	+		√			Gf
Visual puzzles	/	/			+		√			Gf
Object Assembly	+	+		√	/					Gv
Picture Arrangement	+	+		√	/					Gv
Picture Completion <sup>c</sup>	+	+		√	+		√			Gv
Figure Weights <sup>c</sup>	/	/			+		√			Gf
Arithmetic	+	+	√		+			√		Gf
Digit Span	+	+	√		+			√		Gsm
Letter-Number Sequencing <sup>c</sup>	/	+	√		+			√		Gsm
Coding <sup>b</sup>	+	+		√	+				√	Gs
Symbol Search	/	+		√	+				√	Gs
Cancellation <sup>c</sup>	/	/			+				√	Gs

*Note.* VIQ – verbal subtests; PIQ – performance subtests; VCI – verbal comprehension index; PRI – Perceptual Reasoning Index; WMI – Working Memory Index; PSI – Perceptual Speed Index; Gc – crystallized intelligence; Gv – visual processing; Gf – fluid intelligence; Gsm – short-term memory; Gs – processing speed.

<sup>a</sup> Perceptual Reasoning Index (WAIS-IV) was called Perceptual Organization Index (POI) in WAIS-III.

<sup>b</sup> Coding subtest (WAIS-IV) was called Digit Symbol in WAIS-R (VITI), Digit Symbol Coding in WAIS-III.

<sup>c</sup> Supplemental subtests in WAIS-IV.

<sup>d</sup> In WAIS-III, subtests were assigned to VIQ and PIQ. In addition, in WAIS-III subtests were assigned to one of the four indices like in WAIS-IV.

From WAIS-III, 12 out of 14 subtests were retained in WAIS-IV, and three subtests were added: Figure Weight, Visual Puzzles, and Cancellation. However, in these retained subtests, the item content and administration and/or scoring procedures were revised (Wechsler, 2008b). Importantly, the dependency of time bonuses in subtests of WAIS-IV has been significantly reduced or eliminated so that the influence of declining processing speed on the scores of older adults is lessened (Lee, Gorsuch, Saklofske, & Patterson, 2008). Previous studies show that these changes improved the theoretical foundation of the scale (Grégoire, 2013) and that WAIS-IV scale is superior in measurement, scoring, and structural models to measure FSIQ, compared to WAIS-III (Taub & Benson, 2013). Subtests Picture Arrangement and Object Assembly that were part of previous versions (i.e., VITI, Berger et al., 1991, and WAIS-III, Wechsler, 1997) were dropped because of the lengthy administration, and heavy

subtests manipulativeness. In addition, they contained a lot of pieces that could be damaged or lost, or administered inconsistently. Furthermore, subtest Object Assembly was excluded to decrease dependence on time bonus points (Coalson, Raiford, Saklofske, & Weiss, 2010; Larrabee, 2004; Wechsler, 2008b).

To strengthen the measurement of *Gf*, two subtests as measures of PRI were added, i.e., Figure Weights and Visual Puzzles. As the new measure of Processing Speed, Cancellation subtest was added (Wechsler, 2008b). Visual Puzzles subtest was developed as a Perceptual Reasoning subtest, and it was aimed to be a substitute for Object Assembly. Figure Weights subtest was developed as a Perceptual Reasoning subtest for ages 16:0–69:11 aimed to measure quantitative and analogical reasoning, which involves reasoning emphasising inductive and deductive logic (Carroll, 1993). Cancellation subtest (for ages 16:0–69:11) was based on similar existing tasks (e.g., Geldmacher, Fritsch, & Riedel, 2000) and developed as a supplemental subtest targeting Processing speed.

Another significant change was that Information subtest was chosen as the core subtest over Comprehension which was made a supplemental subtest. The reasons for this decision were mostly psychometric, i.e., reliability, subtest floor, gradient, and ceiling. Moreover, the correlations between WISC-IV UK and WAIS-IV UK composites were slightly better for Information compared to Comprehension. Furthermore, administration time and ease of recording and scoring were in favour of Information subtest over Comprehension (Wechsler, 2008b). Subtests assessing working memory, Digit Span, and Arithmetic were revised (Wechsler, 2008b). Arithmetic was chosen over Letter-Number Sequencing as a core subtest. Revisions in Arithmetic subtests were made to increase the applicability across different cultures and countries (Wechsler, 2008b). Specifically, the reference to currency and to UK system units of measurement was eliminated, some items were reworded to increase clarity, some new items were developed to improve floor, ceiling and difficulty gradient, and time bonuses were eliminated (Wechsler, 2008b).

In the subtest Digit Span, a third task Digit Span Sequencing was added. The main reason for this change was to increase the working memory demands since previous research indicated different cognitive demands for the Digit Span Forward and Digit Span Backward tasks (Reynolds, 1997). Another advantage for the practitioners is that the separate process scores for each of the three tasks allow evaluation of differential performance across the tasks.

Psychometric characteristics of the scale were improved through new norms, increase in FSIQ span, improvement of the subtests and the total scale reliabilities, and new evidence on the validity of the scale (for an overview, see Wechsler, 2008b). Important enhancement achieved with WAIS-IV was an improvement of clinical utility through co-norming with Wechsler Memory Scale-Fourth Edition (WMS-IV, Wechsler, 2009), and conducting series of studies on 13 special groups (e.g., several groups of intellectually dysfunctional, intellectually gifted, traumatic brain injury, borderline intellectual functioning, etc.).

