AN EXAMINATION OF THE EFFECTIVENESS OF PREMORBID IQ INDICES AT POSTDICTING IQ SCORES IN NORMAL SUBJECTS

by

Lisa Marie Carswell
B.Sc., University of Western Ontario, 1991

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Psychology

We accept this thesis as conforming to the required standard

[redacted]

Dr.R.E.Graves, Supervisor (Department of Psychology)

[redacted]

Dr.P.Duncan, Departmental Member (Department of Psychology)

[redacted]

Dr.J.Moehr, Outside Member (School of Health Information Science)

[redacted]

Dr. J. Anderson, External Examiner, (Department of Psychological Foundations in Education)

University of Victoria

Supervisor: Dr. Roger E. Graves

ABSTRACT

The present study was conducted to (1) quantitatively evaluate the clinical sensitivity of several measures of premorbid IQ, and to (2) examine the effectiveness of premorbid IQ estimates at postdicting IQ scores in normal, elderly subjects. Selected studies were reviewed that reported the correlation between various IQ predictors (eg. National Adult Reading Test; (NART), and demographics) and Wechsler Adult Intelligence Scale/Revised (WAIS/WAIS-R) IQs for normal subjects. The error in prediction for WAIS/WAIS-R IQ was calculated for these measures as a quantitative estimate of clinical sensitivity. Results indicated, that depending on the measure, a discrepancy between obtained and predicted IQ would have to be at least 13 points (and in some cases as much as 26 points) to be statistically reliable, raising the question of the practical utility of such measures for predicting premorbid IQ. Subsequently, a regression equation was developed utilizing IQ predictors including NART errors, age-scaled WAIS-R Vocabulary scores, and demographics, to postdict WAIS-R Verbal IQ scores (VIQ) obtained more than three years earlier for a sample of 49 cognitively intact elderly normals (mean age = 71). Vocabulary, sex and NART errors contributed significantly to the postdiction of WAIS-R VIQs and accounted for 66% of the

variance in VIQ scores. The new regression equation demonstrated potential clinical usefulness as a postdictor of premorbid IQ as it accounted for 36% more of the variance in WAIS-R VIQ than the NART alone, and it produced a clinically sensitive error range of 15 IQ points.

Examiners:

[redacted]

Dr. R.E. Graves, Supervisor (Department of Psychology)

[redacted]

Dr. P. Duncan, Departmental Member (Department of Psychology)

[redacted]

Dr. J. Moehr, Outside Member (School of Health Information Science)

[redacted]

Dr. J. Anderson, External Examiner (Department of Psychological Foundations in Education)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Dr. Roger Graves for his support and guidance on this project, and would like to express my appreciation to Dr. Gary Snow and Dr. Mary Tierney at Sunnybrook Health Science Centre in Toronto, Ontario, for providing the impetus to begin this project and the subject data to follow it through to completion.

AN EXAMINATION OF THE EFFECTIVENESS OF PREMORBID IQ INDICES AT POSTDICTING IQ SCORES IN NORMAL SUBJECTS

Quantification of intellectual impairment is a critical component of a neuropsychological assessment in clinical, medico-legal and research settings. As pre-existing records are rarely available to provide a pre-injury account of an individual's cognitive ability, an accurate estimate of premorbid intelligence is needed to evaluate potential cognitive decline. A premorbid estimate of functioning is important from a neuropsychological perspective in that it not only determines the pervasiveness of an insult, but also serves to establish whether the individual has sustained losses when compared to his/her prior status (Klesges, Wilkening & Golden, 1981). Although the need to determine premorbid IQ exists for both neurological and psychiatric populations, much of the research has focused on and has developed from estimating premorbid IQ and detecting cognitive decline in brain damaged or dementing patients.

Deterioration Indices

Wechsler (1944) originally attempted to detect cognitive decline by utilizing subtests from the Wechsler-Bellevue Test (WB; Wechsler, 1939). While examining the utility of the Wechsler subtests in the diagnosis of diffuse organic brain damage, Wechsler (1944) noted that subtests that appeared most prone to age-related changes also appeared to be the most sensitive to organic impairment.

Specifically, subtests that measured sustained attention and concentration, immediate memory recall, perseverance on a task, associative learning and abstract and nonverbal concept formation were most sensitive to organic damage and the effects of aging. In contrast, subtests that measured routine learning, overlearned verbal information and general cultural information were the least likely to show decline (Wechsler, 1944). Wechsler (1944) considered that if abilities that appeared to be resistant to the effects of aging were also resistant to the effects of organic damage, then these abilities would most accurately represent a patient's premorbid level of functioning regardless of current organic involvement. Wechsler (1944) thus proposed that "mental deterioration" could be quantified by comparing performance on Wechsler subtests that were considered resistant to the effects of aging and organic impairment ("Hold" tests) to those deemed sensitive to organic damage and age-related changes ("Don't Hold" tests). Wechsler (1944) devised a deterioration quotient (DQ) also called a deterioration index (DI) which compared age-graded scores of "Hold" tests (Information, Comprehension, Object Assembly, Picture Completion) with scores of "Don't Hold" tests (Digit Span, Arithmetic, Digit Symbol, Block Design). The deterioration index was calculated by subtracting the sum of the "Don't Hold" tests from the sum of the "Hold" tests with the difference divided by the sum of the "Hold" tests.

Wechsler (1944) suggested that an individual demonstrated "possible deterioration" if s/he showed a loss above 10% and "definite deterioration" if s/he showed a loss greater than 20% than that allowed by normal decline with age.

With the development of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) which employed a larger, more representative sample than that used for standardization of the Wechsler-Bellevue, test combinations changed for "Hold" vs. "Don't Hold" subtests. Vocabulary replaced Comprehension as a "Hold" subtest and Similarities substituted for Arithmetic as a "Don't Hold" test. Detecting cognitive decline was simplified as deterioration quotients were calculated for all subjects of the standardization sample (N=2052) to avoid the need for any extrapolation or bonus for age (Wechsler, 1958). Additional tables were established to give separate scaled scores for each of the main age groups so that an individual's performance on each test could be compared with that of his/her age peers.

Research on the effectiveness of the deterioration indices has not been very encouraging with regard to detecting cognitive decline in clinical samples. While some data has suggested that the Wechsler Indices may be reasonably successful in discriminating brain damaged patients from normals (Bershoff, 1970; Reitan, 1959; Voght & Heaton, 1977), studies by Meyer (1961), Rabin (1965) and Watson, Thomas, Anderson & Felling (1968) have not supported

Wechsler's (1958) findings and have served to question the utility of the Wechsler Deterioration Index as a reliable screening procedure for patients with brain damage (Livesay, 1986). Controversy also surrounds the stability of Wechsler's four "Hold" tests (Object Assembly, Picture Completion, Vocabulary, Information) when brain damage has been sustained (Klesges et al., 1981). Allen (1948) found that Object Assembly was among the three most seriously impaired of all scales of the Wechsler-Bellevue for a sample of 50 brain damaged patients. Reitan (1959) reported significant differences in performance on the Picture Completion subtest for brain damaged and normal subjects and McFie (1975) indicated that performance on Picture Completion appeared sensitive to secondary area occipital and left parietal occipital lesions. Klesges et al., (1981) suggested that interpretation of Vocabulary and Information subtests may be confounded by education and cultural/environmental variables and Reitan (1959) reported a significant decline in Information performance for brain damaged subjects as compared to normal controls.

