

External Validity of the Indonesian Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV-ID)

Christiany Suwartono
Faculty of Psychology
Atma Jaya Catholic University of Indonesia, Indonesia
Donders Institute for Brain, Cognition, and Behavior
Radboud University, Nijmegen, the Netherlands

Lidia L. Hidajat and Magdalena S. Halim
Faculty of Psychology
Atma Jaya Catholic University of Indonesia, Indonesia

Marc P. H. Hendriks
Donders Institute for Brain, Cognition, and Behavior
Radboud University, Nijmegen, the Netherlands
Department of Behavioral Sciences
Academic Centre of Epileptology
Kempenhaeghe, Heeze, the Netherlands

Roy P. C. Kessels
Donders Institute for Brain, Cognition, and Behavior
Radboud University, Nijmegen, the Netherlands
Department of Medical Psychology
Radboud University Medical Center, Nijmegen, the
Netherlands

In this research, we investigated the external validity of WAIS-IV-ID using other intelligence tests and educational achievement as criteria. We had 194 participants in total. The results showed the Full-Scale IQ score (FSIQ) of the WAIS-IV-ID had moderate yet significant correlation with three intelligence tests, namely Standard Progressive Matrices (SPM, $n = 194$), Cattell's Culture Fair Intelligence Test (CFIT, $n = 134$), and the Wechsler-Bellevue Intelligence Scale (WBIS, $n = 44$). There was also a significant positive correlation between the FSIQ and educational achievement score, the Grade Point Average (GPA, $n = 51$). The four indexes of the WAIS-IV-ID had a significant positive correlation with the SPM, CFIT, and WBIS, except for Verbal Comprehension Index (VCI). We found significant correlations between full-scale IQ with GPA. For the index score, we found significant correlations between Processing Speed Index (PSI) with GPA. We conclude that the WAIS-IV-ID is valid externally.

Keywords: WAIS-IV, Indonesia, academic achievement, SPM, CFIT

Penelitian ini menyelidiki validitas eksternal Wechsler Adult Intelligence Scale-Edisi Keempat (WAIS-IV-ID) menggunakan tes kecerdasan lainnya dan prestasi pendidikan sebagai kriteria. Partisipan berjumlah 194 peserta. Hasil menunjukkan skor skala penuh IQ (FSIQ) WAIS-IV-ID berkorelasi moderat dan signifikan dengan tiga tes kecerdasan, yaitu Standard Progressive Matrices (SPM, $n = 194$), Cattell's Culture Fair Intelligence Test (CFIT, $n = 134$), dan Wechsler Bellevue Intelligence Scale (WBIS, $n = 44$). Ditemukan juga korelasi positif yang signifikan antara FSIQ dan Indeks Prestasi Kumulatif (IPK, $n = 51$). Keempat indeks WAIS-IV-ID memiliki korelasi positif yang signifikan dengan SPM, CFIT, dan WBIS, kecuali Indeks Pemahaman Verbal (VCI). Kami menemukan korelasi yang signifikan antara skor skala penuh IQ dan IPK. Untuk indeks, kami hanya menemukan korelasi yang signifikan antara Indeks Pemrosesan Kecepatan (PSI) dan IPK. Simpulannya adalah bahwa WAIS-IV-ID sah secara eksternal.

Kata kunci: WAIS-IV, Indonesia, prestasi akademik, SPM, CFIT

This research was supported by research grants from Indonesian Ministry of Education and Culture under the International Research Collaboration and Scientific Publication in 2015 number 094/K3/KM/2015. The WAIS-IV (Copyright © 2008 NCS Pearson, Inc. Indonesian translation copyright © 2013. Translated, adapted, and reproduced with permission of the publisher. All rights reserved).

The authors acknowledge the contribution of individuals and institutions (particularly Universitas Pelita Harapan, Universitas Matana, and Bright Consulting) for their help with current data collection. Our gratitude goes to all testers and participants for their contribution in data collection of this research.

Correspondence concerning this article should addressed to Christiany Suwartono, Faculty of Psychology, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia. Email: christiany.suwartono@atmajaya.ac.id

In the history of psychological assessment, especially in the field of intelligence tests, the intelligence test results served as complementary information about developmental, social, educational, and occupational history that can be used as a comprehensive portrayal of a client. At the very least, the results would help to estimate premorbid levels of cognitive functioning, to formulate expectations of performance on other tests, and to determine the level of discourse at which to engage the client (Hiscock, 2007). The Indonesian language of Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV-ID) has proved to have structural validity that excellent and comparable with other internationally published standardized versions of the WAIS-IV (Suwartono, Halim, Hidajat, Hendriks, & Kessels, 2014). However, the structural validity cannot represent all aspects of a test's validity (Canivez, Konold, Collins, & Wilson, 2009). We should be convinced whether the WAIS-IV-ID is also valid externally. Therefore, in the present research, we validated the WAIS-IV-ID using external criteria such as other intelligence tests and Grade Point Average (GPA).

The Raven's Standard Progressive Matrices (SPM) and Cattell's Culture Fair Intelligence Test (CFIT) are among the major intelligence tests currently available in Indonesia. Both tests have less verbal instruction. Raven (2000) mentioned that the SPM test measures the deductive and reproductive ability of general cognitive ability. The deductive ability is the capacity to make meaning out of confusion, to generate high-level schemata that make it easy to handle complexity. The reproductive ability is the ability to absorb, recall, and reproduce information that has been made explicit and communicated from one person to another. The CFIT is assumed to be indifferent to cultural experiences that might differentially influence test taker's responses to its items and measures fluid intelligence. The fluid intelligence is a major measurable outcome of how biological factors influence intellectual development, and is assumed to be unaffected by cultural factors. Even though individual differences might exist within a culture, there are not necessarily any differences in fluid intelligence among cultures (Nenty & Dinero, 1981).

