

WAIS-R Performance Following Closed-Head Injury: A Comparison of the Clinical Utility of Summary IQs, Factor Scores, and Subtest Scatter Indices*

J. R. Crawford¹, D. A. Johnson², B. Mychalkiw³, and J. W. Moore⁴

University of Aberdeen, ²Department of Clinical Neuropsychology, Astley Ainslie Hospital, Edinburgh,
³City General Hospital, Stoke-on-Trent, ⁴Aberdeen Royal Infirmary, Aberdeen, UK

ABSTRACT

WAIS-R performance can be analyzed at the level of the IQs, factor scores, or subtests. This study examined the relative utility of such analyses in a closed-head-injured (CHI) sample ($n = 233$). Comparison of the sample with matched controls ($n = 117$) indicated that factor scores may be the most appropriate level of analysis in clinical practice. This conclusion was supported by a series of discriminant function analyses performed on the WAIS-R scores of the CHI sample and 356 healthy controls drawn from the general population. Factor scores achieved a significantly higher classification accuracy than IQs or indices of subtest scatter (all p 's $< .001$). Combining the scatter indices with the composite measures (i.e., IQs or factor scores) did not produce a significant improvement in discrimination over that achieved by the composites alone. Counter to the assumptions underlying the clinical analysis of subtest profiles, CHI participants failed to exhibit significantly greater *intraindividual* subtest variability in their subtest profiles than the matched control sample.

The Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) is widely employed in clinical practice to assess the cognitive sequelae of closed-head injury (CHI) (Lezak, 1995; McKinlay & Gray, 1992; Richardson, 1990; Walsh, 1991). However, currently there is little published research on the characteristics of WAIS-R performance following a CHI (Crosson, Greene, Roth, Farr, & Adams, 1990). The studies conducted to date have predominantly employed small samples and have been concerned with the WAIS rather than WAIS-R. These WAIS studies have typically reported deficits in both Verbal and Performance IQ but with deficits being more pronounced on the latter scale (see Crosson et al., 1990; and Richardson, 1990 for reviews). Where performance on the individual subtests has been reported, Arithmetic and Digit Span have typically been the most

impaired of the Verbal subtests with Digit Symbol exhibiting the largest Performance Scale deficit (Becker, 1975; Richardson, 1990).

A number of authors have pointed out the dangers of assuming that the results of research obtained using the WAIS will be applicable to the WAIS-R (e.g. Bornstein, 1987; Reitan & Wolfson, 1990). Therefore, there is a need for new data obtained from CHI populations to aid clinicians in their interpretation of WAIS-R performance.

Clinical interpretation of WAIS-R performance can be based on analysis at the level of the summary IQs, factor scores, or the subtests. Authorities on the Wechsler have divided opinions on the usefulness of the summary IQs. Matarazzo and Herman (1985) for example, have suggested that the discrepancy between Verbal and Performance IQs is the best

validated Wechsler index of cerebral dysfunction, whereas Lezak (1995) is dismissive of its importance. Many clinicians base their interpretation of the WAIS-R primarily on the pattern of strengths and weaknesses in a client's subtest profile (e.g., Lezak, 1995; Walsh, 1991). Lezak (1988a) for example, notes that IQs can obscure clinically important strengths and weaknesses and suggests that "although the traditional scoring scheme for the Wechsler exemplifies the IQ problem, these tests also show us a way out of the problem by providing a profile of scores" (p. 359).

However, the interpretation of subtest profiles is not without its own difficulties. The rationale underlying this approach is that, because individuals differ widely in terms of premorbid ability, the clinician needs to employ individual rather than normative comparison standards when attempting to detect impairment. It is assumed that some subtests will be relatively spared following neurological injury and thus will provide these comparison standards. Thus, in inferring impairment, the emphasis is placed more on the presence of large discrepancies among an individual's subtest scores than on comparison of these subtest scores with the relevant norms (e.g., Crawford, 1992; Lezak, 1995).

There has been limited research on the utility of this approach in detecting and quantifying cognitive dysfunction. Furthermore, studies of healthy samples have revealed a surprisingly large degree of *intraindividual* variability; a relatively flat subtest profile is the exception rather than the rule in the general population (e.g., Crawford & Allan, 1996; Kaufman, 1990; Matarazzo, Daniel, Prifitera, & Herman, 1988). These studies suggest that there is a danger of over-inference when discrepancies between subtests are examined in clinical populations and underlines the need for statistical indices to supplement clinical judgment. Two such indices will be considered in the present study. McLean, Reynolds, and Kaufman (1990) have suggested using the variance of an individual's subtests as an index of subtest scatter: they termed this the Profile Variability Index (PVI).