Increase in developmental appropriateness was achieved through adding demo items and examples, and a decrease in the use of professional terms in

instructions since growing body of research suggested that the understanding of the instruction is affected by the mere formulation of instructions and the age of the respondent (Wechsler, 2008b). In addition, developmental appropriateness was achieved through reduction of emphasis on motor skilfulness, and reduction of time bonus (Salthouse, 2004). Furthermore, the adaptation to the participants of advanced age was achieved through reduction of auditory discrimination and visual acuity demands. Finally, testing was made easier for test administrators by shortening the testing time, revising instructions and redesigning record form.

The current study aims to investigate psychometric properties of the WAIS-IV scale adapted for Serbian population. Specifically, the reliability of the subtests, discriminativity of the items and their difficulty was assessed using the IRT, so that quality of the Serbian adaptation of the WAIS-IV is assessed. IRT analysis is considered superior compared to classical test theory especially when developing scales and when psychometric properties of intelligence tests are examined (Bortolotti, Tezza, de Andrade, Bornia, & de Sousa Júnior, 2013). IRT is of special importance for cognitive measures since neither the discrimination along the whole continuum of a dimension nor the item order according to their difficulty, are of such critical importance for non-cognitive measures. The main advantage of IRT is that it relies on the principle of invariance, where the item parameters are not dependent on the respondent's latent traits, and the individual parameters are not dependent on the presented items. Since the central element in the IRT analysis is the item, using IRT in testing psychometric properties of the scale allows comparison of individuals from different populations completing the questionnaires with common items, and it allows for the comparison of individuals submitted to different tests<sup>4</sup>, and between respondents and items (Bortolotti et al., 2013). In other words, with IRT the respondents and the items are located on the same scale, i.e., the respondents are positioned on the latent variable, while the items are positioned according to which level of a latent trait they are best at discriminating. The reliabilities, item sampling adequacy, and homogeneity of the subtests were also analysed within CTT framework.

## Method

### Sample

The WAIS-IV was administered to a sample of 262 respondents – specifically, 104 respondents from sample representative for the larger Belgrade area, 62 schizophrenic, 63 depressive, and 33 patients with intellectual disability from Belgrade clinics. The average age was 39.89 ( $SD = 14.70$ ). Descriptive characteristics of the whole sample and the subsamples are provided in Table 2. The subjects were intentionally sampled from various population

---

4 Within the IRT framework, item parameter estimates are not dependent on the particular sample of examinees, and the person ability estimates are not dependent on the particular sample of test items. Another advantage of using IRT in test comparisons is that it enables comparison of the tests at the item level (unlike CTT, which is test-based and does not assess item parameters independently of a particular sample group).

substrata to increase variability in the achievement scores. Given that the instrument is primarily meant to be used in the clinical setting, affective and schizophrenic patients were included, as the diagnostic groups typical for psychiatric disorders.

Participants from the general population were collected on the sample representative for the Belgrade wider area. The representative sample of the Serbian population was created, and due to financial limitations, only Belgrade wider area was tested. All respondents from the general population were tested by the trained research assistants. Participation of the respondents was on a voluntary basis, and they were paid for their contribution. The sample universe was based on 2002 Census data. A two-stage stratified random representative sample design was employed. In the first stage, sampling units were households, and the method of household selection was a random route technique starting from given address based on the dwelling register. In the second stage, a respondent within a household represented the secondary sampling unit; respondent selection was based on the last birthday in the household in the given age quota.

The household/respondent selection method was defined, and a starting point, i.e., prior to the field survey, the particular address was specified. The maximum number of respondents per starting point was 10. The research assistants began with a given address or at the house nearest to it. If the testing session was successful, the research assistant counted the houses/apartments in a row and walked to the 10th apartment or to the 5th house. If a selected respondent declined to participate or was not found at the address after two attempts, the testing session was considered unsuccessful, and the research assistant chose the next nearest apartment or house. In no case, more than five test sessions were conducted within the same apartment building. In the case that the planned number of testing in a certain street could not be accomplished, the research assistant went to the next nearest street. The sample was created to represent population in Belgrade wider area according to age groups (16-17, 18-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65-69, 70-74, 75-79, 80-84, 85-90), sex, settlement type (urban-rural), and education level. All respondents from the general population were Serbian-native speakers.

Participants from the clinics were inpatients with a diagnosis of schizophrenia and major depression (in remission) from the Institute of Psychiatry and Neuropsychiatric Clinic "Laza Lazarević – Padinska skela, Belgrade. Respondents with intellectual disability were from Military Medical Center – Karaburma and the Institute of Psychiatry, Belgrade. All patients were diagnosed according to ICD-9 (World Health Organisation, 1992). Sociodemographic data were collected for all participants in clinical samples. All patients were Serbian-native speakers. All respondents from the clinical samples were tested by the trained clinicians with long-term experience in administering Wechsler's scales.

## **Ethical Agreement**

Respondents in the sample from clinics were in regular care, and no written consent was therefore required. Each participant was informed about the possible use of the data collected for research purposes, and they could withdraw from participating in the study if wanted. Collected data were made anonymous. The study was conducted in accordance with ethical principles for medical research involving humans (WMA, Declaration of Helsinki).

## **Instrument and procedure**

The WAIS-IV was translated from English to Serbian, and back-translated by authors of this report. In addition, professional proof-editor was consulted during the process of translation. Some items, in verbal subtests, were adapted to fit local requirements with respect to content. Specifically, three items in the Information and four items in the Comprehension (i.e., sayings) subtest had to be adapted to fit local requirements.