Although the Wechsler deterioration indices have not received good empirical support as a reliable means of detecting brain damage, they became widely known and thus introduced the concept of employing tests of present ability for use in estimating premorbid IQ. Other Wechsler subtest formulas were proposed as screening procedures for detecting patients with diffuse brain damage. Allen (1947) suggested a deterioration ratio based on Wechsler-Bellevue scores in which the sum of Information and Comprehension scaled scores was subtracted from the sum of Digit Span and Digit Symbol subtests and a difference of five points or more between the two sums indicated potential organic involvement. Gonen (1970) recommended a deterioration quotient derived by using Information and Comprehension as the "Hold" tests and Digit Symbol and Block Design as the "Don't Hold" subtests. Russell, Neuringer & Goldstein (1970), (as cited in Klesges et al., 1981) produced five Rating Equivalent Scores (RES) based on Digit Symbol scores and their deviations from a combined score of Picture Arrangement, Picture Completion and Block Design subtests. Mahan (1979) suggested comparing the WAIS Verbal and Performance scale scores as an index of impairment. Specifically, he considered that Picture Completion and Vocabulary subtests were resistant to brain damage and could be used independently to provide predicted estimates of Verbal and Performance IQ. Predicted scores could then be compared to actual obtained IQ scores to provide a residual IQ score which would serve as an estimate of impairment. Bauer, Schlottmann, Kane & Johnsen (1984) examined the utility of the Digit Symbol subtest and the RES equations proposed by Russell et al., (1970) in differentiating brain injured (both diffuse and lateralized) subjects from controls. Johnsen, Schlottmann, Kane, Bauer &

Quintana (1985) examined WAIS profiles for 66 brain injured and 50 psychiatric and medical control subjects and calculated four residual IQs using Mahan's method for each subject (Verbal Residual IQs based both on Vocabulary and Picture Completion subtest and Performance Residual IQs based on Vocabulary and Picture Completion subtests). In general, results regarding the accuracy of these modified deterioration indices in identifying brain damaged patients from controls have shown a tendency to a high number of false positives, which makes questionable the predictive reliability of these measures (Livesay, 1986).

Recently, Fuld (1984) suggested that a specific pattern of WAIS subtest scores may be useful in detecting cognitive decline for patients with Alzheimer's dementia. Fuld (1984) reported a characteristic profile of subtest scores on the WAIS associated with drug-induced cholinergic deficiencies in young adults that also differentiated between types of demented patients. Specifically, the profile is expressed by the formula, $A>B>C\leq D$, $A>D$ in which "A" is the mean of the Information and Vocabulary subtest scores, "B" of the Similarities and Digit Span subtest scores, "C" of the Digit Symbol and Block Design and "D" is the Object Assembly subtest score with all subtest scores corrected for age. Fuld (1984) reported that the same subtest profile was found in test data from two groups of dementia patients (61 and 77 patients, respectively) with diagnoses of dementia

Alzheimer's type (DAT), multi-infarct dementia (MID) and other dementias. The profile identified 44% of testable patients with DAT and was 96% specific to DAT with only two false positives. Fuld (1984) thus concluded that the subtest profile could contribute to the differentiation of DAT from other dementias.

Early attempts at cross-validation appeared to support the Fuld formula as an accurate classifier of DAT patients. Brinkman & Braun (1984) administered the WAIS to 23 DAT and 39 MID patients and found that while the Fuld profile was characteristic of 56% of DAT subjects, it was observed in less than 5% of the MID subjects. Studies with normal elderly subjects also revealed that the profile occurred rarely in independently functioning elderly persons, thus lending further support to Fuld's findings that the profile may be relatively specific to DAT (Tuokko & Crockett, 1987; Satz, Van Gorp, Soper & Mitrushina, 1987).

However, recently certain studies have questioned the diagnostic value of the Fuld formula. Filley, Kobayashi & Heaton (1987) examined the diagnostic utility of the Fuld profile for 41 patients with DAT. Filley et al., (1987) reported that only nine (21.9%) of these patients had positive Wechsler profiles and suggested that the diagnostic value of a positive profile was modest. Logsdon, Teri, Williams, Vitiello & Prinz (1989) examined the Fuld profile for a sample of 173 elderly research volunteers and found that the profile occurred in 7% of DAT subjects, 10% of subjects complaining of memory loss who did not meet the diagnostic criteria for DAT, 7% of nondemented major depressed subjects and 7% of normal control subjects. Logsdon et al., (1989) concluded that the Fuld profile was not useful in differentiating DAT from other common disorders among elderly subjects. Further, Gfeller & Rankin (1991) indicated that the Fuld profile was neither sensitive nor specific to DAT for a sample of 16 DAT subjects and 31 MID subjects as similar numbers of DAT subjects (15%) and MID subjects (12%) obtained a positive Fuld profile.

Estimation of Premorbid IQ

The various Wechsler subtest deterioration formulas reviewed above were designed to provide a means for detecting the presence of a decline in cognitive ability. That is, a means for determining that the present IQ is lower than premorbid IQ. A more sophisticated approach is to attempt to estimate premorbid IQ using results from postmorbid tests. Crawford (1989) indicated that for a present ability measure to qualify as a valid means of estimating premorbid IQ it must have adequate reliability, it must correlate highly with IQ in a normal population and it must be resistant to the effects of neurological and psychiatric disorder.

Wechsler Vocabulary Subtest

The Vocabulary subtest of the Wechsler scales has probably been the most commonly used test of present ability to estimate premorbid IQ (Crawford, 1989). With regard to reliability as a testing measure, the Vocabulary subtest has demonstrated the highest split-half reliability of all WAIS subtests with reliability coefficients ranging between .94 and .96 for the three age bands examined by Wechsler (1955). Matarazzo, Carmody & Jacobs (1980) examined the test-retest reliability of WAIS subtests and, of the articles reviewed, determined that the Vocabulary subtest had the second highest test-retest correlation of the 11 subtests, with a median correlation coefficient of .85. Similarly, Vocabulary has also demonstrated the highest reliability of all the Wechsler Adult Intelligence Scale - Revised (WAIS-R; Wechsler, 1981) subtests. The averaged split-half reliability across all age bands in the WAIS-R standardization sample was .96 and test-retest correlation coefficients across the two age bands were .91 and .93, respectively.

As a measure of intelligence, Vocabulary has been identified as the single best measure of both verbal and general mental ability (Lezak, 1983). Vocabulary has been found to correlate very highly with both WAIS and WAIS-R Full Scale IQ (FSIQ). Wechsler (1955) reported that Vocabulary accounted for approximately 74-76% of the variance in WAIS FSIQ for the three age bands in the WAIS standardization sample. Similarly, Wechsler (1981) found that Vocabulary accounted for 72% of the variance in WAIS-R FSIQ when data from the full standardization sample (N = 1,880) were analyzed. Factor analytic studies of the Wechsler scales using the standardization data have demonstrated that Vocabulary loads very highly on the first unrotated factor considered to represent general intelligence (g) (Matarazzo, 1972). Crawford, Cochrane, Besson, Parker & Stewart (1990) recently found that Vocabulary loaded .88 on the WAIS first factor considered to represent general intelligence (g).

Clinically, Vocabulary has not proven to be particularly resistant to neurological or psychiatric dysfunction. Although Vocabulary was considered to be amongst the most resistant of WAIS subtests in clinical conditions, early studies revealed that Vocabulary performance of neurological or psychiatric patients was significantly lower than healthy controls (Crawford, 1989). However, Crawford (1989) cautioned that many of these studies did not take into account premorbid demographic variables known to be related to IQ performance (education, age, socio-economic level) and thus failed to adequately match clinical and control groups on demographic variables.

Russell (1972) calculated a biserial correlation coefficient between the presence or absence of brain damage and Vocabulary performance in a sample of 103 patients. A highly significant correlation was obtained which suggested that Vocabulary performance was severely impaired by brain damage. However, Crawford (1989) indicated that the results of this study were compromised by the fact that Russell (1972) included congenitally brain damaged subjects as part of the clinical sample. Such subjects served little value in assessing Vocabulary to the resistance of adult-onset disorder as these subjects did not experience normal intellectual development. As well, the subjects were not well matched for education as subjects in the control group were of a significantly higher level of education even when the educational level of the congenital cases was removed from the analysis. Vogt & Heaton (1977) investigated the efficacy of Wechsler's "Hold - Don't Hold" index and divided clinical subjects into "impaired" and "nonimpaired" groups on the basis of their performance on the Halstead Impairment Index. Vogt et al., (1977) compared the Vocabulary performance of these two groups and found a significant difference in performance in favour of the "nonimpaired" group. However, further analysis revealed that much of the difference was due to a between group difference in educational level.