A recently proposed alternative approach to the analysis of subtest discrepancies uses the

Mahalanobis Distance as an index (Burgess, 1991). The Mahalanobis Distance Index (MDI) is a multivariate indicator of the abnormality (i.e., rarity) of an individual's scatter and has a chi-square distribution with 11 degrees of freedom when used with a full-length WAIS-R. In the present study it was hypothesized that the MDI would achieve better discrimination between healthy and CHI participants than the PVI as it factors in the estimated population correlations between subtests when evaluating the abnormality of discrepancies. For example, an extreme Digit Span score would be given less weight than an equally extreme Vocabulary score as the former subtest has a lower average correlation with the other subtests; that is, extreme Digit Span scores will occur with greater frequency in the healthy population. The greater sophistication of the MDI has a practical cost. In contrast to the PVI, which can be rapidly calculated by hand, the calculation of an individual's MDI requires a computer as it makes use of the subtest covariance matrix. Thus, the MDI would have to significantly outperform the PVI to justify its routine use in clinical practice.

An alternative to the analysis of both IQs and subtest discrepancies is to examine scores obtained from factor analysis of the WAIS-R (Atkinson, 1991; Canavan, Dunn, & McMillan, 1986; Crawford, 1992; Kaufman, 1990). Although this suggestion has a long history (e.g., Maxwell, 1960), factor scores have not been widely employed in clinical practice or research. The present authors are unaware of any research which has examined WAIS or WAIS-R performance in CHI in terms of factor scores, despite the indications that the factor structure of the WAIS-R is preserved in CHI and other neurological disorders (e.g., Atkinson, Cyr, Doxey, & Vigna, 1989; Leckliter, Matarazzo, & Silverstein, 1986). This neglect of the factor score approach may be unfortunate as it has some advantages over the use of summary IQs or subtests; unlike IQs, factor scores are empirically derived composites rather than composites formed a priori, they therefore have superior construct validity. Because they are composites of correlated components, they necessarily possess superior

reliability to the subtests from which they are formed (e.g., see Guilford, 1954).

Factor analyses of the WAIS and WAIS-R have commonly extracted three factors: a *verbal* (V) factor on which Information, Vocabulary, Comprehension, and Similarities have high loadings, a *perceptual organization* (PO) factor defined by high loadings from Block Design and Object Assembly, and a third factor termed *attention/concentration* (A/C) or *freedom-from-distractibility* consisting of Arithmetic, Digit Span, and Digit Symbol (e.g., Crawford, Allan, Stephen, Parker, & Besson, 1989; Leckliter et al., 1986). The factor score approach may be particularly useful in the interpretation of WAIS-R performance following a CHI. Lezak (1988b) for example, in arguing the case against IQs and for examination of subtest profiles, notes that "Many persons who have suffered a mild concussion with loss of consciousness ... tend to perform relatively poorly on Digits Backward, Arithmetic and Digit Symbol" (p. 358). This clinical observation is in line with the present authors' experience and the research reported above, but it could be argued that it provides a justification for the use of factor scores as much as subtest profile analysis because these three subtests make up the *attention/concentration* factor.

In the present investigation the WAIS-R performance of a sample of CHI patients was compared with a matched sample of healthy controls to identify which components of WAIS-R performance typically reveal the largest deficits and which are relatively preserved. The aim was to update similar data obtained with the WAIS with the exception that factor scores were also examined. The second component of the study employed discriminant function analysis and was prompted by the fact that the principal concern for clinicians is how successful these indices are liable to be at detecting impairment when used with *individual* clients. Furthermore, in clinical work, healthy samples matched for age, sex, and education are obviously not readily available to provide comparison standards against which a client's performance can be compared. Because of this we employed a large sample drawn from the *general* adult population

to serve as the healthy group in the discriminant function analyses.

The foregoing discussion of WAIS-R indices focused on their relative merits. However, in clinical practice interpretation of the WAIS-R will be based on a consideration of these indices *in combination* (analysis of the qualitative features of performance also plays an important role, but is beyond the province of the present investigation). A final aim of the present study was therefore to examine whether the combination of composite scores (i.e., the IQs or factor scores) and indices of subtest scatter would achieve better discrimination between the performance of and healthy individuals and patients with CHI than either of these sets of measures alone.

