

**Two- and Four-Subtest Short Forms of the Wechsler
Adult Intelligence Scale-Revised:
A Comparative Validity Study with Normals**

Susan Plumridge ©

Lakehead University

Thesis Supervisor: Anthony P. Thompson, Ph.D. C. Psych.

**Thesis submitted as partial fulfilment of the
requirements for the Master of Arts Degree in Clinical
Psychology at Lakehead University, Thunder Bay, Ont.**

ProQuest Number: 10611414

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10611414

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-315-97062-6

Acknowledgements

I would like to thank my thesis advisor, Dr. Thompson for his expert guidance and continual support throughout the progress of this project. His professionalism, advice and patience was greatly appreciated. I would also like to thank Dr. Rawana for his helpful suggestions as the second reader for this thesis. In addition I would like to acknowledge the tireless contribution of Mirella Fata as a test administrator. Finally I would like to thank the students who volunteered to participate in this study.

Table of Contents

Acknowledgements	ii
Table of Contents	iii
List of Tables	iv
Abstract	v
1. Introduction	1
2. Method	9
Subjects	9
Procedure	10
Administrative Conditions	11
Scoring	12
3. Results	13
Silverstein's 2SF/4SF and Kaufman's 4KSF	15
Kaufman's 2KSF	19
4. Discussion	21
References	28
Appendix A Ethics Approval	36
Appendix B Information for Subjects	38
Appendix C Consent to Participate	40
Appendix D Questions	42

List of Tables

Table 1	Descriptive Statistics for IQ's by Administrative Condition.....	14
Table 2	Mean and Standard Deviation of WAIS-R Full Scale I.Q. and Estimated I.Q. from Short and Standard WAIS-R Administrations.....	16
Table 3	Correlation Between FSIQ and SFIQ for Each Condition.....	18
Table 4	Estimated IQ Minus Full Scale IQ by Administrative Conditions.....	20

Abstract

The concurrent validity of WAIS-R short forms was evaluated in a sample of normal subjects. The accuracy of short form IQ estimates from a standard administration was compared with the accuracy of an IQ estimate when the short form was administered first followed by the remaining subtests. Two-subtest short forms (Vocabulary/Block Design; Information/Picture Completion) and four-subtest short forms (Arithmetic/Similarities/Picture Completion/Digit Symbol; Vocabulary/Block Design/Arithmetic/Picture Arrangement) were examined. It was found that the performance of the four-subtest short forms was comparable in both administrative conditions. The Vocabulary/Block Design short form was a good estimate of full scale IQ when administered first and overestimated full scale IQ when rescored. The role of motivational factors in short form estimates is discussed. The Information/Picture Completion short form had low validity. The results support clinicians' continued caution in the use of short forms. More research utilizing the methodology of this study is recommended when evaluating short form validity.

Two- and Four-Subtest Short Forms of the Wechsler
Adult Intelligence Scale-Revised:
A Comparative Validity Study with Normals

A full administration of the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) takes approximately 90 minutes. For screening purposes or when a global estimate of IQ is sufficient, WAIS-R short forms can be administered to save time (Kaufman, 1990; Sattler, 1988). The WAIS-R and its predecessor the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) have commonly been shortened by eliminating some of the items but retaining all of the subtests (Cargnello, & Gurekas, 1987; Satz-Mogel, 1962) or by eliminating some of the subtests (Doppelt, 1956; Reynolds, Wilson, & Clark, 1983; Silverstein, 1967; 1982; Ward 1990).

There is some consensus in the literature that shortening the WAIS-R by eliminating subtests is more valid and reliable than short forms that eliminate items (Boone, 1991; Kaufman, 1990, Kaufman, Ishikuma, & Kaufman-Packer, 1991; Sattler, 1988; Silverstein, 1990a; 1990b). Short forms that eliminate items make

it practically impossible to subsequently administer the remaining subtests if deemed necessary (Silverstein, 1990b).

Silverstein (1982) developed short forms of the WAIS-R based on data from the standardization sample. He selected subtests that were commonly used in clinical settings and that correlated highly with the Full Scale IQ (FSIQ). His two-subtest combination consisted of Vocabulary and Block Design subtests (2SF) and his tetrad, originally selected by Doppelt (1956), consisted of the Vocabulary, Block Design, Arithmetic and Picture Arrangement subtests (4SF). Silverstein provided tables for estimating FSIQ's from these short forms.