More recently, Nelson & McKenna (1975) built a regression equation to predict (prorated) WAIS IQ from Vocabulary age-graded scaled scores using a standardization sample of 98 subjects free of neurological disorder. Nelson et al., (1975) compared the Vocabulary estimated IQs of the neurologically normal sample to the Vocabulary estimated IQs of 45 dementing subjects and found that the Vocabulary estimated IQs for the dementing subjects were significantly lower than healthy controls. However, demographic data for the two groups were not presented and the reliability of this measure is questionable as prorating of WAIS IQs has been regarded as inadvisable (Wechsler, 1955). Hart, Smith & Swash (1986) used Nelson & McKenna's (1975) equation to compare the Vocabulary performance of 20 patients with dementia of the Alzheimer's type (DAT) to a control group of 15 normal elderly of comparable age and education. Hart et al., (1986) reported that the DAT patients obtained significantly lower Vocabulary estimated IQs than controls. Crawford, Besson & Parker (1988) further examined the effect of organic impairment upon Vocabulary performance. They found that Vocabulary estimated IQs (Nelson et al., 1975) were significantly lower for groups of patients with DAT, multi-infarct dementia (MID), alcoholic dementia, Korsakoff's dementia and Huntington's disease when compared to their respective controls that were individually matched for age, sex and years of education. Using the same design, Crawford, Besson, Parker, Sutherland & Keen (1987) compared the Vocabulary performance of depressed patients and controls. Crawford et al., (1987) found that the Vocabulary performance of the depressed group was significantly poorer than controls and concluded that impairment on Vocabulary could occur even in the absence of organic pathology.

Although various studies have suggested that the Vocabulary subtest of the WAIS is sensitive to the effects of organic damage across certain clinical conditions, other studies have found that Vocabulary performance is intact in organically impaired patients. Crawford et al., (1988) indicated that there was no difference in Vocabulary performance for patients with closed head injuries and controls matched for age, sex and education. Sullivan, Sagar, Gabrieli, Corkin & Growdon (1989) examined the performance of "demented" Parkinson's disease (PD) patients and "nondemented" PD patients on the Vocabulary subtest of the WAIS-R and found that the "demented" PD group performed equally to the healthy control group, whereas the "nondemented" PD group performed significantly better than the control group. Boyd et al., (1991) compared the performance of PD patients on a battery of tests of cognition, motor function, disability and mood with the performance of healthy controls who were matched for age, sex and premorbid IQ. Boyd et al., (1991) reported that there was no difference in performance between PD patients and controls on the Vocabulary subtest and Digit Span subtest of the WAIS.

Other studies have also examined the extent that

Vocabulary performance may deteriorate in clinical conditions. Martin & Fedio (1983) examined the performance of relatively mildly impaired patients with Alzheimer's disease (AD) on measures of word comprehension, word naming and fluency, and semantic knowledge. Semantic knowledge was evaluated by the Vocabulary and Similarities subtests of the WAIS. Martin et al., (1983) reported that although AD patients achieved significantly lower scores on both the Vocabulary and Similarities subtests as compared to controls matched for age and education, the scores for the AD patients were within the average range for both Vocabulary and Similarities (mean age scaled score = 10.9, S.D. = 2.0 for Vocabulary, mean age scaled score = 10.3, S.D. = 2.1 for Similarities). Martin et al., (1983) also found that the AD group performed at a higher level on both the Vocabulary and Similarities subtests relative to all other WAIS subtests. Crookes (1974) and Whitehead (1973) also reported that of the 11 WAIS subtests, dementing patients achieved the highest score on the Vocabulary subtest.

As Vocabulary has not been found to be consistently resistant to the effects of cerebral dysfunction in clinical samples, more recent attempts at estimating premorbid IQ have included: (1) regression equations that use demographic variables such as education and occupational status which have a strong relationship with measured IQ (Matarazzo, 1972) and (2) tests of overlearned skills such as reading, which are highly correlated with intelligence and are relatively resistant to organic impairment (Grober & Sliwinski, 1991).

Demographic Predictors

Blair & Spreen (1989) reported that the first comprehensive effort to predict premorbid IQ on the basis of demographic variables was made by Wilson et al., (1978). Using the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) standardization sample, Wilson et al., (1978) developed demographically based actuarial prediction equations for IQ by stepwise regression of WAIS Verbal (VIQ), Performance (PIQ) and Full Scale IQ (FSIQ) on variables of age, sex, race, occupation and education. The amount of variance accounted for between all five demographic variables and VIQ, PIQ and FSIQ was 53%, 42% and 54%, respectively. Wilson et al., (1978) indicated that education and race were the most powerful predictors in each regression equation. Education, on its own, accounted for 44%, 31% and 43% of the variance in VIQ, PIQ and FSIQ, respectively.

Several attempts at cross-validation have produced variable results. Initial studies appeared to support the use of demographic equations. Wilson, Rosenbaum & Brown (1979) compared the performance of 140 neurologic and 140 non-neurologic patients and found that the demographic regression equation correctly classified 11% more of the patients than did the WAIS deterioration quotient. Klesges, Sanchez & Stanton (1981) examined the correlations between demographically predicted IQ and obtained WAIS IQ in two neurologically unimpaired samples (60 psychiatric inpatients and 106 outpatients). Klesges et al., (1981) reported highly significant correlations between predicted and obtained IQs in both groups and also determined that the correction for education as proposed by Wilson et al., (1978) significantly reduced the number of scores that were overpredicted by the original equation.

Further studies have revealed mixed results. Bolter, Gouvier, Veneklasen & Long (1982) examined the correlation between predicted and obtained FSIQ in head injured patients at two testing periods. Patients were divided into two groups based on neuropsychological test results at the second testing which indicated that they had recovered (N=11) or were impaired (N=11). A control group of 24 "pseudoneurological" patients was also examined. Although Bolter et al., (1982) found that both adjusted and unadjusted FSIQ predictions correlated significantly with IQs obtained at the second testing (r= .63 to .73), they reported that the demographic estimate misclassified the obtained IQ in 50% of the patients. A misclassification was considered to have occurred when the predicted IQ differed from the obtained IQ by more than one standard error of estimate (SEE). This discrepancy led Bolter et al., (1982) to conclude that use of the demographic regression equation may be inappropriate in clinical samples.

A second study by Gouvier, Bolter, Veneklasen & Long (1983) evaluated the same subjects used in the Bolter et al., (1982) study and used the same protocol, but reported predicted and obtained VIQ and PIQ rather than FSIQ. Similar results were obtained as the correlations between predicted and obtained IQs ranged from .16 to .75 however, the proportion of correct classification of patient groups was still poor.

Klesges, Fisher, Vasey & Pheley (1985) used the Wilson et al., (1978) formulas (adjusted and unadjusted) to compare the actual and predicted IQs of 125 brain injured and 73 normal subjects whose neurological status had been established by CAT scan and neurological exam (Klesges & Troster, 1987). The results of the study indicated significant but low correlations between predicted and actual IQs in the normal group. In addition, the predicted premorbid IQs failed to discriminate reliably between groups of normal and brain damaged subjects (Blair & Spreen, 1989). Klesges et al., (1985) concluded that the Wilson et al., (1978) formula should be used with caution in clinical samples and should be limited to only research purposes.