METHOD

Participants

Individuals who had suffered a CHI requiring a period of inpatient care ($n = 233$; males = 172, females = 61) were recruited from referrals to four clinical neuropsychology services in the UK. Mean years of age in this sample was 32.3 ($SD = 13.07$). Mean years of education was 11.7 ($SD = 2.16$). English was the first language for all participants (this also applied to all controls). Individuals were excluded if there was a preexisting psychiatric or neurological disorder, a history of alcoholism/drug dependency, or if they had not completed a full-length WAIS-R. All participants were judged to have exited post-traumatic amnesia (PTA) at the time of testing. The mean time elapsed between injury and testing was 33.4 months ($SD = 23.29$) with a median of 25 months. Reasonable estimates of length of PTA could be determined for 143 cases although, because of the difficulties in assessing this variable and differences across centers in their approach to its measurement, the data should be treated with caution. In most centers PTA was assessed retrospectively from the case notes and interviews with clients. Mean length of PTA was 21.9 days ($SD = 31.57$) with a median of 10 days. The percentage of cases for whom PTA information was available (61%) was similar to that obtained by Crosson et al. (1990) in their study of WAIS-R cluster patterns in CHI (59%) as was the mean length of PTA (Crosson et al., mean = 28.44 days). A control sample was recruited to match the CHI sample for age, years of education, and sex

ratio ($n = 117$; males = 81, females = 36). The mean years of age in this sample was 34.9 ($SD = 14.04$) and the mean years of education was 11.9 ($SD = 2.21$). The samples did not differ significantly in terms of mean age ($t = 1.67, p > .05$) or mean years of education ($t = 0.73, p > .05$). Additionally, F tests revealed that the samples did not differ significantly in terms of the variances of age ($F = 1.15, /7 > .05$) or years of education ($F = 1.05, p > .05$). A Chi Square test revealed that the sex distribution did not differ significantly between the two samples, $f(1, N = 350) = 0.61, p > .05$.

A further large healthy sample ($n = 356$) was recruited from the general adult population for use in the discriminant function analyses. Mean years of age in this sample was 42.3 ($SD = 17.16$) and mean years of education was 12.15 ($SD = 2.86$).

Potential participants in either of the two healthy samples were screened by interview to exclude those with known neurological, psychiatric, or systemic disorders. The majority received a small honorarium for their participation and were recruited from a variety of sources including commercial and public service organizations, community centers, and recreational clubs.

Procedure

All participants completed a full-length WAIS-R according to standard procedures (Lea, 1986; Wechsler, 1981). In a study of the basic psychometric properties of the WAIS-R, Crawford, Gray, and Allan (1995) reported that there are significant differences among subtest means in the UK. These authors provided a table (Table 9) which adjusts US-derived age-graded scores to have a common mean of 10 and SD of 3 in the UK. The table was derived from a sample of 200 healthy participants recruited to match the UK census in terms of the distributions of age, sex, and occupational code. In the present study these adjusted scores were used in all analyses involving age-graded scaled scores.

Each participant's Profile Variability Index score was calculated by entering their subtest scaled scores into the relevant formula (McLean et al., 1990). Each subject's age-graded subtest scaled scores were entered into a PASCAL program which calculated the Mahalanobis Distance Index scores (see Crawford & Allan, 1994). Because MDI scores are distributed as Chi Square, a square-root transformation was applied prior to conducting any statistical analysis.

To obtain scores on the Verbal, PO, and A/C factors the WAIS-R averaged subtest intercorrelation matrix, obtained from the WAIS-R manual (Table 16; Wechsler, 1981), was subjected to a principal components analysis with a priori extrac-

tion of three factors and varimax rotation. The results of this analysis are presented in Table 1; for each factor the factor loadings appear in the first columns and the factor score coefficients in the second. Examination of the factor loadings reveals the expected pattern of loadings described in previous factor analytic studies in US and UK samples (e.g., Crawford et al., 1989; Leckliter et al., 1986). The steps followed to factor score WAIS-R performance are given below in sufficient detail to allow the reader to use the method in their own practice: (1) each of an individual's age-graded subtest scores were converted to z scores; for example, as age-graded subtest scores are standardized to have a mean of 10 and SD of 3, a subtest score of 6 would produce a z score of -1.33; (2) for each factor, the z scores for each of the 11 subtests were multiplied by their corresponding factor score coefficients (Table 1) and then summed; (3) factor scores are expressed as z scores if they have a mean of zero and SD of one. These scores were transformed so that, in common with the IQs, they had a mean of 100 and SD of 15. Because the factor scores have a mean of zero, this can be achieved by simply multiplying each of the three factor scores by 15 and adding 100 to the products. The dependent correlation method was employed to determine whether a deficit exposed by any particular index was significantly greater than that exposed by the alternative indices. This method was proposed by Baron and Treiman (1980) and endorsed by Strauss and Allred (1987) as a means of establishing evidence for differential deficits following closed-head injury. The point-biserial correlations of group membership (i.e., CHI vs. control in the present case) with scores on competing indices are tested for a significant difference using Williams' (1959) test. This test factors in the correlation between the indices and yields a t value with $TV-3$ degrees of freedom.