In subtest Letter-Number Sequencing, some of the letters used in items had to be changed due to differences between English and Serbian alphabet, but the order of the letters in two alphabets was considered when items were adapted. In addition, special care was paid to create items so that the participants using dominantly Cyrillic or Latin Serbian alphabet (depending on the fact which alphabet was first learned in school) are equally treated. This means that the selection of the letters used in the items was conducted so that the correct answers on the items would always be the same regardless of the alphabet. In other words, adaptation was conducted in a way that preference of the alphabet by the respondent would not influence his/her achievement.

In subtests Similarities, Vocabulary, Digit Span, Arithmetic, Cancellation, Block Design, Matrix Reasoning, Visual Puzzles, Picture completion, Figure Weights, Coding, and Symbol Search content was not changed compared to the original English version of the WAIS-IV (Wechsler, 2008a).

All items of subtests were administered to all respondents, regardless of the general administration instructions, i.e., *discontinue* rule was not applied. The WAIS-IV uses 10 core subtests to produce FSIQ. The Verbal Comprehension Index and Perceptual Reasoning Index are each composed of three core subtests, while Working Memory Index and Processing Speed Index are each composed of two core subtests. Supplemental subtests are provided to substitute core subtest if necessary, but three supplemental subtests (Figure-Weights, Letter-Number Sequencing, and Cancellation) are not available for 70- to 90-year-olds and these were not administered to participants older than 70 years. The WAIS-IV subtests along with corresponding index scores are presented in Table 1.

## Results

Descriptive parameters, i.e., social, demographic, and cognitive characteristics for the subsamples, and the whole sample are displayed in Table 2.

Table 2  
*Social, demographic, and cognitive characteristics of the study sample*

	General population	Schizophrenia	Depression	Intellectual disability	Whole sample
<i>N</i>	104 (39%)	62 (24%)	63 (24%)	33 (13%)	262 (100%)
Sex, % male	47 (45%)	39 (63%)	18 (29%)	19 (57%)	123 (47%)
Age, <i>Mean (SD)</i>	42.6 (18.6)	35.9 (10.6)	42.4 (10.5)	33.7 (10.9)	39.9 (14.7)
Handedness, % right handed	100 (96.2%)	61 (98.4%)	61 (96.8%)	32 (97%)	254 (96.9%)
Education level					
Elementary school	11 (10.6%)	4 (6.5%)	6 (9.5%)	21 (63.6%)	42 (16.1%)
High school	49 (47.1%)	44 (71%)	45 (71.4%)	12 (36.4%)	150 (57.3%)
Bachelor degree	12 (11.5%)	6 (9.7%)	5 (7.9%)	/	23 (8.7%)
University degree	32 (30.7%)	8 (12.9%)	7 (11.1%)	/	47 (17.9%)

Table 3 presents descriptive parameters for all WAIS-IV subtests on the subsample and the whole sample of respondents. All scores were calculated according to the US norms (Wechsler, 2008a). Our results show that average FSIQ for the respondents from the general population (Belgrade wider area) is 102.5 where scores were calculated according to the US norms.



Table 3  
 Descriptive statistics and results of ANOVA with LSD post hoc test for subtests on four groups from the sample

Subtests	General population (N = 104)		Schizophrenia (N = 62)		Major Depression (N = 63)		Intellectual disability (N = 33)		F (df = 3, 258)	P	LSD post hoc
	M	SD	M	SD	M	SD	M	SD			
Information (VCI)	98.7	35.3	96.3	22.6	87.3	30.5	41.9	22.3	31.85	.000	1.2>3>4
Vocabulary (VCI)	93.0	14.1	89.8	8.7	89.6	10.5	71.9	8.6	28.42	.000	1.2.3>4
Similarities (VCI)	95.9	17.5	89.5	14.7	88.8	16.3	64.8	9.7	32.77	.000	1>2.3>4
Comprehension <sup>a</sup> (VCI)	99.7	18.2	83.3	13.8	88.9	11.5	64.5	10.6	50.13	.000	1>3>4; 3>2>4
Block Design (PRI)	97.4	17.6	86.2	12.8	84.7	11.6	74.3	8.7	26.28	.000	1>2.3>4
Matrix Reasoning (PRI)	107.7	36.2	83.4	22.9	81.4	24.4	53.2	18.1	33.67	.000	1>2.3>4
Visual puzzles (PRI)	109.6	41.7	77.1	29.1	82.9	27.9	51.9	22.0	29.37	.000	1>2.3>4
Picture Completion <sup>a</sup> (PRI)	102.0	23.4	83.6	23.7	88.8	28.7	56.6	21.4	29.75	.000	1>2.3>4
Figure Weights <sup>ab</sup> (PRI)	108.0	48.9	85.0	32.1	77.5	30.9	48.4	21.3	22.62	.000	1>2.3>4
Arithmetic (WMI)	106.4	37.7	81.2	28.9	89.8	29.3	48.9	25.5	27.95	.000	1>2.3>4
Digit Span (WMI)	104.4	32.3	89.3	22.2	80.0	23.2	55.2	17.5	32.06	.000	1>2.3>4
Letter-Number Sequencing <sup>ab</sup> (WMI)	103.7	36.5	80.9	26.9	84.5	22.7	57.6	16.2	23.34	.000	1>2.3>4
Coding (PSI)	98.8	15.7	82.5	9.9	84.8	10.1	76.5	8.2	41.02	.000	1>2.3>4
Symbol Search (PSI)	98.8	16.2	82.7	8.8	84.5	9.7	76.8	9.1	39.70	.000	1>2.3>4
Cancellation <sup>ab</sup> (PSI)	107.3	24.3	73.2	19.8	81.2	20.5	67.6	18.3	48.83	.000	1>3>2.4
FSIQ – Full Scale IQ	102.5	17.7	83.4	10.2	83.8	11.3	65.1	4.9	75.99	.000	1>2.3>4

Note. <sup>a</sup> Supplemental subtests; <sup>b</sup> Subtests for ages 16:0–69:11 only.