Crawford (1989) has suggested that major methodological problems exist with the studies reviewed above and argues that concern about the validity of the demographic approach in estimating premorbid IQ may be premature. Crawford (1989) indicates that any attempt to cross-validate a measure of premorbid IQ should: (1) have an adequate sample size as misleadingly high or low correlations could be obtained by chance (2) ensure that the sample reflect as much as possible, the full range in predictor and criterion variables observed in the general population or low estimates of the population correlation between predicted and obtained IQ will result and (3) stress that the purpose of demographic equations is to estimate premorbid IQ and not current IQ. Thus, cross-validation should be carried out with healthy, normal subjects as it is impossible to compare estimates of premorbid IQ with actual premorbid IQ in clinical samples, unless previous records exist.

Recently, Karzmark, Heaton, Grant & Matthews (1985) compared demographically estimated IQ with obtained IQ in a sample of 491 normal subjects. The Wilson et al., (1978) formula yielded similar mean predicted and mean obtained FSIQs (110.9 vs. 112.8), respectively. Predictive accuracy was relatively stable across different age, education and occupational levels. However, the formula tended to underestimate FSIQ at lower IQ levels and overestimate FSIQ at higher IQ levels. Educational adjustment actually decreased predictive accuracy. Kazmark et al., (1985) indicated that one plausible explanation for this discrepant finding may have been that half of the sample had been tested before 1975 and thus changes in the inter-relationship between predictor variables and IQ may have occurred.

In a further cross-validation study, Goldstein, Gary & Levin (1986) examined the accuracy of the Wilson et al., (1978) regression equations for estimating premorbid IQ in a sample of 69 neurologically normal adults. The Wilson et al., (1978) formula yielded similar mean predicted and mean obtained FSIQs (112.0 vs. 109.2), and accounted for 50% and 37% of the variance in WAIS VIQ and PIQ, respectively. Although Goldstein et al., (1986) concluded that the equations provided an adequate overall fit to the data, they indicated that obtained IQs at the extremes of the WAIS scales were found to be most susceptible to underestimation or overestimation by predictive measures.

With the introduction of the Wechsler Adult Intelligence Scale - Revised (WAIS-R), the Wilson et al., (1978) demographic equations became clinically dated. Barona, Reynolds & Chastain developed prediction equations on the basis of the WAIS-R in 1984. In addition to the original demographic variables utilized by Wilson et al., (1978), Barona et al., (1984) originally included geographic region, urban vs. rural residence and handedness as demographic predictor variables. The variable of handedness was not included in the final equation model as it was not found to contribute significantly to VIQ, PIQ or FSIQ. The variance accounted for between the final demographic variables and WAIS-R VIQ, PIQ and FSIQ was 38%, 24% and 36%, respectively. Barona et al., (1984) indicated that education, race and occupation were the most powerful predictors in each equation. Education accounted for 31%, 17% and 28% of the variance in VIQ, PIQ and FSIQ, respectively, whereas race accounted for 8%, 8% and 9% of the variance in VIQ, PIQ and FSIQ, respectively. Occupation explained 14%, 9% and 13% of the variance in VIQ, PIQ and FSIQ, respectively.

Attempts at cross-validation have also produced variable results. Eppinger, Craig, Adams & Parsons (1987) cross-validated the demographic formulas proposed by Barona et al., (1984) on a neurologically normal population and examined the utility of these demographic equations to discriminate between normal and brain impaired individuals. Eppinger et al., (1987) found that for normal subjects, demographic variables accounted for 61%, 36% and 58% of the variance in WAIS-R VIQ, PIQ and FSIQ, respectively. Predictive accuracy further increased to 75%, 71% and 69% for VIQ, PIQ and FSIQ respectively if obtained IQ fell within one standard error of estimate (SEE) of the regression equation. However, Eppinger et al., (1987) cautioned that the demographic equations tended to overestimate IQs of normal subjects.

Eppinger et al., (1987) also examined the ability of the demographic equations to discriminate between normal and brain impaired individuals. They found that while obtained WAIS-R IQs were significantly lower for neurologically impaired subjects, there were no significant differences for equation estimated IQs between groups. Thus, demographic equations sufficiently revealed deterioration from a higher functioning premorbid level in neurologically impaired subjects.

Ryan & Prifitera (1987), (as cited in Paolo & Ryan, 1992) cross-validated the demographic FSIQ equation with 118 normal subjects. They found that it overestimated functioning in persons with obtained FSIQs of less than 89 and underestimated the ability levels of individuals with FSIQ of more than 110.

Ryan & Prifitera (1990) also examined the ability of the Barona Index to predict short form IQs. Two hundred and twenty-five normals were administered the WAIS-R Arithmetic, Vocabulary, Picture Completion and Block Design subtests. Ryan et al., (1990) found that the equation based estimates were significantly correlated with short form IQs, but accounted for a shared variance of only 29%. Further, low IQs were typically overestimated and estimates of high IQ were inaccurate. Ryan et al., (1990) suggested that the Barona Index may be more accurate in predicting IQ based on full administration of the WAIS-R rather than short form administration.

In 1986, Barona & Chastain attempted to improve the accuracy of the original equations and performed regression analyses on the WAIS-R standardization sample utilizing only black and white subjects between 20-74 years of age (Paolo & Ryan, 1992). Races classified as "other" were omitted as were persons 16-19 years of age as their occupational classification had been based on subject's head of household. These modifications appeared to improve the predictive accuracy of the equation as the amount of variance in WAIS-R scores accounted for by demographic variables increased to 47%, 28% and 43% for VIQ, PIQ and FSIQ, respectively (Barona & Chastain, 1986). Once again, education, race and occupation were the most powerful predictors in each equation. Education accounted for 36%, 20% and 33% of the variance in VIQ, PIQ and FSIQ, respectively.

The 1986 equations have not been cross-validated and Paolo & Ryan (1992) recently determined that the 1986 equation may not be as valuable as originally considered when compared to the predictive accuracy of the 1984 equation in estimating WAIS-R IQs in an elderly sample. Paolo & Ryan (1992) compared the 1984 and 1986 Barona equations in a sample of 75 normal elderly and 20 neurologically impaired elderly. In the normal sample, the amount of variance accounted for between actual WAIS-R IQs and estimated IQs were moderate and similar for both estimate procedures. Paolo & Ryan (1992) further evaluated the accuracy of the predicted IQs by calculating the percentage of obtained IQs which fell within one standard error of the estimated IQs. A minimum acceptable level was arbitrarily set at an accuracy of 70%. For normal subjects, VIQ was reasonably predicted by the 1984 method (73.3%) while the 1986 formula failed to meet the minimum criteria (64.0%). Both prediction equations correctly estimated FSIQ within one SSE (1984 = 77.3%),(1986 = 72.0%), while both equations failed to adequately estimate PIQ. Both estimation procedures also tended to underestimate all three actual WAIS-R IQs. Based on results of predictive accuracy, Paolo & Ryan (1992) suggested that for normal elderly, the 1984 equation is slightly better in estimating WAIS-R IQs within one SSE than the 1986 method.

Paolo & Ryan (1992) also examined the results of the neurologically impaired elderly and found that both 1984 and 1986 estimated IQs were significantly higher than obtained WAIS-R IQs. As well, although the normal and neurological groups differed on demographic variables of age and education, no differences were found between groups on formula estimated IQs. Thus, the difference in estimated and obtained WAIS-R IQs for the neurological group as compared to the normal group would suggest the detection of deterioration in cognitive functioning.