The predecessor version of the Wechsler intelligence test, the Wechsler-Bellevue Intelligence Scale (WBIS), was developed by David Wechsler in 1939. The WBIS is still widely used for intelligence testing in Indonesia. The WBIS is an individually administered measure of cognitive ability. The

WBIS consists of 11 subtests, namely Information (I), Comprehension (C), Digit Span (D), Arithmetic (A), Similarities (S), Vocabulary (V), Picture Arrangement (PA), Picture Completion (PC), Block Design (BD), Object Assembly (OA), and Digit Symbol (DSym). The WBIS divided into two parts, verbal and performance (LSP3 FPUI, n.d). It provides a measurement of general intellectual functioning or Full-Scale IQ (FSIQ), Verbal Scale (VS), and Performance Scale (PS). The Verbal scale includes six subtests (I, C, D, A, S, and V). The Performance scale includes five subtests (PA, PC, BD, OA, and DSym). Internationally, WBIS has been revised several times to WAIS, WAIS-R, WAIS-III, and the latest published WAIS-IV (Wechsler, 2008a, 2008b). The WAIS-IV consists of 15 subtests, namely Block Design (BD), Similarity (SI), Digit Span (DS), Matrix Reasoning (MR), Vocabulary (VC), Arithmetic (AR), Symbol Search (SS), Visual Puzzle (VP), Information (IN), Coding (CD), Letter Number Sequencing (LN), Figure Weights (FW), Comprehension (CO), Cancellation (CA), and Picture Completion (PC). The WAIS-IV subtests are identified as core and supplemental subtests. The first ten subtests are the core subtests, and the next five are the supplemental subtests (Wechsler, 2008). The WAIS-IV provides a measurement of general intellectual functioning (FSIQ) and four index scores. The four index scales include Verbal Comprehension (VCI), Perceptual Reasoning (PRI), Working Memory (WMI), and Processing Speed (PSI). The index scales include core and supplemental subtests. The Verbal Comprehension scale comprises three core subtests (SI, VC, and IN) and one supplemental subtest (CO). The Perceptual Reasoning scale includes three core subtests (BD, MR, and VP) and two supplemental subtests (FW and PC). The Working Memory Scale comprises two core subtests (DS and AR) and one supplemental subtest (LN). The Processing Speed scale consists of two core subtests (SS and CD) and one supplemental subtest (CA). We adapted the WAIS-IV into the Indonesian language. We re-arrange the items sequence in each subtest (except for SS, CD, CA because they are speed test) based on index difficulty. The items on the WAIS-IV-ID subtests are identical or equivalent to those of the WAIS-IV-US (Suwartono, Halim, Hidajat, Hendriks, & Kessels, 2014; Wechsler, 2008a). Therefore, in this study, we also investigated about the correlation and the differences between the scores that produced by these two Wechsler's scale.

For external criteria, it is important to know whether information about student's intellectual profile could predict academic achievement (Naglieri & Bornstein, 2003; Parker & Benedict, 2002; Rohde & Thompson, 2007; Watkins, Lei, & Canivez, 2007). From a theoretical perspective, the construct of intelligence is expected to influence the development of academic achievement because learning itself is *g*-demanding (Jensen, 1998). Therefore, we also conducted a correlation study between intelligence (WAIS-IV-ID) and a grade point average (GPA). We hypothesized that there was a significant positive correlation between the WAIS-IV-ID and other intelligence tests results and GPA. Moreover, we assumed that intelligence tests could contribute to the prediction of student success in university studies and thus also serve as the predictive validity of WAIS-IV-ID.

Method

Participants

We used a convenience sampling method. The criteria of participants adapted to the requirements listed in the WAIS-IV Technical and Interpretive Manual (Wechsler, 2008b). The present study involved 194 individuals in total. Of the participants, 66% women and 34% men. The age range from 16 – 61 years old ($M = 23.53$, $SD = 7.75$). Most of them are university students (50.5%), high school students (25.8%), employees (10.8%), consultants (6.2%), housewife (2.1%), lecturers (1.5%), and others (3.1%).

This study is part of a larger study for the evaluation of the psychometric properties of the WAIS-IV-ID. The area that we collect for the present study is from Jakarta (84.5%), Tangerang (6.2%), Bekasi (.5%), and Denpasar (8.8%).

The participants did different tests of intelligence, but all participants did the WAIS-IV-ID. Table 1 presents the demographic characteristics of participants that did both the WAIS-IV-ID and SPM ($N = 194$), the WAIS-IV-ID and CFIT ($N = 134$), the WAIS-IV-ID and WBIS ($N = 44$), and the WAIS-IV-ID and GPA ($N = 51$) samples.

Instruments

We used two measurements of the Wechsler's scale; the Wechsler-Bellevue Intelligence Scale (WBIS;

LSP3 FPUI, n.d) and The Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV-ID; Suwartono et. al., 2014; Wechsler, 2008a). Both of the scales provide information of a general intellectual functioning known as Full-Scale Intelligence Quotient (FSIQ). The WBIS is a predecessor of the WAIS-IV. The WBIS consists of 11 subtests divided into two parts, verbal (WB_VCI) and performance (WB_POI). It took approximately 90 – 100 minutes to finish the WBIS. The WAIS-IV-ID consists of 15 subtests divided into four factor: Verbal Comprehension (VCI), Perceptual Reasoning (PRI), Working Memory (WMI), and Processing Speed (PSI). Usually, the participants took 100 – 150 minutes to do the WAIS-IV. The items on the WAIS-IV-ID subtests are identical or equivalent to those of the WAIS-IV-US. More information of the WAIS-IV can accessed at <http://www.pearsonclinical.com/psychology/products/100000392/wechsler-adult-intelligence-scalefourth-edition-wais-iv.html#tab-details>.