RESULTS

Summary statistics for the three IQ scales (FSIQ, VIQ, and PIQ), the three factors, the PVI, and MDI are presented in Table 2. The results of t tests comparing the CHI and matched control sample on these indices are also presented in Table 2. A Bonferroni correction was applied to reflect the fact that eight comparisons were conducted (i.e., three IQ scales, three factors, and two scatter indices). For example, to maintain the family wise error rate at .05, p for

Table 1. Factor Loadings (FL) Obtained by Subjecting the Averaged Correlation Matrix from the WAIS-R Standardization Sample to Principal Components Analysis with Varimax Rotation.

Subtests	Factor Loadings					
	Verbal		PO		A/C	
	FL	FSC	FL	FSC	FL	FSC
Information	.84	.355	.25	-.150	.29	-.080
Comprehension	.81	.349	.26	-.121	.26	-.110
Vocabulary	.87	.341	.27	-.163	.34	-.031
Similarities	.76	.301	.34	-.047	.25	-.116
Digit Span	.59	-.193	.30	-.222	.83	.756
Arithmetic	.67	.002	.40	-.080	.59	.353
Picture Completion	.43	.084	.64	.353	.11	-.262
Picture Arrangement	.48	.141	.49	.214	.13	-.211
Block Design	.33	-.199	.79	.410	.38	.101
Object Assembly	.14	-.236	.86	.597	.19	-.086
Digit Symbol	.43	-.209	.50	.061	.61	.466

Note. Factor Score Coefficients (FSC) form the second column for each factor.

Table 2. Summary Statistics for IQs, Factors, and Scatter Indices in CHI and Control Samples.

Index	CHI Sample ($n = 233$)		Control Sample ($n = 117$)		t	p^a
	M	SD	M	SD		
Summary IQs						
FSIQ	88.7	14.32	102.0	13.28	8.37	< .001
VIQ	90.6	14.49	102.1	12.82	7.27	< .001
PIQ	87.9	14.83	101.7	13.81	8.42	< .001
Factors						
Verbal	96.5	16.9	99.5	16.2	1.62	ns
PO	93.6	17.3	99.2	14.4	3.20	< .05
A/C	79.6	15.7	100.9	15.4	12.00	< .001
Scatter Indices						
PVI	4.62	2.81	5.26	2.83	2.46	ns
MDI ^b	3.35	0.75	3.32	0.78	0.41	ns

^a The p values are those obtained after applying a Bonferroni correction.

^b The values reported are for the square-root transformation of MDI scores.

any single comparison would have to exceed .006.

It can be seen from Table 2 that the CHI sample obtained significantly lower scores on all

three IQ scales ($p < .001$ for all comparisons). Paired t tests revealed that the means for VIQ and PIQ did not differ significantly from each other in the control sample ($t = 0.45, p > .05$) but

PIQ was significantly lower than VIQ in the CHI sample ($t = 3.85, p < .001$). However, the magnitude of the difference in the CHI sample was modest as indicated by a small effect size ($r^2 = 0.06$).

As can be seen from Table 2 the two samples differed significantly on the attention/concentration and perceptual organization factors but not on the verbal factor. The deficit on attention/concentration was particularly marked ($t = 12.0, p < .001$); the mean difference between the control and CHI samples was 21 points (i.e., $> 1.3 SDs$). Application of the dependent correlation method for testing for differential deficits revealed that the deficit on the attention / concentration factor was significantly greater than the deficits on either of the other two factors and also greater than the deficits observed on all three of the IQ scales; t values ranged from 2.84 to 6.73 ($df = 347$ and $p < 0.001$ for all comparisons).

The differences between the control and head-injured samples on both the PVI and MDI failed to achieve statistical significance. Thus there was no evidence for an increase in subtest scatter in the head-injured sample.

Summary statistics for the WAIS-R subtests in the CHI and control samples are presented in Table 3. Subtest performance is expressed as *age-graded* scaled scores as these are used in clinical practice when examining subtest profiles (e.g., Lezak, 1995).