National Adult Reading Test

Tests of overlearned skills are also being used to estimate premorbid IQ. The National Adult Reading Test (NART; Nelson, 1982) is one such measure of reading skills which is purported to be a useful measure of premorbid IQ. Word reading as an estimate of premorbid intelligence is based on the assumption that reading ability is highly correlated with IQ in the normal population and is preserved in the presence of dementing processes (Nelson, 1982). The NART is a single-word oral reading test which consists of 50 irregular words that do not follow normal grapheme-phoneme correspondence rules of the English language (Crawford, 1989). It requires an individual to correctly pronounce atypical words and in doing so measures previous familiarity with such words independent of present ability to analyze them as a complex visual stimulus (Wiens, Bryan & Crossesn, 1993).

To be effective as a predictive measure of intelligence the NART must demonstrate validity as a measure of IQ and reliability as a testing instrument. Crawford, Stewart, Cochrane, Parker & Besson (1989) examined the construct validity of the NART by principal component analysis and found that it loaded .85 on a WAIS first factor considered to represent general intelligence (g). The NART also has an extremely high test-retest reliability as Crawford, Parker, Stewart, Besson & De Lacey (1989) found that NART errors accounted for 96% of the variance in performance at test and retest. O'Carroll (1987) indicated that the NART has a relatively high degree of inter-rater reliability as Pearson product moment correlations between raters ranged between .89 - .98. Similarly, Crawford et al., (1989) also concluded that the NART has a high degree of inter-rater reliability as Pearson correlations between all possible pairs of experienced raters ranged between .96 - .98.

To be considered a valid measure of intelligence the NART must also correlate highly with IQ in a normal population. The original standardization sample consisted of 120 neurologically normal subjects who were administered seven subtests of the WAIS including Arithmetic, Similarities, Digit Span, Vocabulary, Picture Completion, Block Design and Picture Arrangement (Nelson, 1982). The results of the seven subtests were then used to prorate VIQ, PIQ and FSIQ for each subject. Upon administration, Nelson (1982) found that the NART accounted for 55%, 59% and 32% of the variance in WAIS FSIQ, VIQ and PIQ, respectively. However, the reliability of this result is questionable as prorating of WAIS IQs has been regarded as inadvisable (Wechsler, 1955). Nelson (1982) also reported that NART appears resistant to the normal effects of aging as NART errors did not correlate significantly with age in the standardisation sample. More recently, Starr, Whalley, Inch & Shering (1992) examined the cognitive function of 598 healthy elderly subjects and found that there was no significant correlation between age and NART-predicted IQs.

A number of studies have examined the validity of the NART as a predictive measure of intelligence in normal populations with varying results. Crawford, Parker, Stewart, Besson & De Lacey (1989) found that for 151 normals, the NART accounted for 66%, 72% and 33% of the variance in WAIS FSIQ, VIQ and PIQ, respectively. Willshire, Kinsella & Prior (1991) administered the NART to 104 normal subjects and found that the NART accounted for only 26% of the variance in WAIS-R FSIQs. Willshire et al., (1991) then separately analyzed the results for 28 subjects aged 55 and over and reported that the NART accounted for 53% of the variance in WAIS-R FSIQs. Sharpe & O'Carroll (1991) indicated that for a normal sample of 20 subjects, the NART predicted 60% and 66% of the variance in WAIS-R FSIQ and VIQ, respectively. Ryan & Paolo (1992) found that for 85 normals, the NART accounted for 55%, 61% and 32% of the variance in WAIS-R FSIQ, VIQ and PIQ, respectively.

Clinically, more controversy surrounds the NART as a predictor of premorbid IQ. At question is the NART's ability to provide an accurate estimate of premorbid level of functioning in neurologic and psychiatric populations. To determine if the NART remains a valid estimate of premorbid IQ in a clinical population, performance on the WAIS/WAIS-R and NART of clinical patients is compared to the performance of normal controls on the same measures. If there is: (1) evidence of impaired cognitive ability in that WAIS/WAIS-R scores are considerably lower in the clinical population as compared to controls (2) NART predicted IQ is significantly higher than WAIS/WAIS-R obtained IQ in the clinical sample, and (3) there is no significant difference in NART scores between patients and controls, then the NART is considered to be resistant to the effects of organic impairment and effective in providing an accurate measure of premorbid IQ. If however, in the presence of impaired cognitive ability, performance on the NART is significantly lower in the clinical population than the normal control group, the NART is deemed sensitive to the effects of organic or psychiatric impairment and is not considered to be a valid estimate of premorbid IQ in such a clinical sample. With the assumption of impaired cognitive ability taken into account in clinical samples, performance on the NART as compared to normal controls is one way to determine if the NART is resistant to organic/psychiatric impairment and is a valid measure of premorbid IQ.

A number of studies have examined the validity of the NART as a predictor of premorbid IQ in neurologically and psychiatrically compromised populations with conflicting results. Neurologically, various populations have been investigated including patients with Alcoholic Korsakoff's syndrome, Alzheimer's dementia, multi-infarct dementia, Parkinson's disease, Huntington's disease, head trauma and other dementing processes. Korsakoff's patients generally performed significantly poorer on the NART as compared to normal controls which suggests that the validity of the NART is questionable for this population (Crawford, Parker & Besson, 1988; O'Carroll, Moffoot, Ebmeier & Goodwin, 1992). Various studies have indicated that there is no difference in NART performance between Alzheimer's patients and normal controls (Nebes, Martin & Horn 1984; O'Carroll & Gilleard, 1986; Crawford, Parker & Besson, 1988; Sharpe & O'Carroll, 1991), whereas other studies have found that NART performance in Alzheimer's patients is significantly poorer than that of controls (Fromm, Holland, Nebes & Oakely, 1991; Brayne & Beardsall, 1990; Hart, Smith & Swash, 1986), becomes progressively poorer over time (Fromm et al., 1991), and is negatively affected by the severity of the dementing process and any language problems associated with the disorder (Stebbins, Wilson, Gilley, Bernard & Fox, 1990; Stebbins, Gilley, Wilson, Bernard & Fox, 1990). Boyd et al., (1991) found no significant difference in NART performance between Parkinson's patients and normal controls. Crawford, Parker & Besson (1988) indicated that Huntington's patients performed significantly more poorly on the NART than did controls although patients with multi-infarct dementia and head injured patients did not differ from normal controls on NART performance.

The psychiatric populations that have been investigated have included schizophrenic, psychotic and depressed patients. O'Carroll, Walker et al., (1992) found that there were no differences in NART performance for acute schizophrenics and acute psychotics when compared to normal controls. However, Young et al., (1991) found that schizophrenics demonstrated significantly lower NART scores than controls. Crawford et al., (1992) also found that NART scores were significantly lower for schizophrenic patients in a long stay ward, but found no significant difference in NART performance for schizophrenics who were community residents and normal controls. Crawford, Besson, Parker, Sutherland & Keen (1987) found that there was no significant difference in NART performance for depressed patients as compared to normal controls. Thus, as with the neurological population, the validity of the NART seems variable both within and between psychiatric conditions.

As the original NART regression equations for the prediction of WAIS IQs were developed from prorated WAIS IQ scores and a relatively small standardization sample (N=120), efforts have been made to update these equations based on larger sample sizes and administration of the full WAIS. Crawford, Parker, Stewart, Besson & De Lacey (1989) administered the full WAIS to a sample of 151 neurologically normal subjects. Crawford et al., (1989) then combined the results of their sample (N=151) with the original NART standardization sample (N=120) and generated new regression equations to predict WAIS IQ from NART error scores. Crawford et al., (1989) indicated that these equations predicted 57%, 63% and 37% of the variance in WAIS FSIQ, VIQ and PIQ, respectively and considered that they should be used in preference to the original NART equations as they were based on a larger sample size with a wider IQ and age range.

