Appraising Intellectual Giftedness Using the Malay Version of WISC-R

ABD. MAJID MOHD ISA and OTHMAN MOHAMED
Faculty of Educational Studies
Universiti Pertanian Malaysia
43400 UPM Serdang, Selangor, Malaysia

ABSTRAK

Kajian kesesuaian WISC-R yang diterjemahkan kepada Bahasa Melayu ini dijalankan ke atas 100 pelajar (53 lelaki dan 47 perempuan) yang berumur 11 tahun. Sebelum ianya dijalankan, proses penterjemahan WISC-R ini dilakukan secara sistematik iaitu menggunakan panel dan 'back translation'. Hasil daripada kajian ini mendapati indeks kebolehpercayaan WISC-R bahasa Melayu ialah .91 dan ini selaras dengan syarat yang ditetapkan untuk sesuatu ujian piawai individu. Skor bagi setiap sub-ujian juga tidak berbeza dengan data yang dikemukakan oleh WISC-R. Oleh itu, WISC-R yang diterjemah dalam Bahasa Melayu sesuai untuk digunakan untuk mengukur kecerdasan (IQ).

ABSTRACT

An appropriateness study of the Malay version of the WISC-R was conducted on one hundred 11 year old students (53 boys and 47 girls). Prior to the study, the WISC-R was systematically translated using the back translation method with a panel of experts as moderators. From the results of the study, the reliability index of the Malay version of the WISC-R is .91 and this coefficient is consistent with the condition set for an individual standardised test. The score for each sub-tests too does not differ from the data presented by the original WISC-R. Therefore, the Malay version of WISC-R is appropriate for use as a measure of intelligence (IQ).

INTRODUCTION

Under the present review of the Malaysian Education Act 1961 changes will be made to incorporate a special educational provision for intellectually gifted children. Various measures are used to identify gifted children. The most commonly used measure is intelligence tests. Intelligence tests may not be able to identify all mentally gifted. However, they are the most effective single selection instrument available (Hollingworth 1942; Martinson 1961; Reynold 1962). Karnes and Collins (1978) found that the Stanford Binet Intelligence Test and Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler 1974) are the most commonly used tests in identifying intellectually gifted children in the United States of America. Furthermore, the WISC-R was found to be the most popular and well-researched instrument used for the assessment of intellectual functioning in both clinical and academic settings (Rashed 1989).

The items in the WISC-R have been intensively investigated and found to be clinically and psychometrically sound. The WISC-R Verbal and Performance Scales were found to be correlated with Cattel's (1971) crystallised and fluid abilities (Kaufman 1979). Meeker (1975) indicated that the WISC-R was developed from Guilford's SOI (Structure of Intellect) model. Therefore, the WISC-R had a strong theoretical foundation in quantifying intelligence. Also, intelligence as measured by the WISC-R was found to be the best predictor of school achievement. Sattler (1974) found that the correlations between the Full Scale IQ of WISC-R and a wide variety of achievement measures averaged .61. For the Full Scale IQ of WISC-R, similar magnitudes, in terms of correlations, were reported for groups of white or

predominantly white children(Hale 1978), for groups of minority or primarily minority youngters (Harlage and Steele 1977), and for exceptional populations (Raskin *et al.* 1978). The WISC-R, which is the revised version of the 1949 WISC, is an individually administered intelligence test and has been the major instrument for assessing the intellectual functioning of school-age children (Anderson 1976; Karnes and Collins 1978). Also, the WISC-R has superior reliability; is culturally unbiased, and has up-to-date items (Sattler 1974). Moreover, its norming quality is well accepted because of its diverse representation (Sattler 1974).

The WISC-R reports 3 types of IQ: 1) the Verbal Scale IQ; 2) the Performance Scale IQ; and 3) the Full Scale IQ. The sub-tests for the Verbal Scale are Information (Info), Comprehension (Com), Arithmetic (Arit), Similarities (Sim) and Vocabulary (Voc). The Performance IQ, on the other hand, consists of Picture Completion (PC), Block Design (BD), Picture Arrangement (PA), Object Assembly (OA) and Coding (Cod). All the sub-tests (Verbal and Performance) are used to compute the Full Scale IQ. The administration of the WISC-R to each child requires approximately 1-1 $^1/_2$ hours.