We also used two measurements of intelligence that has less verbal instruction; the Raven's Standard Progressive Matrices (SPM; Raven, 2000, 2008) and the Cattell's Culture Fair Intelligence Test (CFIT; Cattell & Cattell, 1959, 1973; LSP3 FPUI, 2009). SPM consists of 60 items presented in five sets of 12. The test is untimed, but usually, the participant finishes before 25 minutes. It is relatively language free (Raven, 2000). The Cronbach's alpha coefficient is .84 (Suwartono, Amiseso, & Handoyo, 2016). We also used CFIT form 3A, which is designed to be a relatively true indicator of fluid intelligence. The CFIT is highly speeded, takes about 30 minutes to administer, and requires detailed verbal instructions for administration (Colom & Abad, 2007; LSP3 FPUI, 2009). The items of CFIT are entirely non-verbal and consist of four parts: Series, Analogies, Matrices, and Classification (Nenty & Dinero, 1981). The Cronbach's alpha coefficient is .79 (LSP3 FPUI, 2009).

For the educational achievement, we use the Grade Point Average (GPA). The GPA is the grading system employed in the university, ranged from zero to four. The GPA data was obtained only from Psychology students at a private university.

Procedure

This study is part of the larger study. We did cooperation with local offices, consulting firms, foundation, universities, and high schools. We ga-

Table 1
Demographic Characteristics of the Participants

	SPM	CFIT	WBIS	GPA
Sample size	194	134	44	51
Demographic data				
Men (%)	34.02	32.09	20.45	11.76
Women (%)	65.98	67.91	79.55	88.24
Age (years old)				
Age range	16 – 61	17 – 61	19 – 29	18 – 26
<i>M</i>	23.53	23.92	20.33	20.59
<i>SD</i>	7.75	6.89	2.87	1.72
Completed education (%)				
Junior high school	26.8	0.75	-	-
Senior high school	42.78	58.21	86.36	100
Undergraduate	25.77	34.33	11.36	
Master programme	4.64	6.72	2.27	
Ethnicity (%)				
Balinese	7.73	1.49	9.09	1.96
Bataknese	5.15	6.72	4.55	1.96
Javanese	17.53	22.39	15.91	15.69
Tionghoa	33.51	47.01	54.55	56.86
Others*	36.08	22.39	15.91	23.53

Note. SPM = The Raven's Standard Progressive Matrices, CFIT = The Cattell's Culture Fair Intelligence Test, WBIS = The Wechsler-Bellevue Intelligence Scale, GPA = Grade Point Average.

*Others = another ethnicity, like Sunda, Dayak, Minahasa, and many others; and also unanswered ethnicity by the participant.

thered the potential participants in a meeting room and explained that we did test adaptation and validation study. We introduced ourselves and gave an overview of our research. Then, we explained that we need participants to make the test adaptation (WAIS-IV-ID) and hope they willing to spend their time again to do another three-intelligence test (SPM, CFIT, and WBIS). We told them approximately the time required to do each test. After we informed about the time needed to finish each test, not all of them are willing to participate again due to schedule or they just do not want to participate again. If they agree to do other than the WAIS-IV-ID, we made a counterbalancing method to eliminate the potential of test sequence bias. If they want to participate again, we will contact again in two weeks' time to have another intelligence test. So, approximately one participant should allocate their time for two until three sessions; each session is about two hours. We administered the WAIS-IV-ID to all participants.

The WBIS and WAIS-IV-ID are an individually administered intelligence test, but the SPM and CFIT administered in a group setting in the classroom or meeting room at the university. The WBIS and WAIS-IV-ID administered according to the administration rules indicated in the manual (LPSP3 FPUI, n.d.; Wechsler, 2008a). We had separated group that did the SPM, CFIT, and WBIS. Then, to ensure the corresponding type of test sequence, some of them did the WAIS-IV-ID first, and others did the other intelligence test first (SPM or CFIT or WBIS). The administration of the WAIS-IV-ID and other intelligence tests (SPM, CFIT, and WBIS) collected within three months.

Analyses

We performed Levene's test for equality of variances to know whether there is the effect of test sequence received between the two groups: the one who received the WAIS-IV-ID first versus the

Table 2
Descriptive Statistics

	FSIQ_WAIS-IV-ID	SPM	CFIT	FSIQ_WBIS	GPA
Sample size	194	194	134	44	51
<i>M</i>	93.86	50.22	109.50	112.30	3.30
<i>SD</i>	n/a	n/a	n/a	n/a	.32
Minimum	40	36	73	95	2.6
Maximum	133	65	140	129	3.88

Note. n/a = we cannot display the *SD* due to licensing regulation.

group who received other intelligence tests (SPM and CFIT or WBIS) first. Then, we used t-value from t-test independent sample to check the sequence effect for the WAIS-IV-ID of test allocation. We checked whether there is sequence effect between those who did the WAIS-IV first versus those who did SPM and CFIT first (ABBA counterbalancing method). It is an answer whether the counterbalancing method worked or not. Then we present the descriptive statistics of the WAIS-IV-ID and all criteria for external validation study: SPM, CFIT, WBIS, and GPA.

For external validation analysis, we used Pearson product-moment correlation. We correlated the FSIQ, indices, and subtest scores of the WAIS-IV with the total scores on SPM, CFIT, WBIS, and GPA. Then we calculate the coefficient of determination (r^2) to investigate the importance of relationships and how each variable was affected by the other. Since the WBIS is the predecessor of the WAIS-IV-ID, we wanted to know whether this test result yields the same results. Then, we calculated t-test dependent sample comparing the FSIQ between the WBIS and WAIS-IV-ID. *Cohen's d* was calculated based on Lenhard and Lenhard (2016).

For the relationship between intelligence and academic achievement, we did Pearson product-moment correlation. Then, we did two kinds of regression towards the GPA. The first one was a simple linear regression. We estimated the GPA from the FSIQ of WAIS-IV. The second one was multiple linear regression. We estimated the GPA from VCI, PRI, WMI, and PSI of WAIS-IV. We used stepwise method with the following criteria: Probability-of-F-to-enter $\leq .050$, Probability-of-F-to-remove $\geq .100$. In the regression analysis, we checked for the *adjusted R²* and formulated the regression equation for transforming the standardized subtest scores into an estimate of a GPA score (GPA_{Est}). The *adjusted R²* is a modified version of R^2 that has been adjusted for the predictors in the model and increases only if

the new term improves the model more than would be expected by chance; served as a measure of goodness of fit for our prediction model (Field, 2013).