As the NART regression equations were designed for prediction of 1955 WAIS IQs, the original equations have become somewhat outdated with the publication of the WAIS-R. Recently, Ryan & Paolo (1992) generated regression equations to predict WAIS-R IQs from NART error scores using a development sample of 85 normal elderly subjects. Cross-validation with 41 normal elderly subjects indicated that the new NART equations accounted for 69%, 16% and 55% of the variance in WAIS-R VIQ, PIQ and FSIQ, respectively. Since the NART equations demonstrated relatively good predictive ability of IQs within a normal cross-validation sample, Ryan & Paolo (1992) combined the development and cross-validation groups and generated new equations based on the results of all 126 normal elderly subjects. They further assessed the clinical utility of the new equations by applying them to a sample of 20 brain damaged patients. Results indicated that obtained IQs were significantly lower than NART estimated IQs and thus, Ryan et al., (1992) concluded that the NART estimated IQs adequately demonstrated intellectual deterioration in a brain damaged sample.

Revisions have also been developed to improve upon limitations of the original British form of the NART for North American populations (Wiens et al., 1993). Blair & Spreen (1989) standardized a revised word list (NART-R) with a sample of 66 unimpaired American and Canadian subjects. Blair & Spreen (1989) indicated that the NART-R accounted for 69%, 16% and 56% of the variance in WAIS-R VIQ, PIQ and FSIQ, respectively. Efforts to cross-validate these results with normal subjects have not been as successful as the original standardization sample. Wiens et al., (1993) found that for 302 normal subjects, the NART-R accounted for only 21%, 31% and .05% of the variance in WAIS-R FSIQ, VIQ and PIQ, respectively.

Efforts have been made to enhance the potential predictive value of the NART by combining it with demographic variables. Crawford, Stewart, Parker, Besson & Cochrane (1989) developed a regression equation based on the combination of the NART with demographic variables (sex, social class, age) and found that the combined equation accounted for 73%, 78% and 39% of the variance in WAIS FSIQ, VIQ and PIQ. Further, the combined equation accounted for more of the variance in WAIS IQ scores than the NART or demographics alone. Crawford, Nelson, Blackmore, Cochrane & Allan (1990) attempted to replicate this finding with the original NART standardisation sample and also found that combining demographics with the NART significantly improved predicted variance over use of the NART alone. Crawford, Cochrane, Besson, Parker & Stewart (1990) examined the construct validity of the NART/demographic equation (NDE) by principal component analysis and found that it loaded .90 on a WAIS first factor considered to represent general intelligence (g). They concluded that the NDE was superior to the NART alone as a measure of premorbid (g).

Recently, Willshire et al., (1991) developed a regression equation based on the combination of NART errors and demographic variables (education and sex) to predict (prorated) WAIS-R IQs. The regression equation was generated from an Australian sample of 49 normal elderly subjects aged 55 - 69 years. Willshire et al., (1991) indicated that the NART/demographic equation predicted 56% of the variance in WAIS-R FSIQs and accounted for more of the variance in WAIS-R FSIQs than the NART or demographics alone. Willshire et al., (1991) cross-validated the equation with a sample of 104 normal controls aged 20 - 69 years and found that the NART/demographic equation accounted for 46% of the variance in WAIS-R FSIQs.

As both the NART and Vocabulary subtest are measures of present ability and have been used as estimates of premorbid IQ, various studies have compared the predictive validity of each of these measures. Crawford, Besson, Parker, Sutherland

& Keen (1987) compared the performance of a group of depressed patients on the NART and the Vocabulary subtest of the WAIS with controls matched for sex, age and education. Vocabulary age-graded scaled scores were converted to estimated FSIQs by the regression equation developed by Nelson & McKenna (1975). Crawford et al., (1987) reported that the Vocabulary performance of the depressed group was significantly poorer than controls, but there was no significant difference in NART performance for depressed patients and controls. Crawford et al., (1987) suggested that based on the results, the NART may be more resistant to the effects of depression and may therefore provide a more accurate estimate of premorbid IQ for this patient group. Crawford, Parker & Besson (1988) also examined the validity of the NART and Vocabulary as measures of premorbid IQ for patient populations including Korsakoff's psychosis, alcoholic dementia, dementia Alzheimer's type, multi-infarct dementia, Huntington's disease and closed head injury. Each patient group was matched with healthy controls for sex, age and education. Nelson & McKenna's (1975) regression equation was used to convert Vocabulary age-graded scaled scores to estimated WAIS FSIQs. Crawford et al., (1988) found that there was no significant difference in NART performance for controls and patient groups including alcoholic dementia, dementia Alzheimer's type, multi-infarct dementia and closed head injury. Vocabulary performance was significantly poorer for all patient groups, except the closed head injury group, when compared to controls. Based on the results, Crawford et al., (1988) concluded that the NART provided the more accurate estimate of premorbid IQ in organic conditions. However, it should be noted that the sample size was quite small for each clinical condition. Patients size ranged between 6 - 18 patients/group with the closed head injury group (N=18) as the largest clinical group in the study.

Limitations of Existing Premorbid IQ Indices

Although the approaches described above have been the most commonly used measures for estimating premorbid IQ in clinical settings, these approaches are constrained by serious limitations. Not only are these measures controversial at the clinical level as accurate estimators of premorbid level of functioning in neurologic and psychiatric populations, but more importantly, relatively few studies have examined the ability of these measures to predict IQ scores for normal subjects. As Crawford (1989) indicated initially, a measure of premorbid IQ must first correlate highly with IQ in a normal population before it can be used to predict IQ in a clinical sample. Currently, the most popular standard measure of intellectual functioning is the WAIS-R. Thus, for a measure to accurately estimate premorbid IQ, it must correlate highly with WAIS-R IQs in a normal sample. To date, correlations have ranged from .51 (NART; Willshire et al., 1991) to .85 (Vocabulary;

Wechsler, 1981) between predictive measures of IQ and WAIS-R FSIQs. Complicating matters further is the fact that some of the correlational studies have used the WAIS (Wechsler, 1955) as the standard measure of intellectual functioning, while others have used the updated WAIS-R (Wechsler, 1981).

Another limitation with previous studies that have examined the relationship between predictive IQ measures and the WAIS/WAIS-R in normal samples, is that both tests have been administered at the same time. Correlations between tests of similar abilities would be expected to be maximized if the two measures were administered at the same time, but somewhat lower if the measures were administered at different times. Thus, most of the studies examined so far would tend to overestimate the utility of predictive IQ measures as indices of premorbid intelligence. In other words, the critical test for a predictive IQ measure is not how well it predicts current IQ, but how well it postdicts IQ scores at some point in the past.

Recently, the author was involved in a study in which the postdictive ability of the NART was assessed for a sample of normal, elderly subjects. Carswell, Snow & Tierney (1993) administered the NART to a sample of 49 cognitively intact elderly subjects who had been administered the WAIS-R five years earlier, to determine if the NART could accurately postdict IQ in a normal population. Using the original NART regression equations (Nelson, 1982), Carswell et al., (1993) found that the NART only accounted for 28%, 30% and 14% of the variance (r = .53, .54, .37) in WAIS-R FSIQ, VIQ and PIQ, respectively. Based on the poor performance of the NART over time in a normal population, Carswell et al., (1993) suggested the need for caution in using the NART to estimate premorbid IQ in clinical samples.

Berry et al., (1993) recently examined the relationship between NART-R scores and WAIS-R scores derived from testing administered 3.5 years earlier in a normal elderly sample. Berry et al., (1993) reported that the NART-R accounted for 49% of the variance (r = .70) in WAIS-R FSIQ. Based on the results, Berry et al., (1993) suggested that the relationship between the NART-R and WAIS-R FSIQ demonstrated moderate stability over time.

Although the Berry et al., (1993) study reported a higher correlation between FSIQ and NART score than the Carswell et al., (1993) study, various factors require consideration when evaluating the postdictive ability of the NART. Initially, the Carswell et al., (1993) study used the NART while the Berry et al., (1993) study used the NART-R. It should also be noted that the time interval between WAIS-R administration and NART administration was somewhat greater in the Carswell et al., (1993) study, a factor which would be expected to reduce the correlation still further. The subject sample was also somewhat older in the Carswell et al., (1993) study than the Berry et al., (1993) study (71 years of age vs. 67.8 years of age) and had higher WAIS-R FSIQ scores (FSIQ = 116.4 vs. 111.6), although the two samples were similar in terms of education (14.6 vs. 14.8 years of education, respectively).