Compared to the results obtained on the special groups in the US standardisation, our respondents from major depression group show lower FSIQ achievement (98.6 in the US, 83.8 our sample). Respondents from the intellectual disability group show performance in line with those obtained in the US standardisation (Wechsler, 2008b). In the US standardisation study, data were not collected on the schizophrenic patients, but our data are in line with some other studies using WAIS-IV with clinical groups (e.g., Bulzacka, Meyers, Boyer, Le Gloahec, Fond, Szöke, et al., 2016).

Intercorrelations obtained between subtests and total score are provided in Table 4. As can be seen, results indicate that subtests corresponding to the specific index are correlating moderately to highly and that all subtests correlate highly with the general score (FSIQ).

Table 4  
*Intercorrelations of the WAIS-IV subtests*

	IN	VC	SI	CO	BD	MA	VP	PC	FW	AR	DS	LN	COD	SS	CA	FSIQ
IN		.74	.71	.70	.48	.58	.49	.54	.55	.59	.54	.49	.45	.42	.27	.74
VC			.71	.71	.41	.54	.44	.51	.49	.54	.54	.47	.42	.42	.25	.70
SI				.72	.48	.60	.54	.59	.55	.58	.59	.55	.52	.51	.39	.75
CO					.50	.60	.51	.54	.54	.60	.57	.52	.53	.50	.41	.78
BD						.72	.72	.63	.68	.53	.53	.49	.65	.64	.57	.81
MA							.71	.68	.74	.62	.67	.62	.71	.69	.59	.79
VP								.65	.72	.55	.62	.60	.65	.70	.57	.76
PC									.65	.53	.58	.56	.61	.60	.55	.70
FW										.63	.62	.63	.67	.64	.55	.74
AR											.61	.64	.61	.59	.51	.73
DS												.74	.65	.63	.57	.71
LN													.64	.62	.61	.64
COD														.80	.71	.77
SS															.71	.74
CA																.62

*Note.* IN – Information; VC – Vocabulary; SI – Similarities; CO – Comprehension; MA – Matrix Reasoning; VP – Visual Puzzles; PC – Picture Completion; FW – Figure Weights; AR – Arithmetic; DS – Digit Span; LN – Letter-Number Sequencing; COD – Coding; SS – Symbol Search; CA – Cancellation. All correlations are significant at .01 level.

Psychometric properties of WAIS-IV subtests were analysed with WinSteps software (version 3.58.1) for IRT analysis. One-parameter IRT model was used for the analysis of psychometric properties. Analysis showed that FSIQ calculated from all subtests has excellent reliability, i.e., .94. Results show that IRT reliability of subtests ranges from .86 (Comprehension and Picture Completion) to .95 (Cancellation), except for Similarities that had slightly lower IRT reliability coefficient (.82). In addition to IRT, Cronbach alphas for subtests and reliability for the FSIQ calculated as composite reliability for congeneric measures were calculated (Raykov, 1997). When only core subtests are considered, a composite reliability coefficient of total score remains excellent, i.e., .94. A Keiser-Meyer-Olkin measure of representativeness – *KMO*, and Momirović's measure of homogeneity – *h2* were calculated using an RTT10G algorithm (Knežević & Momirović, 1996). All psychometric statistics are displayed in Table 5. Our results on the reliability of the subtests and the whole scale are in line with results obtained in the US standardisation showing that reliability<sup>5</sup> coefficients range from .78 for Cancellation to .94 (Vocabulary), while the total scale on the US sample had reliability coefficient .98 (Wechsler, 2008b).

5 In US standardization IRT analyses were not conducted, but split-half coefficients and Cronbach alphas were used depending on the subtest.

Table 5  
*IRT reliabilities, Misfit order measures, KMO measures of representativeness, Cronbach's  $\alpha$ , and homogeneity of the subtests*

Subtests	IRT reliability	Infit	Outfit	KMO	Cronbach's $\alpha$	$h^2$
Information (VCI)	.91	0.99	0.98	.97	.91	.72
Vocabulary (VCI)	.87	0.99	1.10	.93	.86	.60
Similarities (VCI)	.82	1.0	0.97	.89	.81	.61
Comprehension (VCI)	.86	0.99	1.06	.92	.71	.71
Block Design (PRI)	.92	0.80	0.78	.98	.85	.54
Matrix Reasoning (PRI)	.91	0.95	0.98	.97	.90	.72
Visual puzzles (PRI)	.89	0.89	0.74	.97	.87	.76
Picture Completion (PRI)	.86	0.96	0.90	.93	.82	.48
Figure Weights (PRI)	.92	0.89	0.77	.97	.91	.63
Arithmetic (WMI)	.88	0.91	0.77	.95	.85	.60
Digit Span (WMI)	.92	1.01	0.69	.95	.89	.41
Letter-Number Sequencing (WMI)	.89	0.98	0.89	.95	.87	.46
Coding (PSI) <sup>b</sup>	/	/	/	/	/	/
Symbol Search (PSI) <sup>b</sup>	/	/	/	/	/	/
Cancellation (PSI)	.95	0.86	0.87	.93	.92	.93
FSIQ	.94	1.22	1.19	/	.96 <sup>a</sup>	/

*Note.* Infit – inlier-pattern-sensitive fit statistic (mean-square); outfit – outlier-sensitive fit statistic (mean-square); KMO– Kaiser-Meier-Olkin measure of sampling adequacy,  $h^2$ – Momirovic's measure of homogeneity (Knežević & Momirović, 1996).