Results

We found the Levene's test for equality of variances $F = 3.03$, $p = .09$ between a group that did the WAIS-IV first then other intelligence tests (SPM and CFIT). The t-test independent samples result is $t_{(58)} = -1.40$, $p = .17$. We also found the Levene's test for equality of variances $F = 2.92$, $p = .09$ between a group that did the WAIS-IV first then WBIS. The t-test independent samples result is $t_{(50)} = -.02$, $p = .99$. These indicate an equal variance between groups and ABBA counterbalancing method is successful. For detail information, we present the descriptive statistics of the WAIS-IV-ID and all criteria for an external validation study in Table 2.

For the external validation results, we present the result of correlation analysis and coefficient of determination between WAIS-IV-ID scores (FSIQ, four indices, and 15 subtest scores) and external validity criteria (other measures of intelligence and GPA). Weak but significant correlations were found between SPM and WAIS-IV-ID scores (.19 to .32), whereas significant weak-to-moderate correlations found between CFIT and WAIS-IV-ID scores (.19 to .54). Weak to moderate and significant correlations were found between WBIS and the FSIQ, PRI, WMI, IN, BD, MR, FW, DS, and AR (.30 until .53). Furthermore, we found weak to moderate and significant correlations between GPA and FSIQ, PSI, MR, SS, and CD (.28 until .48).

In Table 3, we present the result of correlation analysis and coefficient of determination between WAIS-IV-ID scores (FSIQ, four index, and 15 subtest scores) and external validity criteria (other measures of intelligence and GPA). Weak but sig-

Table 3
Correlation With Other Tests and Academic Achievement

WAIS-IV-ID	SPM (<i>n</i> = 194)	<i>r</i> ²	CFIT (<i>n</i> = 134)	<i>r</i> ²	WBIS (<i>n</i> = 44)	<i>r</i> ²	GPA (<i>n</i> = 51)	<i>r</i> ²
FSIQ	.28**	.08	.54**	.29	.53**	.28	.38**	.14
VCI	.22**	.05	.34**	.12	.25	.06	.14	.02
SI	.19**	.04	.31**	.10	.14	.02	-.03	.00
VC	.20**	.04	.27**	.07	.04	.00	.20	.04
IN	.24**	.06	.19*	.04	.39**	.15	.13	.02
CO	.24**	.06	.28**	.08	-.02	.00	.08	.01
PRI	.28**	.08	.33**	.11	.45**	0.2	.18	.03
BD	.29**	.08	.26**	.07	.30*	.09	.14	.02
MR	.21**	.04	.30**	.09	.34*	.12	.28*	.08
VP	.28**	.08	.23**	.05	.27	.07	-.03	.00
FW	.32**	.10	.43**	.18	.41**	.17	.16	.03
PC	.26**	.07	.31**	.10	.19	.04	-.10	.01
WMI	.29**	.08	.47**	.22	.49**	.24	.18	.03
DS	.25**	.06	.32**	.10	.40**	.16	.10	.01
AR	.30**	.09	.48**	.23	.49**	.24	.20	.04
LN	.22**	.05	.24**	.06	.22	.05	.01	.00
PSI	.24**	.06	.40**	.16	.15	.02	.48**	.23
SS	.22**	.05	.37**	.14	.12	.01	.42**	.18
CD	.23**	.05	.31**	.10	.16	.03	.38**	.14
CA	.27**	.07	.36**	.13	.17	.03	.15	.02

Note. **. Correlation is significant at the .01 level (2-tailed); *. Correlation is significant at the .05 level (2-tailed).

nificant correlations were found between SPM and WAIS-IV-ID scores (.19 to .32), whereas significant weak-to-moderate correlations found between CFIT and WAIS-IV-ID scores (.19 to .54). Moderate and significant correlations found between WBIS and the FSIQ, PRI, WMI, IN, BD, MR, FW, DS, and AR (.30 until .53). Furthermore, we found moderate and significant correlations between GPA and FSIQ, PSI, MR, SS, and CD (.28 until .48).

The WAIS-IV-ID FSIQ index had significant relationship with SPM ($r_{(192)} = .28$, $r^2 = .08$, $p < .01$). For the indices, WMI had the highest magnitude ($r_{(192)} = .29$, $r^2 = .08$, $p < .01$). The highest magnitude and significant for each factor of WAIS-IV were IN and CO ($r_{(192)} = .24$, $r^2 = .06$, $p < .01$), FW ($r_{(192)} = .32$, $r^2 = .10$, $p < .01$), AR ($r_{(192)} = .30$, $r^2 = .09$, $p < .01$), and CA ($r_{(192)} = .27$, $r^2 = .07$, $p < .01$).

The WAIS-IV-ID FSIQ index had significant relationship with CFIT ($r_{(132)} = .54$, $r^2 = .29$, $p < .01$). As for the indices, WMI had the highest magnitude ($r_{(132)} = .47$, $r^2 = .22$, $p < .01$). The highest magnitude and significant for each factor of WAIS-IV were SI ($r_{(132)} = .31$, $r^2 = .10$, $p < .01$), FW ($r_{(132)} = .43$, $r^2 = .18$, $p < .01$), AR ($r_{(132)} = .48$, $r^2 = .23$, $p < .01$), and SS ($r_{(132)} = .37$, $r^2 = .14$, $p < .01$).

