British Journal of Clinical Psychology • 1990), 29, 105-109 *Printed in Great Britain*

A comparison of the WAIS and WAIS-R in matched UK samples

J. R. Crawford* and **K. M. Allan**

Department of Psychology, King's College, University of Aberdeen, Aberdeen, AB9 2UB, UK

J. A. O. Besson, R. H. B. Cochrane and L. E. Stewart

Department of Mental Health, University of Aberdeen

A matched samples design was employed to compare WAIS and WAIS-R IQ in UK subjects. The WAIS yielded significantly higher mean Full Scale, Verbal and Performance IQs. The mean differences were 7.5, 6.4 and 7.9 IQ points respectively. The WAIS and WAIS-R samples were broadly representative of the UK adult population in terms of age, sex and social class distribution, and therefore provide tentative estimates of population mean scores for both Wechsler scales. Mean WAIS IQ was 108.6, suggesting that the WAIS yields inflated IQ scores in the contemporary UK copulation. Encouragingly, mean WAIS-R Full Scale IQ was 101.1, suggesting that it neither markedly underestimates nor overestimates IQ in the UK.

In the United States, the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) has now been superseded by the Wechsler Adult Intelligence Scale - Revised (WAIS-R; Wechsler, 1981). The WAIS-R has retained the original WAIS format in that it contains the same subtests which are used to derive Full Scale, Verbal and Performance IQs. However, rules for the administration and scoring of some subtests have been modified and around 20 per cent of item content has been updated. This latest version ot the Wechsler scale was standardized on a representative sample of 1880 Americans during 1976-80 (Wechsler, 1981).

Because of the changes noted above, it has been necessary to determine the comparability of WAIS and WAIS-R IQ scores. Wechsler (1981) administered the WAIS and WAIS-R, in counterbalanced order, to a sample of 72 adults aged 35 to 44, and reported that the WAIS-R yielded mean Full Scale (FS1Q), Verbal (VIQ) and Performance (PIQ) IQs which were 7.5, 6.9 and 8.0 points *lower* than those of the WAIS. Subsequent US studies employing the same test-retest design have demonstrated that the WAIS-R yields significantly lower IQ scores in a range of different samples, e.g. college students, vocational guidance clients, neurological patients and psychiatric patients 'Kelly, Montgomery Felleman & Webb. 1984; Mishra *&*. Brown, ;9S3; Prifitera & Ryan, 1983; Quereshi & Ostrowski. 1985; Urbina. Golden & Ariel.1982). The combined data from these studies indicate that the overall mean differences between WAIS and WAIS-R IQs in those samples consisting entirely of non-clinical subiects were 5.2. 5.3 and 6.2 for FSIQ, VIQ ind PIQ respectively.

In the UK. Wechsier's scales nave been the most commonly used IQ measures in both clinical and academic settings. Now that r-e UK version of the WAIS-R is available (Lea, 1986), it is to be expected that this scale will enjoy the same popularity. However, as was the case for ail of Wechsler's previous scales; a UK standardization sample has not been collected. Clearly, then, extensive research is required

*Requests for reprints.

to examine the psychometric properties of the WAIS-R in UK subjects. In the view of the authors of this Study, this research effort should, initially, be aimed at establishing baseline information on the general adult population before turning to an examination of clinical populations.

This report aims to begin this process by comparing WAIS and WAIS-R performance in UK subjects. It was decided to adopt a matched samples design rather than the test-retest design employed in US studies. There were two reasons for this. Firstly, with the test-retest method, inferences about population mean IQ on the WAIS and WAIS-R cannot be drawn from the sample means as these are inflated by practice effects, since half of the subjects tested on the WAIS have had prior exposure to the WAIS-R and vice versa. As the WAIS and WAIS-R were standardized in the US, this is not an important consideration for US studies. However, it is a matter of considerable importance and interest in the UK.

Secondly, it may be that the test-retest method does not provide a valid indication of relative differences between the WAIS and WAIS-R. Smith (1983) employed the test-retest design to compare the two tests in a sample of 72 college students. He reported a significant test x order of administration interaction, such that practice effects were larger when the WAIS followed the WAIS-R than in the opposite order of administration. Thus, the test-retest method may underestimate IQ score differences between the two tests. It is significant that when Smith (1983) compared the scores of subjects administered the WAIS at first testing with those administered the WAIS-R he found a mean difference of nine Full Scale IQ points in favour of the WAIS. This difference is larger than those reported in the studies cited above, where mean IQs were derived by combining the scores of subjects at both test periods.

However, the matched samples design is not without its difficulties. In order to attribute any differences between the samples' IQ scores to differences in the tests, rather than to differences in the samples, stringent matching of subjects on demographic variables is clearly necessary, particularly for those variables with a demonstrated relationship to IQ. In this study, the two samples were compared on an independent IQ measure to determine if matching on demographic variables successfully equated the samples on intellectual ability. The test adopted for this purpose is the National Adult Reading Test (NART; Nelson, 1982). The NART consists of 50 irregular words (e.g. deny, gauche), which subjects have to read and pronounce. In the NART standardization sample (Nelson, 1982), the number of errors made on the test was highly negatively correlated with prorated WAIS IQ. Combined factor analysis of the WAIS and NART (Crawford, Stewart, Cochrane, Parker & Besson, 1989) has revealed that the NART loads very highly (0.85) on the first unrotated principal component which is regarded as representing general intelligence.