<sup>a</sup> Reliability for FSIQ was calculated as composite reliability for congeneric measures (Raykov, 1997).

<sup>b</sup> Psychometric parameters are not displayed since the data were not entered for each item, but an overall number of correct answers were registered instead.

Homogeneity measures are high and indicate that almost all subtests are unidimensional, except for the 3 subtests, i.e., Digit Span, Letter-Number Sequencing, and Picture Completion. This result indicates that some subtests are capturing more than one (CHC) ability, which is not surprising since in these subtests several abilities are engaged, e.g., short-term memory, working memory, and logical reasoning.

Overall, our results show that all subtests discriminate participants well along the whole continuum of intellectual abilities (item-person map and item misfit order measures for all items of all subtests are available in Supplementary materials 1 (<https://osf.io/ydkbm/>) and 2 (<https://osf.io/9uver/>)). As can be seen from item-person maps, all subtests cover the whole range of intellectual abilities and discriminate participants well along the continuum. Analysis of overall misfit order measures for subtests shows that all have very good values, i.e., values range between 0.8 and 1.3 which are considered excellent. Analysis of item misfit order shows that almost all items in all subtests have adequate infit and outfit measures, which indicates that items are good. In the subtests, Digit Span, and Similarities all items demonstrated good measures, both infit and outfit (see Supplementary materials 2 (<https://osf.io/9uver/>)).

In subtest Comprehension, the first item had poor outfit measures, which can indicate that participants with higher abilities can unexpectedly have poor performance. It is possible that the item is unexpectedly easy and that respondents could be confused and provide the response of lower quality. In the subtest Information, items 3 and 5 had poor outfit measures. We have realized that in the case of item 3 improvement in the translation should be done and that it could prevent this item from having poor outfit measure. When it comes to item 5, it is possible that its position in the administration order should be changed and moved more to the middle part of the test.

In subtests where the first item is not a start point, i.e., Block design, Matrix Reasoning, Arithmetic, Visual Puzzles, Figure Weights, and Picture Completion, items preceding start point have poor outfit measures. This indicates that on these items people with higher abilities can occasionally have poor performance, or indicate lucky guesses of extreme respondents. Therefore, as prescribed, these items should not be administered unless reverse rules are required due to poor achievement on the start items.

In the subtest Vocabulary, items 13 and 30 had poorer outfit measures which suggest that for these two items unexpectedly low achievement of highly intelligent respondents can occur, or that respondents who are not high-achievers provide a correct response by lucky guessing. Based on these results, item 13 in Vocabulary should probably be changed. The last item in Vocabulary subtest (i.e., item no. 30 – Palliative) discriminates highly intelligent respondents well (respondents of FSIQ 142 have 50% chance to answer it), but occasionally less intelligent people answer correctly. This could be due to the content of the item itself since the use of the word is widespread in everyday life and language (e.g., palliative care used as the last help for terminal patients), but providing an accurate definition of the term sometimes can be difficult even for the respondents with above-average cognitive abilities. However, *discontinue* rule (after 3 consecutive scores of 0) lowers significantly the possibility of a person with lower abilities to answer correctly on this item.

In the subtest Letter-Number Sequencing, items 2, 30, and 9 had poorer outfit measures. In the case of item 2, reasons for poor outfit measure can lie in the fact that the item itself is extremely easy, and that highly capable respondents are confused with the easiness of the task and provide the low-quality answer. For the last item, i.e., number 30, poorer outfit measure should not pose a problem since *discontinue* rule that must be applied in testing would significantly lower the possibility of lucky guessing. Reasons for poorer outfit index of the item number 9 in this subtest are still not clear, and future studies should inspect this issue in more detail. It is also important to note that the reasons for poorer misfit measures of some items of the subtests may be caused by only a few observations. Therefore, before making a final judgement on the potential lower quality of some items in subtests, data on the complete standardization sample should be considered.

For Symbol Search and Coding reliability was not calculated because the data were not entered for each item, but an overall number of correct answers

were registered instead. In the case of Cancellation, reliability was calculated from the two sums of the two types of correctly identified patterns. Analyses showed that separation index (the signal-to-noise ratio, i.e., the ratio of “true” variance to error variance) of the FSIQ is 3.89, which indicates highly effective discriminative power of the test. It means that the scale can discriminate at least four groups of respondents according to their cognitive abilities.

## Discussion

All results from this study demonstrated that, overall, the WAIS-IV scale enables highly reliable assessment of the full span of intellectual abilities from intellectual disability to intellectual giftedness. These results are in line with other studies (for an overview see, Wechsler, 2008b) showing that the WAIS-IV scale allows for extensive, high-quality assessment of cognitive abilities. This result is a telling one regarding the test itself, but also regarding the nature of intelligence, having in mind that the Serbian adaptation of the WAIS-IV assumed only slight changes in the original items (see Instrument section) and that the scores were calculated relying on the US norms. (Interestingly, the average IQ for the rather small, but random sample from the large Belgrade area, based on these norms was 102.5 – very close to expected value of 100 for general population).