Quantitative Evaluation of the Sensitivity of Premorbid IQ Estimates

In providing an estimate of premorbid IQ for a clinical sample, impairment in cognitive performance is inferred if a patient's results reveal a positive discrepancy between predicted and obtained IQ scores. However, for such a discrepancy to be considered significant, it must be compared to the standard range or variation in scores obtained by a normal population. To date, variations in scores have largely been arbitrarily expressed in terms of the percentage of normal subjects that achieved a certain discrepancy score. Impairment has subsequently been estimated by comparing the discrepancy score of a clinical subject to the percentage of normal subjects that achieved the same discrepancy score. Nelson (1982) indicated that a discrepancy score of 15 points or more between NART estimated FSIQ and WAIS FSIQ was strongly suggestive of intellectual deterioration as less than 2% of the NART standardization sample demonstrated a 15+ point FSIQ discrepancy score. Similarly, Crawford, Stewart, Parker, Besson & Cochrane (1989) indicated that a discrepancy score of 15 points or more between NDE estimated FSIQ and WAIS

FSIQ suggested the presence of intellectual impairment as less than 1% of the NDE standardization sample demonstrated a 15+ point FSIQ discrepancy score.

However, reference to an arbitrary range of discrepancy scores for normals may not be an accurate means of detecting cognitive impairment. Recently, Eppinger et al., (1987) calculated discrepancy scores between demographic predicted IQs (Barona et al., 1984) and WAIS-R IQs for all subjects in a neurologically normal sample. Eppinger et al., (1987) proceeded to evaluate the accuracy of a 10 - 15 point discrepancy range in correctly classifying brain-impaired patients from normal controls. The range was selected as clinically relevant in that it provided a balance between false positives and negatives, and maximized discrimination between normal and brain-impaired groups. Eppinger et al., (1987) reported that discrepancy scores of 12, 14 and 12 IQ points respectively, generated the greatest number of correct patient classifications for VIQ, PIQ and FSIQ, respectively. However, Eppinger et al., (1987) indicated that discrepancy scores were not significantly different from IQ scores in terms of discriminating between patient groups and thus concluded that the diagnostic accuracy of the discrepancy score was not supported by patient classification results.

Therefore, it would appear that to accurately infer cognitive impairment, a patient's results must be compared to a <u>standard range</u> in scores obtained by a normal population. To date, only one study has attempted to estimate a standard range in scores for a normal population. Blair & Spreen (1989) determined a standard range in scores between the NART-R and WAIS-R for normals based on standard errors of estimate at a 95% confidence level. The standard errors of estimate for VIQ, PIQ and FSIQ were 6.56, 10.67 and 7.63, respectively. Blair & Spreen (1989) indicated that for VIQ and FSIQ, a positive discrepancy of 15 points or more between estimated and actual IQ scores suggested the possibility of organic deterioration, while for PIQ, a positive discrepancy of at least 21 points between estimated and actual IQs suggested the possibility of organic deterioration.

Payne & Jones (1957) have investigated the clinical implications of a discrepancy between two test scores and have indicated that when comparing two tests, the magnitude of the variations or range in test scores is dependent on the underlying psychometric properties of the tests being compared. Thus, if two tests were perfectly correlated, any difference between standardized test scores would be significant. Whereas, if two tests were uncorrelated, then the performance on one test would have no relationship or predictive function with regard to performance on the other. Payne & Jones (1957) developed a statistical test for determining the probability for a found difference between two test scores that follows as, zD = (z1 - z2)/square root of (2 - 2r), where zD is the z score for the difference score, z1, z2 are the respective z scores obtained for the two tests by the subject (the discrepancy) and r is the correlation between the two tests in the normal population.

As previous studies have not calculated the variations between predicted and obtained IQ scores for normal subjects based on the correlations between the two measures, the Payne & Jones (1957) equation would provide a new approach to quantifying standard discrepancy so that clinical impairment could be assumed if the difference between a patient's predicted and obtained scores fell outside the normal range. To illustrate, correlations between predictive measures and obtained WAIS/WAIS-R FSIQs have ranged from .51 (NART/WAIS-R; Willshire et al., 1991) to .85 (Vocabulary/WAIS-R; Wechsler 1981, NDE/WAIS; Crawford et al., 1989) for normal subjects. Based on the Payne & Jones (1957) equation, with zD considered to be +1.645 (one tail) and correlations between predictive measures and FSIQs ranging from .85 to .51, a subject could obtain a discrepancy between predicted and actual FSIQs of 14 - 24 IQ points and still be considered to be in the normal range (95% confidence interval) based on the psychometric properties of the two tests. The higher the correlation is between a predictive measure of IQ and actual WAIS/WAIS-R IQ scores, the smaller the standard discrepancy is for a normal population. As indicated above, a correlation of .85 between measures produced a range of 14 IQ points, whereas a correlation of .51 produced a range of 24 IQ points.

From a clinical perspective, one wants to know if an obtained discrepancy in a patient's performance on two tests is due to actual cognitive impairment or rather results from chance fluctuations because of the psychometric properties of the two tests. If the standard discrepancy between two measures is large in a normal population, discrepancies will not be clinically useful since a patient's performance may be impaired but still fall within the standard range. Thus, a restricted standard range, possibly 10 points or less, for the normal population, is necessary for adequate sensitivity in identifying impairment at the clinical level. That is, an obtained IQ 10 points below the postdicted premorbid IQ should, ideally, be statistically significant at the $p = .05$ level. If IQ must fall by 20 points to be statistically discrepant from a postdictor score, then the postdictor would likely not be clinically useful since a drop of 20 points would lead to obvious impairment in daily functioning. However, to achieve a restricted range, the correlation between the predictive measure and actual IQs must be quite high. To date, the variables which have accounted for the greatest amount of variance in WAIS/WAIS-R FSIQs have included the Vocabulary subtest of the WAIS-R (Wechsler, 1981) and the NDE equation (Crawford et al., 1989). Furthermore, the Payne and Jones (1957) formula allows the sensitivity of different premorbid IQ predictors to be directly compared on a meaningful metric (IQ point decline). That is, if predictor "A" can reveal a 15 point IQ drop while predictor "B" can only reveal a 20 point IQ drop ($p = .05$), then predictor "A" is seen to be more sensitive.

An alternate approach to calculating the standard error in prediction between two measures would be to examine the standard deviation of the differences between predicted and obtained WAIS/WAIS-R IQ scores. As most predicted WAIS/WAIS-R IQ estimates have been derived from regression equations utilizing the predictive measure, the error in prediction, or the standard deviation of the difference scores, would be reflected in the standard error of estimate for the regression equation. Given the standard error of estimate (in IQ points), an error range could be calculated based on 95% confidence levels with $z = +1.96$ (two tail), or $z = +1.645$ (one tail). This approach is comparable to the method Blair & Spreen (1989) employed to calculate a standard error range in prediction of WAIS-R IQ based on NART-R errors for normal subjects. Although more sensitive, and hence potentially more useful, the applicability of this method is restricted as it either requires the original data set to calculate the standard deviation of difference scores, or requires the standard error of estimate from the original regression equation. In comparison, the Payne &

Jones (1957) equation can be applied to various data sets including standardization and cross-validation samples, as only the correlation coefficient between the two measures is required to calculate the standard discrepancy in scores.