The WAIS-IV-ID FSIQ index had significant relationship with WBIS ($r_{(42)} = .53$, $r^2 = .28$, $p < .01$). As for the indices, WMI had the highest magnitude ($r_{(42)} = .49$, $r^2 = .24$, $p < .01$). The highest magnitude and significant for each factor of WAIS-IV were IN ($r_{(42)} = .39$, $r^2 = .15$, $p < .01$), FW ($r_{(42)} = .41$, $r^2 = .17$, $p < .01$), AR ($r_{(42)} = .49$, $r^2 = .24$, $p < .01$). There were no significant results between WBIS and subtests of processing speed. We present

Table 4
Correlation Between the WAIS-IV-ID and the Verbal Subtests of the WBIS

	WB_FSIQ	WB_VCI	WB_I	WB_CO	WB_D	WB_A	WB_S	WB_V
FSIQ								
WAIS4	.53**	.55**	.32	-.002	.53**	.60**	-.04	.31*
VCI	.25	.36*	.47**	-.06	.22	.36*	.06	.48**
SI	.14	.24	.14	-.003	.27	.28	-.08	.42**
VC	.04	.06	.24	-.06	-.05	.03	.08	.33*
IN	.39**	.51**	.70**	-.08	.25	.51**	.17	.34*
CO	-.02	.06	.21	.17	.03	.01	-.26	.25
PRI	.45**	.34*	.23	-.01	.23	.49**	-.04	.06
BD	.30*	.16	.17	-.07	.17	.32*	-.24	-.01
MR	.34*	.35*	.30*	-.09	.25	.39**	.14	.10
VP	.27	.16	-.01	.11	.02	.30	-.01	.04
FW	.41**	.34*	.14	.05	.41**	.22	.03	.28
PCm	.19	.03	-.21	.09	.27	-.23	.10	.04
WMI	.49**	.62**	.20	.15	.66**	.59**	-.05	.26
DS	.40**	.56**	.12	.26	.74**	.38**	-.10	.24
AR	.49**	.52**	.25	-.03	.40**	.71**	.02	.24
LN	.22	.38*	.14	.02	.50**	.30	.004	.24
PSI	.15	.12	-.01	-.10	.26	.08	-.01	.07
SS	.12	.10	-.03	-.08	.30	.01	-.04	.03
CD	.16	.11	.02	-.09	.16	.11	.03	.08
CA	.17	.17	.09	.09	.24	-.01	.03	-.05

Note. **. Correlation is significant at the .01 level (2-tailed); *. Correlation is significant at the .05 level (2-tailed).

the details about correlation between indexes and subtests of the WAIS-IV-ID and the verbal subtests of the WBIS in Table 4.

In Table 5, we present the details about the correlation between indexes and subtests of the WAIS-IV-ID and the performance subtests of the WBIS.

We also compared the WAIS-IV-ID and WBIS for each individual using the t-test for the dependent sample. There was a significant differences in the scores for the WAIS-IV-ID ($M = 103.68$) and the WBIS ($M = 112.30$) to measure their intelligence score; $t_{(43)} = 7.41$, $p < .01$, $d = 1.09$. The results of WBIS are higher than WAIS-IV. However, these results must be carefully interpreted, due to the major lack of updated information of the norm and items of WBIS. We will discuss further about the result in the discussion section.

The WAIS-IV FSIQ index had significant relationship with GPA ($r_{(49)} = .38$, $r^2 = .14$, $p < .01$). Among all the indices, only PSI that had significant correlation with GPA ($r_{(49)} = .48$, $r^2 = .23$, $p < .01$).

For subtests, the GPA showed significant correlation with MR ($r_{(49)} = .28$, $r^2 = .08$, $p < .05$), SS ($r_{(49)} = .42$, $r^2 = .18$, $p < .01$), and CD ($r_{(49)} = .38$, $r^2 = .14$, $p < .01$). The details of the relationship of the WAIS-IV-ID with GPA can be seen at Table 3.

Table 6 presents the correlations between GPA and several intelligence tests. GPA is not correlating significantly with SPM and CFIT. However, GPA correlates significantly with FSIQ and PSI of the WAIS-IV-ID. Then, both score can be further analysed as criteria for predictive validation.

Regarding predictive validity, our analysis showed that the FSIQ of the WAIS-IV-ID can predict the GPA ($Adjusted R^2 = .13$, $F_{(1,49)} = 8.22$, $p < .01$). The formula is $GPA_{Est} = 1.71 + .02$ FSIQ of WAIS-IV. Then, we estimated the GPA from the four indexes: VCI, PRI, WMI, and PSI of WAIS-IV as predictors with the stepwise method; we got $Adjusted R^2 = .21$, $F_{(1,49)} = 14.31$, $p < .01$. The formula is $GPA_{Est} = 2.06 + .01$ PSI. However, PSI is the only significant predictor of GPA.

Table 5
Correlation Between the WAIS-IV-ID and the Performance Subtests of the WBIS

	FSIQ_WB	WB_POI	WB_PA	WB_PC	WB_B	WB_OA	WB_Dsym
FSIQ_WAIS4	.53**	.20	.05	.18	.34*	-.09	.22
VCI	.25	-.02	-.07	.22	-.10	-.05	.06
SI	.14	-.02	-.09	.15	-.17	.01	.05
VC	.04	-.01	-.04	.20	-.11	.06	.01
IN	.39**	.01	-.01	.15	.05	-.19	.09
CO	-.02	-.13	-.26	.19	-.03	-.05	-.12
PRI	.45**	.36*	.09	.28	.52**	.004	.10
BD	.30*	.33*	.01	.06	.62**	.16	.04
MR	.34*	.17	.09	.26	.10	-.16	.18
VP	.27	.26	.09	.25	.38*	-.01	-.001
FW	.41**	.25	.08	.14	.36*	-.09	.23
PCm	.19	.29	-.15	.22	.36*	.29	.17
WMI	.49**	.07	.04	.05	.29	-.31*	.09
DS	.40**	-.01	-.12	-.02	.27	-.23	.06
AR	.49**	.16	.24	.14	.23	-.33*	.11
LN	.22	-.07	-.19	-.002	.24	-.32*	.001
PSI	.15	.06	.04	-.09	.07	.16	.34*
SS	.12	.02	-.13	-.04	.23	.15	.10
CD	.16	.09	.20	-.12	-.10	.13	.49**
CA	.17	.08	-.18	.01	.06	.25	.25

Note. **. Correlation is significant at the .01 level (2-tailed); *. Correlation is significant at the .05 level (2-tailed).