Others have also developed short forms of the WAIS-R. Cyr and Brooker (1984), for example, averaged reliability and validity coefficients to determine the best short forms. Using data from the standardization sample they concurred that Silverstein's dyad of Vocabulary and Block Design was the best two subtest short form. However, they identified the Vocabulary, Block Design, Arithmetic, and Similarities subtests as the best tetrad. Brooker and Cyr (1986) provided

tables of age scaled scores for calculating FSIQ for their short forms. Also using the standardization sample, Reynolds, Wilson and Clark (1983) developed a four-subtest short form consisting of Information, Arithmetic, Picture Completion and Block Design.

Research to assess the best and most valid short forms has routinely utilized three criteria developed by Resnick and Entin (1971). According to these criteria: a) the correlation between the short form estimate and the FSIQ should be significant and account for the majority of shared variance, b) the mean of the short form estimate and the FSIQ should be comparable and c) the short form estimate and the FSIQ should agree on the intellectual classification according to Wechsler's (1981) seven categories. The validity of WAIS-R short forms has been tested according to these criteria in recent studies with a variety of clinical samples. Some studies support the utility of short forms (Donders, 1992; Hoffman, & Nelson, 1988; Paolo, & Ryan, 1993); others question their utility or suggest caution (Cyr, & Atkinson, 1991; Benedict, Schretlen, & Bobholz, 1992; Ryan, Georgemiller, & McKinney 1984; Watkins, Mckay, Parra, & Polk, 1987; Watkins, Himmel,

Polk, & Reinberg, 1988).

Silverstein (1985, 1990b) drew three conclusions about the criteria used in short form validity studies. First, he argued that the correlation between short form estimates and the FSIQ is always significant and usually high because this was the basis for selection of the short form. Second, he contended that the mean of the short form estimate is often different from the FSIQ. "Even the most trivial difference from a clinical perspective will prove significant with a sufficiently large sample size, and 'sufficiently large' need not be very large at all when a short form is highly correlated with the full scale" (Silverstein, 1985, p. 678). Finally, Silverstein stated that the intellectual classifications determined by the short forms are often different from those determined by the FSIQ. This is a questionable criterion since a 1 IQ point difference can result in misclassification in some instances, whereas a discrepancy as large as 19 IQ points may result in similar classification. It should be noted that short forms have not been recommended for classification purposes. They are in fact primarily endorsed as screening devices (Haynes, 1983, 1985;

Kaufman et al. 1991; Margolis, Taylor, & Greenlief 1986; Ryan, 1983; Ryan, Georgemiller, & McKinney 1984; Ryan, Larsen, & Prifetara 1983; Silverstein, 1985, 1990b).

Silverstein (1985) concluded that the Resnick and Entin (1971) criteria are not particularly useful for evaluating short forms. However, these criteria are frequently applied (Cyr, & Atkinson, 1991) and provide potential short form users with useful information. Thompson et al. (1986) also recommended reporting the confidence interval, or the accuracy of the short forms in terms ± 5 IQ points of the true IQ.

Short form validity studies have traditionally used a rescoring technique to obtain the short form IQ estimate. Thus, data from a standard full administration of the WAIS-R have been used. This method of validating short forms with clinical samples is not faithful to the conditions of a bona fide abbreviated administration during which attention and motivation are quite likely different from a full administration. During the administration of a full test, that usually takes 90 minutes, subjects may become bored, fatigued or distracted. Thompson, Howard

and Anderson (1986) were the first to point out the limitation in existing short form validity research and Kaufman (1991) has subsequently referred to this phenomenon as the "motivation effect".

To investigate the motivation effect, Thompson and his colleagues (1986) designed an empirical validity study in which the subtests that comprised Silverstein's short forms (2SF and 4SF) were administered at the beginning of the session followed by the remaining subtests to constitute a full administration. By placing the subtests of the short forms first, Thompson and his colleagues sacrificed the 'purity' of the administrative order but were able to investigate the influence of shortened administration time on validity. Thompson et al. (1986) found that the 2SF and 4SF overestimated the FSIQ but that the 2SF was more of an overestimate when administered first. This was attributed to attentional/motivational factors. It was suggested that 2SF embedded in a standard administration may have been influenced by reduced attention and motivation, whereas attention and motivation may have been optimal for the two subtests when they were administered first. Attention and

motivation did not seem to affect the 4SF.