Method

Subjects

Subjects were recruited from a wide variety of sources, e.g. non-medical health service employees, employees ot local and national companies, members of local clubs (e.g. pensioners' clubs, angling and pigeon clubs), communitv centres etc. Most were paid a small honorarium. A total of 275 subjects free from neurological, psychiatric, or sensory disability were administered either the WAIS (.V = 155) or the WAIS-R ?.V = 120) during 1987/88. All subjects also completed the NART. Administration and scoring followed standard procedures as set out in the relevant test manuals (Lea. 1986; Nelson, 1982; Saville, 1971; Wechsler. 1955, 1981).

Subjects who completed the WAIS-R were individually matched for sex, age ±4 year and years of education (± 2 years with a WAIS subject, to yield 100 matched pairs (53 males. 47 females). Pairing of subjects was carried out by a researcher to whom only the demographic data were made available.

Results

Mean age of the WAIS sample was 41.9 years (SD 18.0), and the age range 16 to 85. The corresponding figures in the WAIS-R sample were 41.7 years (SD 18.2) and 17 to 83. Mean years of education in the WAIS sample were 11.9 (SD 2.7) with a range of 7 to 19 years. The corresponding figures in the WAIS- f sample were 11.6 (SD 2.7) and 7 to 20 *y*ears.

Table 1. Social class distribution (percentages) in WAIS and WAIS-R samples and in the general UK population

	Social class				
	1	2	3	4	5
WAIS sample	8	25	41	16	10
WAIS-R sample	7	24	46	17	6
General population	5	23	48	18	6

The OPCS (1980) Classification of Occupations was used to determine social class being coded by their husband's occupation. The social class distributions of the WAIS and WAIS-R samples are presented in Table 1. A chi squared test revealed that the distribution of social class did not differ significantly in the two samples ($\chi^2(4) = 1.40$, n.s.).

The 1981 census was consulted to determine the extent to which the two samples were representative of the UK population. The distribution of social class in the general UK population is also presented in Table 1. Chi squared tests revealed that neither the WAIS nor WAIS-R sample differed significantly from the census derived distribution (WAIS sample $\chi^2(4) = 2.44$, n.s.; WAIS-R sample $\chi^2(4) = 0.43$, n.s.). A similar procedure was adopted to examine age distribution. The proportion of the UK population falling within three arbitrary age bands (16—35, 36—55, 56—85) was determined. Chi squared tests revealed that these proportions did not differ significantly from those observed in either sample (WAIS $\chi^2(2) = 0.77$, n.s.; WAIS-R $\chi^2(2) = 1.01$, n.s.).

Mean scores, standard deviations and ranges for the Wechsler scales and the NART are presented in Table 2. The Pearson product moment correlation between NART errors and Full Scale IQ (FSIQ) was - 0.78 ($p < .001$) in the WAIS sample and -0.72 ($p < .001$) in the WAIS-R sample, thus demonstrating that NART performance was a good predictor of IQ score in both samples. Mean NART errors in the two samples were compared by a matched samples *t* test, and were found not to differ significantly ($t = 1.88$, n.s.).

Table 2. Means, standard deviations and ranges for the NART and Wechsler scales in subjects administered either the WAIS or WAIS-R

	NART errors	Full-scale IQ	Verbal IQ	Performance IQ
WAIS sample				
M	22.9	108.6	108.4	107.6
SD	10.9	12.0	13.4	11.4
Range	1–45	75–132	72–135	80–130
WAIS-R sample				
M	21.1	101.1	102.0	99.8
SD	9.3	13.5	13.1	13.1
Range	3–41	72–140	74–134	71–139

It can be seen from Table 2 that mean FSIQ in the WAIS sample was 108.6 (SD 12.0). In the WAIS-R sample, mean FSIQ was 101.1 SD 13.5). Paired samples *t* tests revealed that the WAIS yielded significantly higher IQ scores than the WAIS-R for FSIQ ($t = 7.46$. $p < .001$.) Verbal IQ ($t = 6.39$, $p < .001$) and Performance IQ ($t = 7.90$. $p < .001$). The mean differences were 7.5. 6.4 and 7.9 IQ points, respectively.

Discussion

The highly significant differences between WAIS and WAIS-R IQs in this study strongly suggest that the two tests do not yield comparable results in the UK. The alternative explanation that these differences reflect underlying differences in the intellectual abilities of the samples is not convincing. As noted, subjects were individually matched for age, education and sex, and did not differ significantly in terms of social class. More importantly, they did not differ significantly in their performance on the NART, an independent estimate of IQ.

The mean differences in favour of the WAIS of 7.5, 6.4 and 7.9 IQ points for FSIQ, VIQ and PIQ, respectively, are broadly consistent in direction and magnitude with those reported in US studies. The study under report, however, avoids the problems of interpretation associated with the test-retest design, has the largest sample size to date, and is the only study in which the sample can be viewed as at all representative of a general adult population. It could be argued, therefore, that it provides a more valid indication of IQ score differences for the UK than is currently available for the US.