Table 6
Correlations Between the GPA and the Intelligence Tests

	SPM	CFIT	FSIQ_WAIS-IV-ID	VCI	PRI	WMI	PSI
	(n = 37)	(n = 35)	(n = 51)	(n = 51)	(n = 51)	(n = 51)	(n = 51)
GPA	-.06	.15	.38**	.14	.18	.18	.48**

Note. **. Correlation is significant at the .01 level (2-tailed); *. Correlation is significant at the .05 level (2-tailed).

Discussion

The current research provides further evidence for the Indonesian Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV-ID) validity. The Full-Scale Intelligence Quotient (FSIQ), indexes, and all subtests of WAIS-IV-ID had positive and significant correlations with the Raven's Standard Progressive Matrices (SPM) and the Cattell's Culture Fair Intelligence Test (CFIT). Moderate correlations found between the FSIQ of the WAIS-IV-ID and other measures of intelligence. Those results were expected because the WAIS-IV-ID covers four areas: verbal, perceptual reasoning, working memory,

and processing speed. However, the SPM and CFIT only cover the non-verbal (reasoning) area.

We found the Verbal Comprehension (VC) of the WAIS-IV-ID has a moderate and significant relationship with CFIT but the weak and significant relationship with SPM. Such results may stem from the fact that SPM and CFIT put less emphasize on verbal content, as both tests heavily measure fluid intelligence (Nenty & Dinero, 1981; Raven, 2000). Other than the fluid intelligence which is measured by the WAIS-IV-ID as PR factor, WAIS-IV-ID also measures VC, Working Memory (WM), and Processing Speed (PS). The interesting result, we found a weak and not significant relationship between VC

of the WAIS-IV-ID and the FSIQ of Wechsler-Bellevue Intelligence Scale (WBIS). From our experience in the administration and scoring process, the items on the WBIS are outdated and the time limit is very convenient. Participants got the zero score not because of they did not have the ability to answer the questions, but the content of the questions is no longer applied, as the questions about words that no longer used in every day in educational setting or everyday communication nowadays.

We found significant and positive correlations between WM and its subtests with SPM and CFIT. This result can be explained by Ackerman, Beier, Boyle (2002) and Kanerva & Kalakoski (2016) findings; the working memory had shared variance with general reasoning. Moreover, Kyllnonen and Christal (1990) and Tourva, Spanoudis, & Demetriou (2016) found that WM performance is positively and significantly related to tasks of reasoning or fluid intelligence.

The SPM and CFIT also had significant and positive correlations with PS and its subtests. Our results support Jensen (1980), Tillman, Bohlin, Sorensen, & Lundervold (2009), and Vernon (1983) findings. Jensen (1980) found that speed or efficiency of neural transmission in the brain affects performance on elementary cognitive tasks as well as reasoning task. Therefore, PS contributed to the intelligence performance (Tillman et al., 2009) and considered as the bridge between working memory and general cognitive ability (Vernon & Jensen, 1984).

The FSIQ of WAIS-IV-ID had significant and positive correlation with the Wechsler-Bellevue Intelligence Scale (WBIS). The moderate strength of correlation (shared variance was 28%) between the WAIS-IV-ID and the WBIS indicated that the shared part might be explained by the reasoning aspects measured in both tests. The interesting result, we found a weak and not significant relationship between PS of the WAIS-IV-ID and the FSIQ of Wechsler-Bellevue Intelligence Scale (WBIS). Overall performance from participants when they did WBIS, most of them answered correctly in the performance subtests in time. Unfortunately, WBIS still used as the main intelligence test among other psychological instruments in Indonesia. From our research results, we suggest that continued usage of WBIS in Indonesia is not recommended. From the data collection experience, we observed that the time limit of performance subtests from WBIS was too lenient as most of the participants have an-

swered correctly within the time limit. As an example in Digit Symbol (DSym; subtest 10) of the WBIS, Time limit is 90 seconds, max score = 67. From our data: the time limit range 58 – 90 s ($M = 84.68$, $SD = 7.12$). The score's range 34 – 67 ($M = 63.66$, $SD = 6.18$). Most of participants (52.3%) got the maximal score.

We found significant mean differences of FSIQ from WBIS and WAIS-IV-ID; WBIS ($M = 112.30$) yielded higher scores than WAIS-IV-ID ($M = 103.68$). This condition is similar to previous research (Hiscock, 2007) where it was found that WAIS yielded higher scores than the new test, WAIS-R. Our result suggests that there would be significant differences if someone used the WBIS for intelligence score; the individual would get a higher result if the IQ were derived from the WBIS. However, the major concern is the WBIS norm (LPSP3, n.d). There is no date that the table norm is produced. Therefore, it might be the WBIS score is higher because there were no updated norms. As the time flies, some changes might have explained a favorable effect on intelligence: people strive for better condition and rising standards of living like from working-class to middle-class homes. This condition also leads to a better environment, smaller family size, improved health, better nutrition, and improvement in the education (Hiscock, 2007; Nisbett, Aronson, Blair, Dickens, Flynn, Halpern, & Turkheimer, 2012; Rindermann, Becker, & Coyle, 2017; Williams, 1998). In Indonesia, the improvement in education was described in increased percentage of completed junior high school education rate (from 14.51% to 19%) and decreased illiteracy rate for adolescents (from 14.84% to 10.21%) from 1994 (Badan Pusat Statistik Indonesia, 2016). Therefore, WBIS could estimate a higher score than WAIS-IV-ID. However, current research is the first time we tested individuals with the WAIS-IV-ID. We need further research and must wait for some period to have that conclusion.