As short forms of the WAIS-R are intended to save time, Kaufman, Ishikuma and Kaufman-Packer (1991) were interested in identifying subtest combinations that had both validity and minimal administration time. Administration times vary depending on the situation, although estimates of subtest administration times have been published (Ryan & Rosenberg, 1984; Ward, Selby & Clark, 1987). Kaufman and his colleagues selected Information and Picture Completion as a two subtest short form (2KSF) because it required only 12 minutes to administer and had adequate validity and reliability based on the standardization data. Since 2KSF is comprised of the first and second subtests in a normal administration of the WAIS-R, Kaufman pointed out that the validity data from the standardization sample maintained faithfulness to the conditions of a bona fide shortened administration. For the same reason, 2KSF can be readily validated with clinical samples by rescoring protocols.

Kaufman, Ishikuma and Kaufman-Packer (1991) selected Arithmetic, Similarities, Picture Completion and Digit Symbol as an "amazingly short" tetrad (4KSF)

The 4KSF took 19 minutes to administer and avoided Vocabulary as a subtest because of considerable subjective judgement, administration and scoring times which are potential sources of examiner error with this subtest. Kaufman argued that attentional/motivational factors would not influence 4KSF since this had not been a factor in the Thompson, et al. (1986) evaluation of Silverstein's tetrad (4SF). Hence, validity of 4KSF was based only upon its performance in the standardization sample. A validity study comparing 2KSF, 4KSF and 4SF was conducted with a psychiatric sample utilizing the rescoring methodology (Boone, 1992). The results revealed that 4SF and 4KSF were comparable in terms of validity, reliability and classification. However, 2KSF was found to have relatively low validity and poor classification (Boone, 1992).

It is possible that the brief time for administering 4KSF (i.e., 19 minutes) may influence the validity of the tetrad if its accuracy were validated other than by rescoring. Silverstein's (1982) 2SF which is of comparable duration was affected by attentional/motivational factors when evaluated using

the Thompson et al. (1986) methodology. For this reason, 4KSF should be evaluated by administering the tetrad subtests first followed by remaining subtests from a complete WAIS-R. This was one objective of the current study.

The present study was also designed to attempt to replicate the Thompson et al. (1986) results with 2SF and 4SF for a different sample. Normal subjects were chosen because of their availability and since WAIS-R short forms may perform differently with intellectually more heterogeneous samples; particularly subjects with higher intellectual functioning (Thompson et al., 1986). Finally, the methodology of the current study afforded an opportunity to evaluate the validity of 2KSF.

Method

Subjects The subjects were 80 volunteers, 50 female and 30 male, enrolled in undergraduate psychology courses at Lakehead University. The mean age of the subjects was 23 years, 10 months (range 18 to 34 years). The sample was primarily caucasian but also

included 4 native subjects.

Procedure Subjects were recruited in class and contacted later to make specific arrangements. All subjects read an introductory letter outlining the purpose of the study and consent forms were signed before participation. Subjects were not provided with personal feedback regarding their individual tests. Those subjects who received course credit for participating in the study responded to two questions pertaining to the study which were designed to facilitate their learning experience.

Subjects were assigned alternately to one of two test administrators in the order in which they appeared on the sign-up list. If a subject did not keep the experimental appointment an attempt was made to reschedule. Lost subjects were replaced by the next person on the list who had not yet been scheduled for an appointment. Prior to arrival, each subject was randomly assigned to one of the four administrative conditions. To ensure that all conditions had an equal number of subjects each block of four conditions was filled before starting a subsequent block.

The Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) was administered by two females. Each of these examiners had completed a Master's level course in Psychological Assessment and was trained in the administration of the WAIS-R. In addition, they reviewed administration procedures with a registered psychologist and were alerted to common administration errors (Moon, Blakey, Gorsuch, & Fantuzzo 1991, Slate, & Jones, 1988; Thompson, & Bulow, 1994).

Administrative Conditions

Condition 1: A two-subtest WAIS-R short form (2SF; Silverstein, 1982) consisting of Vocabulary and Block Design was administered first followed by the remaining subtests to comprise a complete WAIS-R.