The WAIS and WAIS-R samples used in this study were both broadly representative of the UK population in age, sex, and social class distribution. They therefore provide estimates of *population* mean scores on the Wechsler scales, albeit *very tentative ones.* Mean WAIS FSIQ was 108.6, suggesting that the WAIS overestimates IQ in the contemporary UK population. To our knowledge, the only previous attempt to obtain a UK population estimate for the WAIS was that of Wilson, McCullough & Morris (1987). This study (n = 63) indicated that the WAIS yielded an even more inflated IQ than that observed in the study reported here (mean FSIQ = 116.5 ± 9.74). It is not clear, however, how representative the sample was. Mean age was only 27.7 (SD 12.2) and, although the authors reported that there was no white-collar bias, the social class distribution was not presented.

Encouragingly for UK users of the WAIS-R, mean FSIQ in this study (101.1) was very close to 100. In order to apply normal interpretive rules to WAIS-R IQ in the UK, the population standard deviation should be 15. It can be seen from Table 2 that, despite the apparently successful attempt to recruit subjects who varied considerably in terms of demographic characteristics, the sample SD was 13.5. An F test, however, revealed that this SD was not significantly different from 15 for a sample size ot 100($F(99, x) = 1.23, > .1$).

It-must be stressed that the above conclusions must be regarded as tentative, since they are based on samples in which subjects came from only one geographical region and were almost entirely urban dwellers. There is a need for additional WAIS-R samples recruited from other UK centres, and for an eventual pooling of results. As noted, the WAIS has been widely used in the UK, despite the virtual absence of research on its appropriateness for a UK population. It is to be hoped that the use of the WAIS-R will proceed on a more sound empirical footing.

The practical implications of these findings could be stated as follows. Firstly, as the WAIS and WAIS-R do not yield comparable IQ scores, retesting of subjects to monitor decline or recovery should be conducted with the same version of the Wechsler scales (see Flynn [1984] for a discussion of the problems encountered in longitudinal investigations where the IQ tests vary in the obsolescence of their norms). Secondly. Since in clinical decision making involving the Wechsler scales. IQ scores are compared with other test results and with other clinically relevant information, clinicians switching from the WAIS to WAIS-R should take close account of the score differences reported here between the two scales. Finally, the results obtained in this report suggest that, unlike the WAIS, the WAIS-R does not markedly overestimate IQ in the UK. On the evidence presented here, the WAIS-R would appear to yield a more valid IQ, and it should therefore be used in preference to the WAIS in clinical practice and research.

References

Crawford. J. R., Stewart. L. E.. Cochrane. R. H. B.. Parker. D. M. & Besson, J. A. O. (1989). Construct validly of the National Adult Reading Test: A factor analytic study. *Personality and Individual Differences.* 10. 585-58". Flynn, I. R. ,1984). The mean IQ of Americans: Massive gains 1932 to 19~8. *Psychological Bulletin,* 95,
29-51.

Kelly, M. P., Montgomery, M. L., Felleman, E. S. & Webb, \X'. W. (1984). WAIS and WAIS-R in a neurologically impaired population, *journal of Clinical Psychology,* 40, "88—790. Lea, M. (1986). *A British Supplement to the Wechsler Adult Intelligence Scale - Revised.* Windsor: N'FER-Nelson. Mishra, S. P. & Brown, K. H. (1983). The comparability of WAIS and WAIS-R IQ and subtest scores.
journal of Clinical Psychology, 39, 754-757.
Nelson, H. E. (1982). *National Adult Reading Test f\'ART): Test Manual.* Windsor: NFER-Nelson.
Office of Population, Censuses and Surveys (1980). *Classification of Occupations.* London: HMSO.
Prifitera, A." & Ryan, J. H. (1983). WAIS-R/WAIS comparisons in a clinical sample. *Clinical Neuropsychology, 5,* 97-99. Quereshi, M. Y. & Ostrowski (1985). The comparability of three WAIS's in a college sample. *Journal of Clinical Psychology,* 44, 397-407. Saville, P. (1971). *A British Supplement to the Manual of the Wechsler Adult Intelligence Scale.* Windsor: NFER-Nelson. Smith, R. S. (1983). A comparison study of the WAIS and the WAIS-R in a college population. *Journal of Consulting and Clinical Psychology,* 51, 414—U9. Urbina, S. P., Golden, C. J. & Ariel, R. N. (1982). WAIS/WAIS-R: Initial comparisons. *Clinical Neuropsycholofy.* 4 145—146. Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale.* New York: Psychological Corporation. Wechsler, D. (1981). *Manual {or the W'echsler Adult Intelligence Scale — Revised.* New York: Psychological Corporation. Wilson, A., McCullough, J. A. & Morris, R. C. (1987). WAIS scores for a non-clinical British sample.
British Journal oj Clinical Psychology, 26, 57—58.

Received} December 1988; revised version received 10 April 1989