Regarding whether the WAIS-IV-ID could contribute to the prediction of student success in university studies, Naglieri and Bornstein (2003) found that cognitive test that measures basic psychological processes has considerable validity for prediction of achievement. Present findings established the validity of the WAIS-IV-ID as an intelligence test, which can be seen by its significant correlation with educational achievement represented by GPA, as Pluck, Ruales-Chieruzzi, Paucar-Guerra, Andrade-Guimaraes, and Trueba (2016) found. The FSIQ,

PSI, MR, SS, and CD had positive and significant correlations with GPA. These results also aligned with Parker and Benedict (2002) findings that FSIQ is predictive of IQ-achievement correlation. We found FSIQ influenced 14% to GPA. Moreover, the construct of intelligence is expected to precede and influence the development of academic achievement since learning itself is *g*-demanding (Jensen, 1998). As for the WAIS-IV-ID indices, only PSI had a positive and significant correlation with GPA. PSI reflects the resources of mental and motor speed to solve nonverbal problems. PSI also requires a person to be able to plan, organize, and develop relevant strategies. This is also reflected in its subtests, as MR measures nonverbal abstract reasoning abilities and visual information processing; SS and CD require capacity to absorb information as well as integrate and respond to this info, eye-hand coordination, attention, and capacity to work under pressure (Groth-Marnat, 2009). Those abilities were highly needed to survive in the university, which may explain the significant positive correlation between PSI and GPA. PSI could also be applied when the students need to prioritize and develop strategies between class, assignments, exams, and even the extracurricular activities. Therefore, a high PSI would help students to achieve a satisfactory grade at the university (GPA).

The strength of correlations between WAIS-IV-ID and GPA was weak to moderate. This is because the WAIS-IV-ID scales measure only a limited range of abilities compare to GPA; in fact, no battery can ever give a complete picture. Tests in a battery only assess the specific area of functioning that intended to be measured. An IQ is an estimate of a person's current level of functioning as measured by the various tasks required in a test; there are factors like motivation, persistence, personal adjustment, and family support which may contribute to the result (Groth-Marnat, 2009). More factors which can influence a person's performance in intelligence tests are beyond the scope of this article. Similarly, many factors other than IQ could contribute to the GPA. For instance, students are often given tasks to be solved by groups. His/her success in completing the task does not always come from their cognitive abilities. Additional capabilities such as leadership, delegation of tasks, and work together in teams are also influential.

Validity is not determined by a single evidence of validation study, but by a body of research that demonstrates the relationship between the test and the

behavior it is intended to measure. Brown (2010) described that validity is a unitary factor known as construct validity that consists of five sources of evidence. That evidence includes test content, response processes, internal structure, relations to other variable, and consequences of testing. Future validation research can examine the diagnostic utility of WAIS-IV with special (clinical) groups (Pintea & Moldovan, 2009). Clinical groups examined may include individual identified as intellectually gifted, as well as intellectually disabled, mild cognitive impairment and alzheimers. Group differences were desirable as one way to know whether the WAIS-IV could be useful as a diagnostic test.

The present research has some limitations. Even though we designed to use quota sampling, due to time and situation when we did the data collection, we used convenience sampling. Therefore, this study has covered limited areas and limited participants from Java and Bali islands. Further research is necessary to assess the generalization of the findings to people from other geographical regions in Indonesia.

At the same time, this research involved limited size participants. This external validation research is subject to the availability of the participant's time and their willingness to participate in the different occasion. We always offered all participants to have another test but not all of them agreed due to the schedule, or they did not want to participate again. If they are willing to participate again, we set an appointment approximately in two weeks' time to have another intelligence tests. So, approximately one participant should allocate their time for three meetings; each meeting lasted approximately two hours. The time and sometimes place restriction of current data collection were the main challenges that we could not control. Thus, future research should verify whether investigations with larger participants and more diverse in age range would reveal comparable results.

Addressing the concern about the Flynn effect, it is not applicable in our results. The Flynn effect is a theory which emphasizes the fact that average intelligence quotient (IQ) scores have risen over generations (Hiscock, 2007). The younger generations scored better than the older generation. Our result showed that the participants score higher in WBIS compare to WAIS-IV now. Unfortunately, we did not have their past scores with either of those scales. So, we cannot investigate about Flynn effect. They scored higher in old test (WBIS) due to

their speed is increased, and they have more access to a source of information, so they scored better in WBIS. Moreover, we have a limited age range, most of the participants are young generation.

With consideration of those limitations and concerns, this present research provides preliminary evidence that the WAIS-IV-ID is valid as it has a positive and significant correlation with other intelligence tests. Moreover, the WAIS-IV-ID could predict future achievement in university (GPA). Therefore, the WAIS-IV-ID is valid regarding external criteria such as other non-verbal intelligence tests and academic achievement.

References

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, *131*, 567–589.
- Badan Pusat Statistik Indonesia. (2016). *Indikator Pendidikan, 1994-2015* [Education Indicator 1994 - 2015]. Retrieved <https://www.bps.go.id/linkTabelStatistik/view/id/1525>
- Brown, T. (2010). Construct validity: A unitary concept for occupational therapy assessment and measurement. *Hong Kong Journal of Occupational Therapy*, *20*(1), 30–42.
- Canivez, G. L., Konold, T. R., Collins, J. M., & Wilson, G. (2009). Construct validity of the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test: Convergent and structural validity. *School Psychology Quarterly*, *24*, 252–265. <http://dx.doi.org/10.1037/a0018030>
- Cattell, R. B., & Cattell, A. K. S. (1959). *Culture fair intelligence test (Scale 3)*. Champaign, IL: Institute for Personality and Ability Testing, Inc.
- Cattell, R. B., & Cattell, A. K. S. (1973). *Measuring intelligence with the culture-fair tests: Manual for Scales 2 and 3*. Champaign, IL: Institute for Personality and Ability Testing, Inc.
- Colom, R., & Abad, F. (2007). Culture Fair Intelligence Test. In Neil J. Salkind, & K. Rasmussen (Eds.), *Encyclopedia of Measurement and Statistics*. (pp. 205-206). Thousand Oaks, CA: Sage Publications, Inc. <http://dx.doi.org/10.4135/9781412952644.n116>
- Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics*. London: Sage Publication.
- Groth-Marnat, G. (2009). *Handbook of psychological assessment*. New Jersey: John Wiley & Sons.
- Hiscock, M. (2007). The Flynn effect and its relevance to neuropsychology. *Journal of clinical and experimental neuropsychology*, *29*(5), 514-529.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: The Free Press.
- Kanerva, K., & Kalakoski, V. (2016). The Predictive Utility of a Working Memory Span Task Depends on Processing Demand and the Cognitive Task. *Applied Cognitive Psychology*, *30*, 681–690.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*, 389–433.
- Lenhard, W., & Lenhard, A. (2016). Calculation of Effect Sizes. Retrieved from http://www.psychometrica.de/effect_size.html. Bibergau (Germany): Psychometrica.
- Lembaga Pengembangan Sarana Pengukuran dan Pendidikan Psikologi Fakultas Psikologi Universitas Indonesia (LSP3 FPUI). (n.d). *Manual Wechsler-Bellevue Intelligence Scale for Adolescents and Adults Form I*. Depok: LSP3 FPUI.
- Lembaga Pengembangan Sarana Pengukuran dan Pendidikan Psikologi Fakultas Psikologi Universitas Indonesia (LSP3 FPUI). (n.d). *Tabel IQ WBIS*. Depok: LSP3 FPUI.
- Lembaga Pengembangan Sarana Pengukuran dan Pendidikan Psikologi Fakultas Psikologi Universitas Indonesia (LSP3 FPUI). (2009). *Buku Petunjuk Penggunaan: CFIT Culture Fair Intelligence Test Skala 3 Bentuk A dan Bentuk B*. Depok: LSP3 FPUI.
- Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment*, *21*, 244-260. <http://dx.doi.org/10.1177/073428290302100302>
- Nenty, H. J., & Dinero, T. E. (1981). A cross-cultural analysis of the fairness of the Cattell Culture Fair Intelligence Test using Rasch model. *Applied Psychological Measurement*, *5*(3), 355–368. <http://dx.doi.org/10.1177/014662168100500309>
- Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, *67*(2), 130-159.
- Parker, D. R., & Benedict, K. B. (2002). Assessment and intervention: Promoting successful transitions for college students with ADHD. *Assessment for Effective Intervention*, *27*(3), 3-24.