Condition 2: A four-subtest WAIS-R short form (4SF; Silverstein, 1982) consisting of Picture Arrangement, Vocabulary, Block Design and Arithmetic was administered first followed by the remaining subtests to comprise a complete WAIS-R.

Condition 3: A four-subtest WAIS-R short form (4KSF; Kaufman et al. 1991) consisting of Arithmetic, Similarities, Picture Completion and Digit Symbol administered first followed by the remaining subtests to comprise a complete WAIS-R.

Condition 4: A standard WAIS-R was administered. This condition included the Kaufman et al. (1991) short form consisting of Information and Picture Completion (2KSF) as these are the first two subtests of a standard administration.

Scoring Scoring of full scale IQs (FSIQ) followed the WAIS-R manual (Wechsler, 1981) with consideration of additional information regarding common scoring errors from Slate and Jones (in press). Each protocol was scored by the examiner who administered the test and uncertain items were discussed with the other tester in order to reach a decision about the scoring. The protocols were checked for computational/clerical errors by a Registered Psychologist. In addition, the Vocabulary, Comprehension and Similarities subtests

which require the most judgement were rescored by the Registered Psychologist to ensure consistency. Changes in scoring were made as required.

Short form IQ (SFIQ) estimates were calculated according to procedures and normative tables found in Silverstein (1982) for the 2SF and 4SF, and Kaufman et al. (1991) for the 2KSF and 4KSF.

Results

Calculations were based on the eleven WAIS-R subscales recognizing that due to the experimental design, some scales were administered out of their standard order. For the entire sample, the mean FSIQ was 102.87 (SD = 11.08, range = 77 - 132); the mean Verbal IQ (VIQ) was 100.69 (SD = 10.96, range = 78 - 127); and the mean Performance IQ (PIQ) was 104.8 (SD = 13.03, range = 74 - 140). Overall summary statistics for FSIQ, VIQ and PIQ are reported in Table 1. The distribution of males and females by administrative condition did not differ significantly, $\chi^2(3) = 5.81$, $p = .12$.

Table 1

Descriptive Statistics for WAIS-R IQs by Administrative
Conditions¹

Administrative Condition	VIQ		PIQ		FSIQ		Range
	M	SD	M	SD	M	SD	
2SF First	99.5	9.2	102.7	11.9	100.7	8.9	84 - 117
4SF First	98.6	12.4	102.8	15.1	102.1	13.2	77 - 123
4KSF First	99.8	11.7	105.6	10.9	102.2	10.5	85 - 132
Standard	104.9	10.1	108.1	14.0	106.7	11.3	85 - 127

¹These are based upon 11 subscales recognizing that due to the experimental design some scales are administered out of their standard order.

Table 2 provides descriptive statistics for the various short forms of this study by administrative condition.

Silverstein's 2SF/4SF and Kaufman's 4KSF

To compare the short form estimate with the FSIQ taking into account administrative condition, a repeated measures ANOVA was performed for each short form. In each analysis, administration order (short form first versus standard) was the between subjects factor and the IQ estimate (SFIQ versus FSIQ) was the within subjects factor. For the 2SF, the results showed a main effect due to administrative order, $F(1,38) = 6.85$; $p = .01$; a main effect for IQ, $F(1,38) = 10.00$; $p < .01$ and an administrative order by IQ interaction, $F(1,38) = 6.45$; $p = .02$. Post-hoc paired t -tests were conducted to investigate the significant interaction. SFIQ and FSIQ were not significantly different when the short form was administered first followed by the remaining subtests, $t(19) = .55$, $p = .59$. However, when the short form was rescored from the standard administration, the SFIQ was significantly larger than the FSIQ, $t(19) = 3.47$, $p = .01$; $\eta^2 = .22$. For the 4SF and the 4KSF, there were no significant

Table 2

Mean and Standard Deviation of WAIS-R Full Scale I.Q.
and Estimated I.Q. From Short and Standard WAIS-R
Administrations

Administrative Condition	N	Full Scale IQ		Estimated IQ	
		M	SD	M	SD
Short Forms First					
2SF	20	100.6	8.9	101.3	10.1
4SF	20	102.1	13.2	102.3	13.8
4KSF	20	102.2	10.5	102.	9.3
<hr/>					
Standard	20	106.7	11.3		
Short Forms Rescored					
		2SF		112.6	13.2
		4SF		108.2	11.1
		4KSF		105.4	10.
		2KSF		101.4	12.4

main or interaction effects.