Two sets of discriminant function analyses (DFAs) were run with the CHI sample and large,

healthy sample. Five DFAs were conducted in the first set to examine the classification accuracy achieved by (1) FSIQ; (2) the combination of VIQ and PIQ; (3) the combination of the three factor scores; (4) the PVI, and (5) the MDI.

The results of these analyses are presented in Table 4 which records the percentages of participants correctly classified as head-injured or healthy, χ^2 and its probability level for each discriminant function, and the canonical correlations. It can be seen from Table 4 that the factor scores achieved a higher classification accuracy than either the IQs or scatter indices. To determine if the factors achieved *significantly* better discrimination than the summary IQs or scatter indices, forced entry regression analyses were performed with each set of variables as the predictors and group membership as the criterion. The correlations of the regression scores derived from these analyses with group membership are directly equivalent to the canonical correlations in Table 4. However, by using regression scores it is possible to compute the correlation *between* the competing sets of discriminating variables and thus employ the significance test for dependent correlations outlined previously. The results of this procedure are presented in the last two columns of Table 4 from which it can be seen that the factor scores achieved significantly better discrimination than the summary IQs and scatter indices ($p < .001$ for all comparisons).

Table 3. Summary Statistics for WAIS-R Subtests in CHI and Control Samples.

Subtest	CHI Sample ($n = 223$)		Control Sample ($n = 117$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Information	8.03	3.23	10.00	3.09
Comprehension	8.10	3.61	10.11	2.99
Vocabulary	7.89	3.44	9.80	3.20
Similarities	8.60	3.58	9.72	2.93
Digit Span	6.94	3.14	9.92	3.29
Arithmetic	6.92	3.43	10.53	2.86
Picture Completion	8.22	2.87	9.62	2.95
Picture Arrangement	7.44	3.07	9.90	3.08
Block Design	8.11	3.21	10.19	2.80
Object Assembly	8.51	3.76	9.89	2.91
Digit Symbol	5.27	3.48	9.73	3.07

The second set of analyses involving the CHI and large healthy sample employed hierarchical discriminant function analysis to determine whether the addition of either of the two scatter indices would improve upon the discrimination achieved by the composite variables alone (i.e., the IQs and factor scores). Thus, three hierarchical analyses were run in which either FSIQ, the combination of VIQ and PIQ, or the three factor scores were entered into a DFA with subsequent entry of the PVI. The same analyses were then re-run using the MDI in place of the PVI. When either the PVI or MDI was entered into the DFAs the change in the percentage of cases correctly classified by the composite variables was uniformly modest (< 2% for all comparisons). McNemar-repeated-measures Chi Squared tests revealed that the addition of either of the scatter indices did not significantly improve upon the classification accuracy achieved by any of the composites alone (all $ps > .05$). This method was recommended by Tabachnick and Fidell (1989) and assessed whether the number of participants incorrectly classified initially and subsequently correctly classified differed significantly from the number in which the converse occurred (participants in whom there is no change are ignored).

DISCUSSION

In the present study the comparison of the CHI sample with matched controls confirm the pres-

ence of significant intellectual deficits following closed-head injury. It also confirms WAIS findings of a significant discrepancy between Verbal and Performance IQ in favor of the latter (e.g., see Richardson, 1990). However, the mean difference (2.7 points) is modest, particularly when set against the large discrepancies between the factorially derived composites; for example, the discrepancy between the attention / concentration and verbal factors exceeded 16 points. The small discrepancies in the means (and correspondingly small effect size) suggests that the VIQ/PIQ discrepancy will be of little or no value as an index of impairment when used with individual clients. This was confirmed by the results from the discriminant function analyses. The classification accuracy achieved by the combination of the VIQ and PIQ scales (71.7%) was only marginally greater than that achieved by FSIQ (70.8%); thus the contrast between VIQ and PIQ is not an important discriminating feature of CHI performance in the individual case. Comparison of the results for the IQs and factor scores provides empirical support for Le-zak's (1988a) point that summary IQs can obscure important cognitive strengths and weaknesses. It will be noted that a highly significant difference in VIQ was obtained when the CHI group was compared with matched controls. However, examination of the factor scores (and individual subtest scores) indicates that the principal reason for this difference arises from the inclusion of the two Verbal subtests which index the attention/concentration factor in the compu-

Table 4. Results from Discriminant Function Analyses Comparing Classification Accuracy of Summary IQs, Factor Scores, and Scatter Indices.