- Pintea, S., & Moldovan, R. (2009). The Receiver-Operating Characteristic (ROC) analysis: Fundamentals and applications in clinical psychology. *Journal of Cognitive and Behavioral Psychotherapies*, 9(1), 49-66.
- Pluck, G., Ruales-Chieruzzi, C. B., Paucar-Guerra, E. J., Andrade-Guimaraes, M. V., & Trueba, A. F. (2016). Separate contributions of general intelligence and right prefrontal neurocognitive functions to academic achievement at university level. *Trends in Neuroscience and Education*, 5, 178-185.
- Raven, J. (2000). The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cognitive Psychology*, 41, 1-48.
- Rindermann, H., Becker, D., & Coyle, T. R. (2017). Survey of expert opinion on intelligence: The Flynn effect and the future of intelligence. *Personality and Individual Differences*, 106, 242-247.
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, 35, 83-92.
- Suwartono, C., Amiseso, C. P., & Handoyo, R. T. (2016). Uji Psikometri Raven's Standard Progressive Matrices. *Manuscript in preparation*.
- Suwartono, C., Halim, M. S., Hidajat, L. L., Hendriks, M. P., & Kessels, R. P. (2014). Development and reliability of the Indonesian Wechsler adult intelligence scale-fourth edition (WAIS-IV-ID). *Psychology*, 5, 1611-1619.
- Tillman, C. M., Bohlin, G., Sorensen, L., & Lundervold, A. J. (2009). Intelligence and specific cognitive abilities in children. *Journal of Individual Differences*, 30(4), 209-219.
- Tourva, A., Spanoudis, G., & Demetriou, A. (2016). Cognitive correlates of developing intelligence: The contribution of working memory, processing speed and attention. *Intelligence*, 54, 136-146.
- Vernon, P. A., & Jensen, A. R. (1984). Individual and group differences in intelligence and speed of information processing. *Personality and Individual Differences*, 5(4), 411-423.
- Watkins, M. W., Lei, P. W., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence*, 35, 59-68.
- Wechsler, D. (1939). *Wechsler-Bellevue Intelligence Scale*. New York, NY: The Psychological Corporation.
- Wechsler, D. (2008a). *Wechsler Adult Intelligence Scale-Fourth Edition Administration and Scoring Manual*. San Antonio, TX: Pearson.
- Wechsler, D. (2008b). *Wechsler Adult Intelligence Scale-Fourth Edition Technical and Interpretive Manual*. San Antonio, TX: Pearson.
- Williams, W. M. (1998). Are we raising smarter children today? School- and home- related influences on IQ. In U. Nesser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 125-154). Washington, DC: American Psychological Association.

Glossary

SPM	=	The Raven's Standard Progressive Matrices
CFIT	=	The Cattell's Culture Fair Intelligence Test
GPA	=	Grade Point Average
WBIS	=	The Wechsler-Bellevue Intelligence Scale
I	=	Information
C	=	Comprehension
D	=	Digit Span
A	=	Arithmetic
S	=	Similarities
V	=	Vocabulary
PA	=	Picture Arrangement
PC	=	Picture Completion
BD	=	Block Design
OA	=	Object Assembly
DSym	=	Digit Symbol
POI	=	The Perceptual Organization Index
WB	=	The Wechsler-Bellevue Intelligence Scale

WAIS-IV	=	The Wechsler Adult Intelligence Scale-Fourth Edition
WAIS-IV-ID	=	The Indonesian version of Wechsler Adult Intelligence Scale-Fourth Edition
WAIS-IV-US	=	The American of version Wechsler Adult Intelligence Scale-Fourth Edition
BD	=	Block Design
SI	=	Similarity
DS	=	Digit Span
MR	=	Matrix Reasoning
VC	=	Vocabulary
AR	=	Arithmetic
SS	=	Symbol Search
VP	=	Visual Puzzle
IN	=	Information
CD	=	Coding
LN	=	Letter Number Sequencing
FW	=	Figure Weights
CO	=	Comprehension
CA	=	Cancellation
PC	=	Picture Completion
VCI	=	Verbal Comprehension Index
PRI	=	Perceptual Reasoning Index
WMI	=	Working Memory Index
PSI	=	Processing Speed Index
FSIQ	=	General intellectual functioning, Full-Scale Intelligence Quotient