The accuracy of short form estimates by administrative condition was further examined by calculating correlation coefficients. The correlations between the SFIQ and FSIQ were uniformly high across all administrative conditions and are summarized in Table 3. Using Fisher's Z statistic, none of these coefficients were found to differ significantly across administrative conditions.

The proportion of SFIQ's with the same Wechsler intelligence classification (Wechsler, 1981; p. 28) as that of the FSIQ was 85% for 2SF administered first and 45% for 2SF rescored; 60% for 4SF administered first and 85% for 4SF rescored; 80% for 4KSF administered first and 60% for 4KSF extracted. The distribution of intelligence classification by administrative condition was significantly different for 2SF, $\chi^2 (1) = 7.04$, $p < .01$. The distribution of intelligence classification by administrative condition was not significantly different for 4SF or 4KSF.

The 95% confidence interval for FSIQ averaged across age ranges is ± 5.06 (Wechsler, 1981; Table 12, p.33). Short form accuracy by administrative condition

Table 3

Correlation Between FSIO and SFIO for Each Condition

Administrative				
Condition	N	r		P
<hr/>				
Short Forms First				
2SF	20	.85		<.01
4SF	20	.93		.83
4KSF	20	.89		.857
<hr/>				
Standard	20			
Short Forms Rescored				
2SF		.81		<.01
4SF		.95		<.01
4KSF		.82		<.01
2KSF		.77		<.01

was further examined by calculating the number of subjects whose SFIQ fell within the 95% confidence limits of FSIQ. These statistics are reported in Table 4. The distribution of accuracy totals by administrative condition was significantly different for 2SF, $\chi^2 (2) = 9.096, p < .01$. An inspection of the data in Table 3, suggests that this significant difference was a result of a substantial number (i.e., 7 out of 20) of the 2SF rescored short forms being more than 10 IQ points inaccurate compared with no 2SF estimates this inaccurate when administered first. The distribution of accuracy totals by administrative conditions was not significantly different for 4SF or 4KSF.

Kaufman's 2KSF

As Kaufman pointed out, his 2KSF is faithful to the conditions of a short form administration when rescored from a standard administration since the first two WAIS-R subtests (Information and Picture Completion) are used. Consequently, the standard administration of this study allowed an evaluation of 2KSF accuracy. It was found that the mean 2KSF IQ

Table 4

Estimated IQ Minus Full Scale IQ by Administrative
Conditions

Administrative Condition	Difference in IQ points/ SE_m			N
	± 5 $1SE_m$	± 6 to ± 10 $2SE_m$	$> \pm 10$ $> 2SE_m$	
Short Forms First				
2SF	13	7	0	20
4SF	16	2	2	20
4KSF	16	3	1	20
Standard				
Short Forms Rescored				
2SF	6	7	7	
4SF	18	2	0	
4KSF	12	5	3	

($M = 101$, $SD = 12.35$) was significantly different from the FSIQ (106.65), $t(19) = 3.19$, $p = .005$. The FSIQ correlated significantly with the SFIQ but was the lowest correlation of the short forms in the study (Table 3). For 2KSF, 50% of the subjects were assigned the same Wechsler Intelligence Classification (Wechsler, 1981; p. 28) as with the FSIQ. Ten of the 20 subject's SFIQ estimated by 2KSF fell within the 95% confidence limits of FSIQ (± 5 IQ points; $1SE_m$); 5 subjects were within 6 to 10 IQ points ($2SE_m$) and 5 were more than 10 IQ points ($2SE_m$) from the FSIQ.

Discussion

The results of the study showed that the validity of Silverstein's two-subtest short form (Vocabulary and Block Design) was different when the short form was administered first versus when the short form was rescored following a standard administration. Specifically, when administered first, 2SF performed better in terms of mean IQ, classification, and confidence interval accuracy. Silverstein's four-

subtest short form (4SF) and Kaufman's four-subtest short form (4KSF) did not differ according to the validity criteria when administered first versus rescored from a standard administration.

In a previous study using the same methodology as here, Thompson et al. (1986) also found that 2SF was influenced by the administrative condition but not 4SF. Although the findings in the present study were consistent in identifying the two-subtest short form as susceptible to administrative conditions, the direction of influence was different than in the earlier study. With Thompson's psychiatric sample, the 2SF mean over-estimate was 5 IQ points when administered first versus 1.5 IQ points when rescored. In the current study, when administered first 2SF was within 1 IQ point of FSIQ on average versus a 6 point over-estimate when rescored.