Variables	% Correctly Classified	Chi Square	<i>P</i>	Canonical Correlation	<i>t</i>	<i>P</i>
FSIQ	70.8	150.4	< .0001	.476	5.40	< .001
VIQ & PIQ	71.7	156.6	< .0001	.484	4.86	< .001
Factor Scores	76.9	247.2	< .0001	.587	-	-
PVI	52.3	3.9	< .05	.081	12.91	< .001
MDI	53.7	3.5	ns	.077	9.3	< .001

^a Results of testing for a significant difference in dependent correlations for factor scores (*r*) versus the other indices (*o*)\ $df = N - 3 = 586$ for all comparisons.

tation of VIQ. Scores on the verbal factor did not differ significantly between the CHI and matched control samples demonstrating that CHI does not typically seriously impair the largely well-consolidated verbal abilities tapped by this empirically derived composite; as can be seen from the factor score coefficients in Table 1, Arithmetic and Digit Span make little contribution to this factor.

In contrast to the verbal factor results, a severe deficit was apparent on the attention/concentration factor (to which Digit Span and Arithmetic are two of the three principal contributors). Indeed the difference between the CHI and control groups on this factor (21 points) was significantly greater than the difference obtained on PIQ. Thus, an apparently moderate deficit on VIQ masked an absence of deficit on the verbal factor and a severe deficit in attention/concentration.

The results for the attention/concentration factor and examination of the means for the subtests which contribute to this factor also provide empirical support for Lezak's (1988a) clinical observation which identified Arithmetic, Digit Span and Digit Symbol as the subtests which typically expose problems following a CHI. It can be seen from Table 3 that these subtests yielded the largest differences between the CHI and control samples. This result is also consistent with the findings of earlier, smaller scale, studies which employed the WAIS rather than WAIS-R (Richardson, 1990).

The discussion of results to this point has mainly concentrated on the comparison of mean differences between the CHI and matched control samples in order to identify the typical effects of a CHI on WAIS-R indices. However, although the results provide useful information to guide assessment, the principal concern for clinicians is how useful these indices are liable to be when used with *individual* clients. The discriminant function analyses addressed this latter issue.

The discriminant function analyses were not aimed primarily at determining the *absolute* level of classification accuracy achieved by WAIS-R indices (nor at deriving discriminant functions for actuarial purposes). The principal

aims were to assess the *relative* utility of competing indices in detecting deficits in CHI and to assess empirically whether combining clinical interpretation of composites and subtest variability is liable to have advantages over focusing interpretation on either of these sets of variables alone. However, some brief commentary on the absolute levels of accuracy achieved by the indices is warranted. It can be seen from Table 4 that, even for the factorially derived composites, the classification accuracy achieved is modest (77%). Such a result is not unexpected. The large and highly significant differences between mean scores for the CHI and matched control samples on the majority of the composites indicate that the present CHI sample exhibited marked cognitive deficits. However, the sample was heterogeneous with regard to injury severity and substantial WAIS-R deficits would not be anticipated following a mild CHI. Furthermore, although as Lezak (1988b) notes, the Wechsler is "the single most utilized component of the neuropsychological repertory" (p. 53), it constitutes only a small part of the armamentarium that the clinical neuropsychologist will utilize in assessing the cognitive consequences of a CHI. On the basis of the comparison of mean scores for the CHI and matched control sample it was concluded that a factor score approach to the clinical interpretation of WAIS-R performance is to be preferred over the use of summary IQs or examination of subtest scatter. The results of the discriminant function analyses support these conclusions. Factor scores were significantly better than conventional summary IQs at discriminating between the WAIS-R performance of individuals with a CHI and a large sample of the general, healthy population ($p < .001$ for all comparisons). Given that factor scores also possess superior construct validity (i.e., they are empirically derived composites rather than composites formed in an ad hoc fashion) a strong case can be made for their use in clinical practice. Furthermore, although some authorities have stressed the analysis of subtest scatter and subtest profiles (e.g., Lezak, 1995), the WAIS-R could be viewed as primarily providing broad indicators of current functioning against which more specific neuropsychological mea-

asures are compared. The superior construct validity of the factorially derived composites suggest that they are better suited to provide these broad intraindividual comparison standards than are IQs.