The rationale for this study to test the applicability of WISC-R in Malaysian classrooms is because it has not been done before. Therefore, basic research and development for establishing the reliability and the validity of WISC-R in Malaysia was necessary. This study will justify its potential as a selection tool for identifying intellectually gifted children in Malaysia and also ascertain that the Malay version of the WISC-R is able to do so.

METHOD

The English version of WISC-R was initially translated into Malay and then given to a translation panel comprised of 5 local experts. They were two associate professors in Malay Studies, two lecturers in Educational Psychology and a teacher with 12 years' experience teaching Malay children in a rural area.

Item Analysis and Back Translation

The amendments suggested by the panel were made after receiving separate comments from each translation panel member. The WISC-R

was then administered to a class of 25 primary six pupils in a school in Selangor. The respondents were instructed not only to respond to every item but also to mark whether they did or did not understand each item. All pupils indicated that they understood the items. The final version of the WISC-R was given to 5 final year Bachelor of Education students (Teaching English as a Second Language) at Universiti Pertanian Malaysia. Each item was translated into English. This 'back translation procedure' was essential to ensure that the content of the final Malay version of the WISC-R had not deviated from the original English version. From the back translation, only one item had been mistranslated. Item 18 of the vocabulary sub-test was translated as 'lagenda' meaning 'heroic act' in place of the original concept of 'fable'. Subsequently, the word 'lagenda, was changed to 'dongeng', which means 'fairy tale or folk story'.

The subjects in the study were 100 primary six pupils comprising 53 boys and 47 girls from a single school in Rembau. This sample size was adequate to minimiz Type II error consistent with the t-test statistic set at a power level .8 and alpha at .05 (Cohen 1992). The average age of the subjects was 11 years 5.8 months (SD=2.8 months). Although the boys were slightly older (11.7 years) than to the girls (11.5 years), there was no significant difference in age. The WISC-R was administered in July, 1992 to an average of 5 pupils a day. After a lapse of 30 days, 30 pupils were randomly selected and administered another WISC-R.

RESULTS

The Malay version of the WISC-R total scores and its sub-tests, with the exception of Vocabulary, very close to a normal distribution (Table 1). The means of the sub-test scores were generally close to the 'Western' values, X=10 and SD=3 for every sub-test. The sub-test with lower values than the values published in the manual (Wechsler 1974) were Comprehension and Object Assembly. On the contrary, the mean for Block Design and Coding were comparatively higher. However, the differences between the mean of these four scores and the values published in the manual were not significant as determined by t-test statistics.

The mean and the standard deviation for the Verbal Scale IQ, Performance Scale IQ and Full Scale IQ of the Malay version of WISC-R

were similar to those in the manual, 50 and 10 for Verbal and Performance IQ and 100 and 15 for the Full Scale IQ. Although the mean for the Verbal IQ of the Malay version WISC-R was lower than the English version, the mean for the Performance and the Full Scale IQ was higher than the value reported in the manual. As in the case of its sub-tests, the differences were not significant as determined by t-test statistics.

The consistency of the Malay version of the WISC-R score, with the exception of Coding, was measured by Cronbach's alpha. The Cronbach's alpha values ranged from .65 to .91. As the number of items for each sub-test was between 4 and 30, no greater Cronbach alpha values could be expected. Based on these data, it could be concluded that the WISC-R score has modest internal consistency (Table 2).

The stability of the WISC-R score was measured by the test-retest procedure after a lapse of 30 days using 30 randomly selected subjects. The correlation between the first and the second administration was between .73 and .91. Although the scores of the second administration were higher than the initial administration, there were no significant differences indicated by the t-test result for each sub-test (Table 2).

It is the objective of test and inventory developers to produce an instrument that is free from gender bias. However, there are still cultures that encourage differences between feminine and masculine behaviour. In this respect, response differences between gender are to be

expected (Maccoby and Jacklin 1975).

Therefore, in appraising intellectual quotient, it is essential to study the gender differences. In this regard, different norms have to be proposed for the two genders, if the differences are significantly large.