It is difficult to account for this discrepancy. Thompson et al. (1986) reasoned that increased attention/motivation for the 2SF when administered first was a plausible explanation for accuracy differences by administrative condition. Psychiatric patients presumably would perform better during the

early part of an intellectual assessment. Their attention and motivation might be expected to decline as the length of the assessment progressed. The university students of this study may also have been more highly motivated at the beginning of the testing session. However, they also seemed invested in doing well. It is possible that anxiety detracted from their early performance during the two-subtest short form. By the middle of the testing session, the subjects of this session seemed more relaxed and possibly more likely to do well on one or both of the subtests comprising 2SF.

Support for this line of reasoning was found by tallying the number of times that Vocabulary or Block Design was the highest (or tied for the highest) scaled score among Verbal or Performance subtests respectively. Vocabulary was the highest verbal score 5 times under the standard administration and 3 times when administered first. Block Design was the highest performance score 8 times under the standard administration versus 5 times when administered first. Further calculations were made to compare the scaled score means of the first two subtests administered in

each condition with the sixth and seventh subtests administered in each condition. Results supported the theory that subjects were more likely to do well on subtests administered later in the testing, $t(79) = -2.82, p < .01$.

A number of other studies have shown that two-subtest short forms overestimate FSIQ (Margolis, Taylor, & Greenlief, 1986; Roth, Hughes, Monkowski, & Crosson, 1984; Ryan et al., 1983; Ryan, Utley, & Worthen, 1988; Slater, & VanWagoner, 1988). As these studies utilized the rescoring methodology, the short form subtests were administered around the middle of the testing session. Some samples may be more likely to obtain higher performance on one or two subtests once they have settled into the assessment, thereby explaining the embedded overestimate of FSIQ. Thus, attentional and motivational factors may operate differently across samples with two-subtest short forms more susceptible to such influences.

The four-subtest short forms under investigation performed well as estimates of FSIQ regardless of whether administered first or embedded. Thus, both the Silverstein and Kaufman tetrad can be used with some

confidence for screening, particularly as others have also supported their validity in clinical trials (Banken, & Banken, 1987; Boone, 1991, 1992; Haynes, 1985; Ryan, Georgemiller, & McKinney, 1984; Thompson, et al., 1986). Kaufman's "amazingly short" tetrad was not susceptible to the attentional/motivational factors hypothesized to influence 2SF even though they both require about 19 minutes to administer. Presumably, the broader sampling of subtests ameliorates such influences. Although 4SF and 4KSF are reasonably valid, examiners must remain aware of the margin of error possible with abbreviated intellectual assessments. In the various administrative conditions of this study, between 10% and 40% of tetrad estimates were in error by 6 or more IQ points.

Kaufman's short form dyad which is comprised of the first two subtests of a standard WAIS-R administration did not perform well. Correlation between FSIQ and SFIQ was significant but the lowest in the study. This finding is similar to the low validity reported by Boone (1992) with a psychiatric sample. Thus, although 2KSF is fast and the short form subtest order consistent with a standard administration, it

cannot be recommended.

There are limitations to this study and the conclusions. First, the sample was comprised of only 20 subjects for each administrative condition. Second all subjects were drawn from a population of students enrolled in undergraduate psychology courses. Consequently, the results may not generalize broadly and, in particular, may be of limited relevance to clinical samples for whom abbreviated intellectual screening is most likely.

One of the strengths of this study was the meticulous checking of protocols for clerical and computational errors and for consistency in scoring subtests that require subjective judgement. A number of computational errors were found in the protocols and such mistakes are a common source of error with the WAIS-R (Slate, & Jones, in press). Changes in IQ's of this study also resulted from checking for consistent application of the scoring criteria of Vocabulary, Similarities and Comprehension. Changes to scores frequently resulted and these were usually downgrading by 1 or 2 scaled scores. Other investigations of short form validity must ensure that scoring inaccuracies do

not contribute to poor validity.