Our analyses showed that smaller changes in some of the subtests could improve psychometric properties of some subtests. Namely, in the subtest Vocabulary, one item (no. 13) should be changed in the final version of the Serbian WAIS-IV. Future studies should clarify the quality of the item number 30 in the Vocabulary subtests, and of the items 2, 9, and 30 in the subtest Letter-Number Sequencing. These and other issues (such as the introduction of a couple of new items in the verbal reasoning tests to improve discrimination of the respondents with high abilities) will be addressed in the ongoing Serbian standardisation of the instrument.

The special quality of the WAIS-IV is the variety of high-quality supplemental subtests which allows the scale to be used when assessing cognitive abilities of different clinical groups, and to obtain additional information on cognitive functioning. The high quality of all subtests in Serbian adaptation indicates that the use of the core WAIS-IV subtests, when the time for administration is limited or when testing respondents older than 70 years would also be appropriate. Moreover, further reduction of the number of subtests in assessing IQ is a frequent trend, especially in research settings (Axelrod, 2002). There are recommendations for use of seven subtests versions of WAIS-III and WAIS-IV (Meyers, Zellinger, Kockler, Wagner, & Miller, 2013; Wymer, Rayls, & Wagner, 2003), as well as for assessment with four subtests (WASI, The Psychological Corporation, 1999; Axelrod, 2002). Due to high reliability of these subtests, correlations of IQ measures derived from these abbreviated versions and full WAIS scales are above .95 (Ryan, Carruthers, Miller, Souheaver, Gontkovsky, & Zehr, 2003).

Consistent with previous findings, in the group of schizophrenic patients processing speed was most affected, followed by working memory, perceptual reasoning and verbal comprehension (Bulzacka et al., 2016; Michel, Goldberg, Heinrichs, Miles, Ammari, & McDermid Vaz, 2013). Previous results showed that individuals with major depression consistently show lower performance on processing speed tasks relative to other types of cognitive tasks and compared to nonclinical respondents (Gorlyn, Keilp, Oquendo, Burke, Sackeim, & John Mann, 2006; Gualtieri, & Johnson, 2006; Iverson, Turner, & Green, 1999; Mandelli, Serretti, Colombo, Florita, Santoro, Rossini, et al., 2006). In addition, consistent with previous findings (Naismith, Hickie, Turner, Little, Winter, Ward, et al., 2003) our respondents from major depression groups had the lower achievement on timed subtests, e.g., Figure Weights and Visual Puzzles.

When discussing achievement of the respondents from the intellectual disability group, our results are in line with the results obtained in US standardisation of the WAIS-IV (Wechsler, 2008b). Like our results, results of previous studies demonstrated that individuals with intellectual disability have the poorest performance on Working Memory tasks (Baddeley & Jarrold, 2007; Conners, Rosenquist, Arnett, Moore, & Hume, 2008; Van der Molen, Van Luit, Jongmans, & Van der Molen, 2007), and Perceptual Reasoning tasks (Caffrey & Fuchs, 2007; Fontana, 2004).

Future studies should be conducted on the sample representative of the whole Serbian population, and provide norms. In addition, studies should give evidence of the factor structure of the scale, and the predictive validity of WAIS-IV in the assessment of cognitive strengths and weaknesses of respondents from different clinical groups.

## References

- Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 185–202). New York, NY: Guilford Press.
- Axelrod, B. N. (2002). Validity of the Wechsler abbreviated scale of intelligence and other very short forms of estimating intellectual functioning. *Assessment, 9*(1), 17–23. doi: <https://doi.org/10.1177/1073191102009001003>
- Baddeley, A., & Jarrold, C. (2007). Working memory and Down syndrome. *Journal of Intellectual Disability Research, 51*(12), 925–931. doi: [10.1111/j.1365-2788.2007.00979.x](https://doi.org/10.1111/j.1365-2788.2007.00979.x)
- Berger, J., Marković, M., & Mitić, M. (1991). *Vekslerov Individualni Test Inteligencije – Priručnik* [Wechsler Adult Intelligence Scale – Manual]. Beograd, Društvo Psihologa Srbije.
- Bortolotti, S. L. V., Tezza, R., de Andrade, D. F., Bornia, A. C., & de Sousa Júnior, A. F. (2013). Relevance and advantages of using the item response theory. *Quality & Quantity, 47*(4), 2341–2360. doi: [10.1007/s11135-012-9684-5](https://doi.org/10.1007/s11135-012-9684-5)
- Bulzacka, E., Meyers, J. E., Boyer, L., Le Gloahec, T., Fond, G., Szöke, A., ... & Schürhoff, F. (2016). WAIS-IV seven-subtest short form: validity and clinical use in schizophrenia. *Archives of Clinical Neuropsychology, 31*(8), 915–925. doi: <https://doi.org/10.1093/arclin/acw063>
- Caffrey, E., & Fuchs, D. (2007). Differences in performance between students with learning disabilities and mild mental retardation: Implications for categorical instruction. *Learning Disabilities Research & Practice, 22*(2), 119–128. doi: [10.1111/j.1540-5826.2007.00236.x](https://doi.org/10.1111/j.1540-5826.2007.00236.x)



- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.
- Coalson, D. L., Raiford, S. E., Saklofske, D. H., & Weiss, L. G. (2010). WAIS-IV: Advances in the Assessment of Intelligence. In Weiss, L. G., Saklofske, D. H., Coalson, D. L., & Raiford, S. E. (Eds.) *WAIS-IV Clinical use and interpretation*. UK, Elsevier.
- Connors, F. A., Rosenquist, C. J., Arnett, L., Moore, M. S., & Hume, L. E. (2008). Improving memory span in children with Down syndrome. *Journal of Intellectual Disability Research, 52*(3), 244–255. doi: 10.1111/j.1365-2788.2007.01015.x
- Evans, J. J., Floyd, R. G., McGrew, K. S., & Leforgee, M. H. (2002). The relations between measures of Cattell–Horn–Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review, 31*(2), 246–262.
- Flanagan, D. P. (2000). Wechsler-based CHC cross-battery assessment and reading achievement: Strengthening the validity of interpretations drawn from Wechsler test scores. *School Psychology Quarterly, 15*, 295–329. doi:10.1037/h0088789
- Fontana, S. (2004). The Instructional Approach: Improving the Performance of a Person with Moderate Mental Retardation in a Reasoning Task. *Journal of Cognitive Education and Psychology, 4*(1), 148–150. doi: <https://doi.org/10.1891/194589504787382857>
- Geldmacher, D. S., Fritsch, T., & Riedel, T. M. (2000). Effects of stimulus properties and age on random-array letter cancellation tasks. *Aging, Neuropsychology, and Cognition, 7*(3), 194–204. doi: [http://dx.doi.org/10.1076/1382-5585\(200009\)7:3;1-Q;FT194](http://dx.doi.org/10.1076/1382-5585(200009)7:3;1-Q;FT194)
- Gorlyn, M., Keilp, J. G., Oquendo, M. A., Burke, A. K., Sackeim, H. A., & John Mann, J. (2006). The WAIS-III and major depression: absence of VIQ/PIQ differences. *Journal of clinical and experimental neuropsychology, 28*(7), 1145–1157. doi: <http://dx.doi.org/10.1080/13803390500246944>
- Grégoire, J. (2013). Measuring components of intelligence: Mission impossible? *Journal of Psychoeducational Assessment, 31*, 138–147. doi:10.1177/0734282913478034
- Gualtieri, C. T., & Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology, 21*(7), 623–643. doi: <https://doi.org/10.1016/j.acn.2006.05.007>
- Iverson, G. L., Turner, R. A., & Green, P. (1999). Predictive validity of WAIS-r VIQ-PIQ splits in persons with major depression. *Journal of Clinical Psychology, 55*(4), 519–524. doi: 10.1002/(SICI)1097-4679(199904)55:4<519::AID-JCLP14>3.0.CO;2-B
- Keith, T. Z. (1990). Confirmatory and hierarchical confirmatory analysis of the Differential Ability Scales. *Journal of Psychoeducational Assessment, 8*, 391–405. doi: <https://doi.org/10.1177/073428299000800314>
- Keith, T. Z., & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we learned from 20 years of research. *Psychology in the Schools, 47*(7), 635–650. doi: 10.1002/pits
- Knežević, G., & Momirović, K. (1996). RTT9G, program za analizu metrijskih karakteristika kompozitnih memorijskih instrumenata [RTT9G, software for metric properties analysis of composite measuring instruments]. In P. Kostić (Ed.), *Problemi merjenja u psihologiji, 2*. Beograd: Centar za primenjenu psihologiju.
- Larrabee, G. (2004). A review of clinical interpretation of the WAIS-III and WMS-III: Where do we go from here and what should we do with the WAIS-IV and WMS-IV? *Journal of Clinical and Experimental Psychology, 26*(5), 706–717.
- Lee, H. F., Gorsuch, R. L., Saklofske, D. H., & Patterson, D. A. (2008). Cognitive differences for ages 16 to 89 years (Canadian WAIS-III). *Journal of Psychoeducational Assessment, 26*, 382–394. <https://doi.org/10.1177/0734282908323628>
- Lichtenberger, E. O., & Kaufman, A. S. (2009). *Essentials of WAIS-IV Assessment*. New York, NY: Wiley
- Mandelli, L., Serretti, A., Colombo, C., Florita, M., Santoro, A., Rossini, D., ... & Smeraldi, E. (2006). Improvement of cognitive functioning in mood disorder patients with depressive