In clinical practice a client's subtest profile is examined to assist in the process of identifying, quantifying, and specifying the nature of any acquired deficits. The presence of acquired deficits is inferred when there are marked discrepancies between subtests (e.g., Crawford, 1992; Lezak, 1995, 1988b; Walsh, 1991). However, clinicians have not normally had the opportunity to administer the WAIS-R, or indeed their other assessment instruments, to significant numbers of non-clinical individuals. Because of this, we may form inaccurate or distorted impressions of the degree of subtest variability that occurs in the general, healthy population. It is, therefore, important to consult relevant base-rate data as it would be difficult to justify drawing inferences regarding the presence and nature of deficits from a clinical subtest profile in which the degree of subtest variability is likely to be equaled or exceeded by a substantial proportion of the healthy population.

The results of the present study are disappointing for clinicians who emphasize the importance of analyzing subtest profiles and run counter to the present authors' own expectations. The differences between the CHI and matched control sample on the PVI and MDI scatter indices failed to achieve significance, thereby failing to support the assumption that intraindividual subtest variability will be increased following a CHI. Indeed, on the PVI the trend was for the matched controls to exhibit greater variability than the CHI cases although this difference failed to achieve significance after applying the Bonferroni correction. The results of the discriminant function analyses underlined that intraindividual variability in subtest performance is unlikely to be of clinical significance; these indices performed at essentially chance levels in discriminating CHI participants from the large sample of participants drawn from the general healthy population.

Scatter indices are sensitive to *intraindividual* variability and are relatively insensitive to abso-

lute levels of performance. For example, the MDI score for an individual who obtained age-graded scaled scores of 1 on all subtests would not be statistically significant (Burgess, 1991) and the variance of the profile (i.e., the PVI score) would be zero. It might, therefore, have been expected that combining these indices with composites which express averaged performance in absolute terms would have led to an improvement in classification over the use of either of these sets of variables alone.

In view of the clear superiority of the composites over the scatter indices (see Table 4), this issue reduces to the question of whether inclusion of the scatter indices improves upon the discrimination achieved by the composites alone. The results of the hierarchical discriminant function analyses revealed that neither of the scatter indices produced a significant improvement in classification.

These results suggest that giving weight to information on a client's subtest scatter may do little to aid the accuracy of clinicians' attempts to identify the presence of acquired deficits. However, they do not undermine the *general* utility of evaluating a client's strengths and weaknesses as the factorially derived composites were significantly better at discriminating between healthy and CHI samples than was FSIQ. Thus, there is value in incorporating information on relative strengths and weaknesses on the WAIS-R into clinical decision-making but the results suggest that this information should be gleaned from differences between the factors rather than between the IQs (i.e., VIQ vs. PIQ) or subtests.

It should also be stressed that the implications of the present results depend on the purposes of assessment. Where an important aim is to identify the presence of acquired deficits (e.g., where a medico-legal assessment is conducted for cases of personal injury litigation) interpretation of subtest differences may be best avoided. However, subtest profile analysis has a legitimate role in assessment conducted for rehabilitation purposes as the nature of a client's strengths and weaknesses at the subtest level may still have important implications for the design of remedial interventions regardless of

whether the weaknesses are, or are not, the result of any injury.

The CHI participants in the present study were recruited from amongst neurological clients referred through routine channels to four geographically diverse services. These features plus the large n (and correspondingly large n for the healthy comparison samples) suggest that the present results may be generalizable to CHI populations serviced by other centers. However, given that the WAIS-R remains a widely used instrument in neuropsychological assessment (Lezak, 1988a) and the present results cast doubt on assumptions underlying a widely endorsed approach to its clinical interpretation (e.g., Crawford, 1992; Lezak, 1988b; Walsh, 1991), we would encourage attempts at replication in clients with head injuries and other neurological disorders.

One obvious limitation of the present study was the quality of information on indicators of the severity of injury. Any future multi-center studies of this issue could attempt to achieve standardization of methods for assessing and recording PTA and thereby examine whether any of the conclusions drawn from the present findings require refinement when subgroups differing in severity of injury are examined.