In conclusion, the methodology of this study seems important for examining the concurrent validity of WAIS-R short forms. The initial evidence suggests that some two-subtest short forms can perform differently when administered on their own as opposed to being rescored from standard administrations. Four subtest short forms seem unlikely to be influenced by attentional and motivational factors associated with an abbreviated administration.

References

- Benedict, R.H.B., Schretlen, D., & Bobholz, J.H. (1992). Concurrent validity of three WAIS-R short forms in psychiatric inpatients. Psychological Assessment, 4, 322-328.
- Boone, D.E. (1991). Item-reduction vs. Subtest-reduction short forms on the WAIS-R with psychiatric inpatients. Journal of Clinical Psychology, 47, 271-276.
- Boone, D. E. (1992). Evaluation of Kaufman's short forms of the WAIS-R with psychiatric inpatients. Journal of Clinical Psychology, 48, 239-245.
- Brooker, B.H., & Cyr, J.J. (1986). Tables for clinicians to use to convert WAIS-R short forms. Journal of Clinical Psychology, 42, 982-985.
- Cargnello, J.C., & Gurekas, R. (1987). The clinical use of a modified WAIS procedure in a geriatric population. Journal of Clinical Psychology, 43, 286-290.
- Cyr, J.J., & Atkinson, L. (1991). Use of population-specific parameters in generating WAIS-R short forms. Psychological Reports, 69, 151-167.

- Cyr, J.J., & Brooker B.H. (1984). Use of appropriate formulas for selecting WAIS-R short forms. Journal of Consulting and Clinical Psychology, 52, 903-905.
- Donders, J. (1992). Validity of two short forms of the WISC-R in children with traumatic brain injury. Journal of Clinical Psychology, 48, 364-370.
- Doppelt, J.E. (1956). Estimating the full scale score on the Wechsler Adult Intelligence Scale from scores on four subtests. Journal of Consulting Psychology, 20, 63-66.
- Haynes J.P. (1983). Comparative validity of three Wechsler short forms for delinquents. Journal of Clinical Psychology, 39, 275-278.
- Haynes, J.P. (1985). Comparative validity of two WAIS-R short forms with clients of low IQ. Journal of Clinical Psychology, 41, 283-284.
- Hoffman R.G., & Nelson, K.S. (1988). Cross-validation of six short forms of the WAIS-R in a healthy geriatric sample. Journal of Clinical Psychology, 44, 952-957.
- Kaufman, A.S. (1990). Assessing Adolescent and Adult Intelligence. New York: Allyn and Bacon.

- Kaufman, A.S., Ishikuma, T., & Kaufman-Packer, J.L. (1991). Amazingly short forms of the WAIS-R. Journal of Psychoeducational Assessment, 9, 4-15.
- Margolis, R.B., Taylor, J.M., & Greenlief, C.L. (1986). A cross-validation of two short forms of the WAIS-R in a geriatric sample suspected of dementia. Journal of Clinical Psychology, 42, 145-146.
- Moon, G.W., Blakey, W.A., Gorsuch, R.L., & Fantuzzo, J.W. Frequent WAIS-R administration errors: An ignored source of inaccurate measurement. Professional Psychology: Research and Practice, 22, 256-258.
- Paolo, A.M., & Ryan, J.J. (1993). WAIS-R abbreviated forms in the elderly: A comparison of the Satz-Mogel with a seven subtest short form. Psychological Assessment, 5, 425-429.
- Roth, D.L., Hughes, C.W., Monkowski, P.G. and Crosson, B. (1984). Investigation of validity of WAIS-R short forms for patients suspected to have brain impairment. Journal of Counseling and Clinical Psychology, 52, 722-723.

- Resnick R.J., & Entin, A.D. (1971) Is an abbreviated form of the WISC valid for Afro-American children? Journal of Counseling and Clinical Psychology, 36, 97-99.
- Reynolds, C.R., Wilson, V.L., & Clark, P.L. (1983). A four-test short form of the WAIS-R for clinical Screening. Clinical Neuropsychology, 5, 111-116.
- Ryan, J.J. (1983). Clinical utility of a WAIS-R short form. Journal of Clinical Psychology, 39, 261-262.
- Ryan, J.J., Georgemiller, R.J., & McKinney, B.E. (1984). Application of the four-subtest WAIS-R short form with an older clinical sample. Journal of Clinical Psychology, 40, 1033-1036.
- Ryan, J.J., Larsen, J., & Prifetara, A. (1983). Validity of two- and four-subtest short forms of the WAIS-R in a psychiatric sample. Journal of Consulting and Clinical Psychology, 51, 460.
- Ryan, J.J., Utley, A.P., & Worthen V. E. (1988). Comparison of two IQ conversion tables for the Vocabulary-Block Design short-form Journal of Clinical Psychology, 44, 950-952.