- symptomatic recovery during treatment: an exploratory analysis. *Psychiatry and Clinical Neurosciences*, 60(5), 598-604. doi: 10.1111/j.1440-1819.2006.01564.x
- McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 151–180). New York: Guilford Press.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37, 1–10. doi:10.1016/j.intell.2008.08.004
- Meyers, J. E., Zellinger, M. M., Kockler, T., Wagner, M., & Miller, R. M. (2013). A validated seven-subtest short form for the WAIS-IV. *Applied Neuropsychology: Adult*, 20(4), 249–256. doi: 10.1080/09084282.2012.710180
- Michel, N. M., Goldberg, J. O., Heinrichs, R. W., Miles, A. A., Ammari, N., & McDermid Vaz, S. (2013). WAIS-IV profile of cognition in schizophrenia. *Assessment*, 20(4), 462–473. doi: 10.1177/1073191113478153
- Naismith, S. L., Hickie, I. B., Turner, K., Little, C. L., Winter, V., Ward, P. B., ... & Parker, G. (2003). Neuropsychological performance in patients with depression is associated with clinical, etiological and genetic risk factors. *Journal of clinical and experimental neuropsychology*, 25(6), 866–877. doi: http://dx.doi.org/10.1076/jcen.25.6.866.16472
- Newton, J. H., & McGrew, K. S. (2010). Introduction to the special issue: Current research in Cattell-Horn-Carroll-based assessment. *Psychology in the Schools*, 47(7), 621–634. doi: 10.1002/pits.20495
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, 21(2), 173–184. doi: https://doi.org/10.1177/01466216970212006
- Reynolds, C. R. (1997). Forward and backward memory span should not be combined for clinical analysis. *Archives of Clinical Neuropsychology*, 12(1), 29–40. doi: https://doi.org/10.1093/arclin/12.1.29
- Ryan, J. J., Carruthers, C. A., Miller, L. J., Souheaver, G. T., Gontkovsky, S. T., & Zehr, M. D. (2003). Exploratory factor analysis of the Wechsler Abbreviated Scale of Intelligence (WASI) in adult standardization and clinical samples. *Applied Neuropsychology*, 10(4), 252–256. doi: http://dx.doi.org/10.1207/s15324826an1004\_8
- Salthouse, T. A. (2004). What and when of cognitive aging. *Current directions in psychological science*, 13(4), 140–144. doi: https://doi.org/10.1111/j.0963-7214.2004.00293.x
- Taub, G. E., & Benson, N. (2013). Matters of consequence: An empirical investigation of the WAIS-III and WAIS-IV and implications for addressing the Atkins intelligence criterion. *Journal of Forensic Psychology Practice*, 13(1), 27–48. doi:10.1080/15228932.2013.746913
- The Psychological Corporation. (1999). *Manual for the Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Author
- Van der Molen, M. J., Van Luit, J. E. H., Jongmans, M. J., & Van der Molen, M. W. (2007). Verbal working memory in children with mild intellectual disabilities. *Journal of Intellectual Disability Research*, 51(2), 162–169. doi: 10.1111/j.1365-2788.2006.00863.x
- Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale (WAIS)*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition: Administration and scoring manual*. San Antonio, TX: Psychological Corporation
- Wechsler, D. (2008a). *Wechsler Adult Intelligence Scale—Fourth Edition: Administration and scoring manual*. San Antonio, TX: Psychological Corporation
- Wechsler, D. (2008b). *Wechsler Adult Intelligence Scale—Fourth Edition: Technical and Interpretative Manual*. San Antonio, TX: Psychological Corporation
- Wechsler, D. (2009). *Wechsler Memory Scale—Fourth Edition: Technical and Interpretative Manual*. Pearson; San Antonio, TX.

- World Health Organisation. (1992). *International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10)*. Geneva: WHO.
- Wymer, J. H., Rayls, K., & Wagner, M. T. (2003). Utility of a clinically derived abbreviated form of the WAIS-III. *Archives of Clinical Neuropsychology, 18*(8), 917–927. doi:10.1016/S0887-6177(02)00221-4

RECEIVED 09.07.2017.

REVISION RECEIVED 01.10.2017.

ACCEPTED 02.11.2017.

## Psihometrijske karakteristike srpske verzije Vekslerove skale za procenu inteligencije kod odraslih – četvrta revizija (WAIS-IV)

Ljiljana B. Lazarević<sup>1,3</sup>, Goran Knežević<sup>2,3</sup>, Marija Mitić<sup>4</sup>, i  
Dragana Đurić Jočić<sup>4,5</sup>

<sup>1</sup>Institut za psihologiju, Filozofski fakultet, Beograd

<sup>2</sup>Odeljenje za psihologiju, Filozofski fakultet, Beograd

<sup>3</sup>Laboratorija za istraživanje individualnih razlika, Filozofski fakultet, Beograd

<sup>4</sup>Fakultet za medije i komunikaciju, Univerzitet Singidunum

<sup>5</sup>Institut za psihijatriju, Klinički centar Srbije

Četvrta revizija Vekslerove skale za procenu inteligencije kod odraslih (WAIS-IV) je klinički instrument koji se zadaje individualno i koji je konstruisan za procenu intelektualnih sposobnosti adolescenata i odraslih. WAIS-IV se sastoji od 15 suptestova (10 sržnih i 5 dopunskih) koji procenjuju kognitivno funkcionisanje u četiri domena – verbalno razumevanje (eng. verbal comprehension – VCI), perceptivno rezonovanje (eng. perceptual reasoning – PRI), radna memorija (eng. working memory – WMI) i brzina procesiranja (eng. processing speed – PSI), kao i opštu intelektualnu sposobnost (eng. Full-Scale IQ – FSIQ). WAIS-IV je zadat uzorku od 262 ispitanika: 104 ispitanika iz uzorka koji je reprezentativan za širu oblast Beograda, 62 ispitanika sa dijagnozom shizofrenije, 63 sa dijagnozom depresije i 33 ispitanika sa intelektualnim teškoćama. Psihometrijske karakteristike WAIS-IV suptestova su analizirane u okvirima teorije stavskog odgovora (eng. IRT) i klasične testne teorije (eng. CTT). Rezultati su pokazali da se IRT pouzdanost kreće od .86 (Razumevanje i Dopune) do .95 (Precrtavanje), dok je IRT pouzdanost FSIQ bila .94. Sveukupno, svi suptestovi dobro diskriminišu ispitanike duž celog kontinuuma intelektualnih sposobnosti. WAIS-IV omogućava visoko pouzdanu procenu intelektualnih sposobnosti na srpskoj populaciji.

**Ključne reči:** Četvrta revizija Vekslerove skale za procenu inteligencije kod odraslih, IRT, psihometrijske karakteristike

© 2018 authors



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution ShareAlike 4.0 International license