The WISC-R scores (Table 3) indicate that it was only in the Arithmetic sub-test that the scores for boys were significantly higher than those for girls. This phenomenon is common, indicating boys are better than girls in mathematical skill and achievement (Maccoby and Jacklin 1975). In the Verbal, Performance and Full Scale IQ, the differences were not significant. Therefore, the IQ appraised by the Malay version of WISC-R has no gender bias.

The intercorrelation among the test or an inventory's sub-tests is a common procedure used to re-validate the test or an inventory (Sattler 1974). The intercorrelation coefficients of the Malay version WISC-R sub-tests were modest, ranging from .20 to .58 (Table 4). The correlation coefficients of Verbal sub-tests with Verbal IQ were generally higher than the correlation coefficients of Performance sub-tests with Performance IQ. These data confirm that each sub-test measured related constructs.

Also, the Verbal Scale IQ and Performance Scale IQ were strongly correlated with the Full Scale IQ, with values of .92 and .89 respectively. Based on these findings, it can be concluded that the Malay version of WISC-R is similar to the original version in determining IQ, thus showing validity.

TABLE 1
Descriptive statistics of WISC-R
(N=100)

Subtests	Mean	SD	Median	Mode	Kurtosis
Info	10.98	2.54	11	12	.63
Sim	9.49	2.98	10	10	-.15
Arit	10.87	3.01	11	13	-.13
Voc	9.27	3.18	10	10	1.48
Com	9.95	3.67	8	9	-.29
PC	10.31	3.37	10	10	-.05
PA	11.25	3.68	11	14	-.26
BD	11.01	2.97	12	13	.12
OA	9.56	4.03	9	11	-.46
Cod	10.18	3.21	10	10	.66
Verbal	49.13	11.31	52	54	.08
Performance	52.32	11.29	53	56	.02
Full IQ	103.12	14.92	104	101	-.28

TABLE 2
Reliability of the Malay version of WISC-R

| Sub-test | No of Items | First Administration | | | After 30 Days | | | | |
| | | Mean | SD | Alpha* | Mean | SD | r | SEM | t |
		(N=100)			(N=30)				
Info	30	10.98	2.54	.75	11.33	2.63	.81	1.68	-0.65
Sim	17	9.49	2.98	.66	9.71	2.51	.76	2.23	-0.41
Arit	18	10.87	3.01	.86	11.66	2.96	.86	1.52	-1.28
Voc	32	9.27	3.18	.89	10.09	2.88	.86	1.78	-1.33
Com	17	9.95	3.18	.84	10.50	2.05	.84	2.01	-1.13
PC	26	10.31	3.37	.87	10.72	2.66	.88	1.66	-0.71
PA	12	11.25	2.97	.65	12.11	2.87	.73	1.80	-1.43
BD	11	11.01	2.97	.78	11.56	3.00	.81	1.84	-0.89
OA	4	9.56	4.03	.74	10.75	3.66	.79	2.72	-1.52
Cod	93	10.18	3.21	NA	11.12	3.33	.83	1.79	-1.36
Verbal		49.13	11.3	.89	52.74	10.8	.90	5.17	-1.59
Performnce		52.31	11.3	.87**	57.06	12.0	.91	5.57	-1.63
Full IQ		103.12	14.9	.91**	106.77	13.9	.91	6.19	-1.24

*   Calculates based on Cronbach Alpha except in for Cod where it was calculated based on test-retest procedure.

**   Excluding Cod.

TABLE 3
Inter-correlation coefficient* of WISC-R
(N=100)

Sub-tests	1	2	3	4	5	6
Verbal:						
1. Info						
2. Sim	49					
3. Arit	20	30				
4. Voc	47	43	33			
5. Com	45	50	44	58		
6. Verbal	68	75	59	77	79	
7. Full IQ	64	73	50	73	72	71
Performance:						
1. PC						
2. PA	39					
3. BD	46	29				
4. OA	37	43	40			
5. Cod	21	30	25	31		
6. Performance	71	68	66	74	60	
7. Full IQ	57	59	65	57	92	89

All r are significant at p<.01
*decimal point is omitted

TABLE 4
Gender differences of WISC-R score
(Boys=53 Girls=47)