- Ryan, J.J. & Rosenberg, S.J. (1984). Administration time estimates for the WAIS-R subtests and short forms in a clinical sample. Journal of Psychoeducational Assessment, 2, 125 -129.
- Sattler, J.M. (1988). Assessment of Children (3rd ed.). San Diego: Sattler.
- Satz, P., & Mogel, S. (1962). An abbreviation of the WAIS for clinical use. Journal of Clinical Psychology, 18, 77-79.
- Silverstein, A.B. (1967). A short form of Wechsler's scales for screening purposes. Psychological Reports, 20, 682.
- Silverstein, A.B. (1982). Two- and Four-subtest short forms of the Wechsler Intelligence Scale-Revised. Journal of Consulting and Clinical Psychology, 50, 415-418.
- Silverstein, A.B. (1985). An appraisal of three criteria for evaluating the usefulness of WAIS-R short forms. Journal of Clinical Psychology, 41, 676-680.

- Silverstein, A.B. (1990a). Notes on the reliability of Wechsler short forms. Journal of Clinical Psychology, 46, 194-196.
- Silverstein, A.B. (1990b). Short forms of intelligence tests. Psychological Assessment, 2, 3-11.
- Slate, J.R., & Jones, C. H. (1988). Strategies to reduce examiner errors on the Wechsler Scales. Social and Behavioural Documents, 18, 1-2.
- Slate, J.R., & Jones, C.H. Errors on the Wechsler scales: Commonly mis-scored examinee responses. In Press
- Slater E.J., & VanWagoner, S.L. (1988). Validity and Utility of WAIS-R and WISC-R short forms with adolescents recovering from closed head injury. Journal of Adolescent Research, 3, 217-225.
- Thompson, A. P. (1987). Methodological issues with WAIS-R short forms. Journal of Clinical Psychology, 43, 142-144.
- Thompson, A.P., & Bulow C. A. (1994). Administration error in presenting the WAIS-R blocks: Approximating the impact of scrambled presentations. Professional Psychology: Research and Practice, 25, 89-91.

- Thompson, A.P., Howard, D., & Anderson, J. (1986). Two- and four-subtest short forms of the WAIS-R: Validity in a psychiatric sample. Canadian Journal of Behavioural Science, 18, 287-292.
- Ward, L.C. (1990) Prediction of verbal, performance and full scale IQs from seven subtests of the WAIS-R. Journal of Clinical Psychology, 46, 436-440.
- Ward, I.C., Selby, R.B. & Clark, B.I. (1987). Subtest administration times and short forms of the Wechsler Adult Intelligence Scale-Revised. Journal of Clinical Psychology, 43, 276 - 278.
- Watkins, C.E., McKay, B.L., Parra, R.A., & Polk, N.P. (1987). Using WAIS-R short forms with clinical outpatients: A cautionary note. Professional Psychology: Research and Practice, 18, 397-398.
- Watkins, C.E., Himmel, C. D., Polk, N.P., & Reinberg, J.A. (1988). WAIS-R short forms with mentally retarded adults: A cautionary note. Journal of Mental Deficiency Research, 32, 397-398.
- Wechsler, D. (1955). Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corporation.

Wechsler, D. (1981). Manual for the Wechsler Adult Intelligence Scale-Revised. New York: Psychological Corporation.

Appendix A



Appendix B

Appendix C

CONSENT FORM

I, _____ have read and understand the introductory letter of the research study entitled "Two- and Four-Subtest Short Forms of the WAIS-R: A validity study". I agree to participate in this study. I understand, however, that my participation is voluntary and that I can withdraw from the study at any time. I further understand that I will not be provided specific feedback about my own performance but will be able to obtain a summary of the overall results upon completion of the study

Signature of Participant

Date

Appendix D

Please Answer the following questions and provide your name and student number in order that you may be assigned credit for participating in this study.

1. What does the WAIS-R measure?

2. What is the value of short forms of the WAIS-R?

NAME: _____

STUDENT NUMBER _____