REFERENCES

- Atkinson, L. (1991). Some tables for statistically based interpretation of WAIS-R factor scores. *Psychological Assessment*, 3, 288-291.
- Atkinson, L., Cyr, J. J., Doxey, N. C. S., & Vigna, C. M. (1989). Generalizability of WAIS-R factor structure within and between populations. *Journal of Clinical Psychology*, 45, 124-129.
- Baron, J., & Treiman, R. (1980). Some problems in the study of differences in cognitive processes. *Memory and Cognition*, 4, 313-321.
- Becker, B. (1975). Intellectual changes after head injury. *Journal of Clinical Psychology*, 31, 307-309.
- Bornstein, R. A. (1987). The WAIS-R in clinical practice: Boon or bust? *The Clinical Neuropsychologist*, 1, 185-190.
- Burgess, A. (1991). Profile analysis of the Wechsler Intelligence Scales: A new index of subtest scatter. *British Journal of Clinical Psychology*, 30, 257-263.
- Canavan, A. G. M., Dunn, G., & McMillan, T. M. (1986). Principal components of the WAIS-R. *British Journal of Clinical Psychology*, 25, 81-85.
- Crawford, J. R. (1992). Current and premorbid intelligence measures in neuropsychological assessment. In J. R. Crawford, D. M. Parker, & W. McKinlay, (Eds.), *A handbook of neuropsychological assessment* (pp. 21-49). London: Lawrence Erlbaum.
- Crawford, J. R., & Allan, K. M. (1994). The Mahalanobis distance index of WAIS-R subtest scatter: Psychometric properties in a healthy UK sample. *British Journal of Clinical Psychology*, 33, 65-69.
- Crawford, J. R., & Allan, K. M. (1996). WAIS-R Subtest Scatter: Base rate data from a healthy UK sample. *British Journal of Clinical Psychology*, 35, 235-247.
- Crawford, J. R., Allan, K. M., Stephen, D. W., Parker, D. M., & Besson, J. A. O. (1989). The Wechsler Adult Intelligence Scale - Revised (WAIS-R): Factor structure in a UK sample. *Personality and Individual Differences*, 10, 1209-1212.
- Crawford, J. R., Gray, C. D., & Allan, K. M. (1995). The WAIS-R (UK): Basic psychometric properties in an adult UK sample. *British Journal of Clinical Psychology*, 34, 237-250.
- Crosson, B., Greene, R. L., Roth, D. L., Fair, S. P., & Adams, R. L. (1990). WAIS-R pattern clusters after blunt-head injury. *The Clinical Neuropsychologist*, 4, 253-262.
- Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw-Hill.
- Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.
- Lea, M. (1986). *A British supplement to the manual for the Wechsler Adult Intelligence Scale - Revised*. San Antonio TX: Psychological Corporation.
- Leckliter, I. N., Matarazzo, J. D., & Silverstein, A. B. (1986). A literature review of factor analytic studies of the WAIS-R. *Journal of Clinical Psychology*, 42, 332-342.
- Lezak, M. D. (1988a). IQ: R.I.P. *Journal of Clinical and Experimental Neuropsychology*, 10, 351-361.
- Lezak, M. D. (1988b). Neuropsychological tests and assessment techniques. In F. Boiler & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 1, pp. 47-68). Amsterdam: Elsevier.
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.
- Matarazzo, J. D., Daniel, M. H., Prifitera, A., & Herman, D. O. (1988). Inter-subtest scatter in the WAIS-R standardization sample. *Journal of Clinical Psychology*, 44, 940-950.
- Matarazzo, J. D., & Herman, D. O. (1985). Clinical uses of the WAIS-R: Base rates of differences between VIQ and PIQ in the WAIS-R standardization sample. In B. Wolman (Ed.), *Handbook of intelligence: Theories, measurement and applications* (pp. 899-932). New York: Wiley.

- Maxwell, A. E. (1960). Obtaining factor scores on the Wechsler Adult Intelligence Scale. *Journal of Mental Science, 106*, 1060-1062.
- McKinlay, W. W., & Gray, J. M. (1992). Assessment of the severely head-injured. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 363-378). Hove, UK: Lawrence Erlbaum.
- McLean, J. E., Reynolds, C. R., & Kaufman, A. S. (1990). WAIS-R subtest scatter using the profile variability index. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*, 289-292.
- Reitan, R. M., & Wolfson, D. (1990). A consideration of the comparability of the WAIS and WAIS-R. *The Clinical Neuropsychologist, 4*, 80-85.
- Richardson, J. T. E. (1990). *Clinical and neuropsychological aspects of closed head injury*. London: Taylor & Francis.
- Strauss, M. E., & Allred, L. J. (1987). Measurement of differential cognitive deficits after head injury. In H. S. Levin, J. Grafman, & H. M. Eisenberg (Eds.), *Neurobehavioural recovery from head injury* (pp. 88-107). New York: Oxford University Press.
- Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper & Row.
- Walsh, K. W. (1991). *Understanding brain damage* (2nd ed.). Edinburgh: Churchill Livingstone.
- Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.
- Williams, E. J. (1959). The comparison of regression variables. *Journals of the Royal Statistical Society (Series B), 21*, 396-399.