Subtests	Gender	Mean	Std. Dev.	t
Info	Boy	11.08	2.60	.40
	Girl	10.87	2.49	
Sim	Boy	9.41	3.42	-.27
	Girl	9.57	2.42	
Arit	Boy	12.55	2.71	2.18*
	Girl	11.26	3.16	
Voc	Boy	9.23	3.59	-.14
	Girl	9.32	2.68	
Com	Boy	7.75	3.59	-.15
	Girl	8.17	3.55	
PC	Boy	10.58	3.42	.87
	Girl	10.00	3.31	
PA	Boy	11.26	2.74	.02
	Girl	11.25	2.66	
BD	Boy	12.55	2.98	1.95
	Girl	11.40	2.87	
OA	Boy	8.39	4.24	-.43
	Girl	8.74	3.83	
Cod	Boy	12.75	3.35	-1.42
	Girl	13.66	3.02	
Verbal IQ	Boy	48.38	11.97	-.70
	Girl	49.98	10.58	
Performance IQ	Boy	55.17	12.55	-.14
	Girl	55.49	9.80	
Full IQ	Boy	102.66	16.73	-.33
	Girl	103.64	12.74	

*$p<0.05$

CONCLUSION

The study indicated similarities in the reliability and validity of the Malay version of the WISC-R with the data of the original WISC-R. Although there are variations in the findings, these values are expected because of the small, yet adequate sample size. Nevertheless, the National Foundation of Educational Research (NFER) (1977) suggests a reliability coefficient of .9 for an individual appraisal instrument. The Malay version of WISC-R in this study had a test-retest reliability and an internal consistency of .91 and therefore meets this criterion. On the basis of these data, the usage of the Malay version of WISC-R to assess general intellectual functioning among the Malaysian children is justified.

REFERENCES

ANDERSON, K.E. 1976. *Research on the Academically Talented Student*. Washington D.C.: National Education Association.

CATTEL, R.B. 1971. *Abilities: Their Structure, Growth, and Action*. Boston: Houghton-Mifflin.

COHEN, J. 1992. A power primer. *Psychological Bulletin* **112(1)**: 155-159.

HALE, R.L. 1978. The WISC-R as a predictor of wide range achievement test performance. *Psychology in the Schools* **15**: 172-175.

HARLAGE, L.C. and C.T. STEELE, 1977. WISC and WISC-R correlates of academic achievement. *Psychology in the Schools* **14**: 15-18.

HOLLINGWORTH, L.S. 1942. *Children Above 180 IQ*. Yonkers, N.Y.: World Book.

KARNES, F.A. and E.C. COLLINS. 1978. States definitions on the gifted and talented. *Journal for the Education of the Gifted* **22:** 44-62.

KAUFMAN, A.S. 1979. *Intelligent Testing with the WISC-R.* New York: Wiley.

MACCOBY, E.E. and C.N. JACKLIN. 1975. *The Psychology of Sex Differences.* Stanford, Calif.: Stanford University Press.

MARTINSON, R.A. 1961. *California State Study of Educational Programs for Gifted Children.* Saccremento: California State Department of Education.

MEEKER, M.N. 1975. *WISC-R Template for SOI.* California: SOI Institute.

NATIONAL FOUNDATION of EDUCATIONAL RESEARCH. 1977. *Catalogue of Tests for Educational Guidance and Assessment.* Windsor.

RASHED, A.G. 1989. Standardization of the WISC-R for Students aged 12-15 years in Iraq. Unpublished Ph.D. Thesis, University of Hull.

RASKIN, L.M., A.S BLOOM, S.H. KLEE and A. REESE. 1978. The Assessment of developmentally disabled children with WISC-R, Binet and other tests. *Journal of Clinical Psychology* **34:** 111-114.

REYNOLD, M.C. (ed) 1962. *Early School Admission for Mentally Advanced Children: A Review of Research and Practice.* Washington D.C.: Council for Exceptional Children.

SATTLER, J.M. 1974. *Assessment of Children's Intelligence.* Philadelphia: Saunders.

WECHSLER, D. 1974. *Manual for WISC-R.* New York: Psychological Corporation.