Purpose

The purpose of the present study was to examine the discrepancy between predictive IQ measures and obtained WAIS/WAIS-R IQ scores in normal populations using the Payne & Jones (1957) equation, to evaluate the potential clinical usefulness of the measures for detecting cognitive impairment. The Payne & Jones (1957) equation was selected as the primary approach for examining the discrepancy between predicted and obtained IQ scores as it could be applied to a wider range of standardization and cross-validation samples, and consequently allowed for greater comparison of measures. Initially, this study determined the standard discrepancy between predictive IQ measures and obtained WAIS/WAIS-R IQ scores for studies that examined the correlations between these measures for normal subjects. Particularly, the standard discrepancy between NART predicted and WAIS-R obtained IQs in the Carswell et al., (1993) study was examined, as that study attempted to provide a more accurate portrayal of a predictive measure of IQ by examining the ability of the NART to postdict WAIS-R IQs over a five year period. Then, using data from the Carswell et al., (1993) study, the present study attempted to reduce the discrepancy in scores between predictive and obtained IQ measures by developing a regression equation to account for a greater proportion of variance in WAIS-R VIQ. This study examined predictive measures of IQ including the NART, demographics and the Vocabulary subtest of the WAIS-R to determine which measures, either alone or in combination, would produce a regression equation that would most accurately postdict WAIS-R VIQ scores. Since all previous studies have shown that the NART and Vocabulary scores are most accurate at estimating VIQ, this study focused on estimating only VIQ. The study also evaluated the predictive ability of the WAIS-R VIQ regression equation developed by Ryan & Paolo (1992). From previous research it was considered that:

(1) NART and demographic data would account for more of the variance in WAIS-R VIQs than the NART alone or demographic data alone

(2) the Vocabulary subtest of the WAIS-R would account for more of the variance in WAIS-R VIQ than the NART alone or demographics alone

(3) the amount of variance accounted for by the combination of the Vocabulary subtest of the WAIS-R with the NART or demographic data was yet to be determined.

The underlying aim of the study was to determine the most accurate way of postdicting WAIS-R IQ scores in normal subjects in order to make detecting cognitive impairment more reliable in clinical samples. This study attempted to demonstrate the potential usefulness of the various pre/postdictors in normal subjects. Although it is necessary to have a high correlation in normals between obtained IQs and predicted values, this is not sufficient for an estimate of premorbid IQ. To be reliable as an estimate of premorbid IQ, the measure must not decline with cognitive impairment. This study did not assess the clinical reliability of the regression equation developed to postdict VIQ in a normal sample, thus, future cross-validation is necessary to determine clinical usefulness.

Method

Initially, discrepancy scores were calculated for studies that examined the correlations between predictive IQ measures and obtained WAIS/WAIS-R IQ scores for normal subjects. Discrepancy scores were determined by use of the Payne & Jones (1957) equation:

$$(z1-z2)=zD\sqrt{(2-2r)}$$

For this study, zD was set at +1.645 (one tail) to represent the 95% confidence level, r was the correlation between IQ predictors and WAIS/WAIS-R IQs, and z1 and z2 were not calculated individually, but rather were considered as the difference (discrepancy) score. This discrepancy score was converted to IQ points by multiplication of 15 points.

The remainder of the current study was largely based on the original study by Carswell et al., (1993) which was conducted at Sunnybrook Health Science Centre in Toronto, Ontario. The original study served as part of a prospective study of dementia in which a sample of over 100 cognitively intact individuals were administered a neuropsychological test battery, including the full WAIS-R, three times at yearly intervals. Five years after the study had begun, 49 of these individuals (27 females and 22 males) were located who were willing to come in for further brief testing. At the time the NART was administered, the mean age and education of the subjects was 71 (S.D.=7.7) and 14.6 (S.D.=2.7) years, respectively.

The present study used the data collected at Sunnybrook Hospital to develop a regression equation that would be more accurate at postdicting WAIS-R VIQs in normal subjects and more useful for detecting cognitive impairment in clinical subjects. The study examined the predictive measures of IQ including the NART (administered at year 5), demographic variables (age, education and sex) and the Vocabulary subtest of the WAIS-R (administered at year 3) to develop a regression equation that would be more accurate at postdicting WAIS-R VIQ scores (obtained at year 1) in the above sample of 49 elderly normals. The NART was evaluated in terms of number of NART errors, age and education were considered in years, the Vocabulary score was age-scaled and sex was dummy variable coded with males = 1 and females = 2.

Discrepancy scores were calculated for the new regression equation based on the Payne & Jones (1957) equation. As the new regression equation also produced a standard error of estimate, a second approach to calculating discrepancy scores was employed, with the standard error of estimate for postdicted WAIS-R VIQ multiplied by $z = +1.645$ (one tail).

The accuracy of NART errors and demographics (age, education, sex), both alone and in combination, at postdicting WAIS-R VIQ was also examined for the above sample. Corresponding discrepancy scores were calculated utilizing both the Payne & Jones (1957) equation and the standard error of estimate.

The current study also examined the predictive utility of the WAIS-R VIQ regression equation developed by Ryan et al., (1992) at postdicting WAIS-R IQs in the above sample. The equation developed by Ryan et al., (1992) was as follows:

estimated WAIS-R VIQ = 132.3893 + (NART errors)(-1.164).

Discrepancy scores were determined for the above equation utilizing the Payne & Jones (1957) equation.

Results

Standard discrepancy scores based on the Payne & Jones (1957) equation with $zD = +1.645$ (one tail) are presented in Table 1 for studies that examined the correlations between predictive IQ measures and obtained WAIS/WAIS-R IQs for normal subjects. These studies involved administration of the predictive IQ measure and WAIS or WAIS-R at the same testing period.

Table 1

Standard Discrepancy Scores for Premorbid IQ Estimates calculated using the Payne & Jones Equation

Author	IQ Estimate	IQ Measure	IQ Test	Corr.	Discr. Score
Wechsler (1955)	Vocab.	WAIS	VIQ	.90	11
			FSIQ	.87	13
Wechsler (1981)	Vocab.	WAIS-R	VIQ	.90	11
			FSIQ	.85	14
Wilson et al., (1978)	Demogr.	WAIS	VIQ	.73	18
			PIQ	.65	21
			FSIQ	.74	18
Goldstein et al., (1986)	Wilson et al., ('78) equation	WAIS	VIQ	.71	19
			PIQ	.61	22

Karzmark et al., (1985)	Wilson et al., ('78) equation	WAIS	FSIQ	.68	20
Barona et al., (1984)	Demogr.	WAIS-R	VIQ PIQ FSIQ	.62 .49 .60	22 25 22
Eppinger et al., (1987)	Barona et al., ('84) equation	WAIS-R	VIQ PIQ FSIQ	.78 .60 .76	16 22 17
Barona et al., (1986)	Demogr.	WAIS-R	VIQ PIQ FSIQ	.69 .53 .66	19 24 20
Nelson (1982)	NART	WAIS	VIQ PIQ FSIQ	.77 .57 .74	17 23 18
Crawford, Parker et al., (1989)	NART	WAIS	VIQ PIQ FSIQ	.85 .57 .81	14 23 15
Willshire et al., (1991)	NART	WAIS-R	FSIQ	.51	24

Sharpe et al., (1991)	NART	WAIS-R	VIQ FSIQ	.81 .77	15 17
Ryan et al., (1992)	NART	WAIS-R	VIQ PIQ FSIQ	.78 .57 .74	16 23 18
Blair et al., (1989)	NART-R	WAIS-R	VIQ PIQ FSIQ	.83 .40 .75	14 27 17
Wiens et al., (1993)	NART-R	WAIS-R	VIQ PIQ FSIQ	.56 .22 .46	23 31 26
Crawford, Stewart, Parker et al., (1989)	NDE	WAIS	VIQ PIQ FSIQ	.88 .62 .85	12 22 14
Crawford, Nelson et al., (1990)	NDE	WAIS	VIQ PIQ FSIQ	.82 .61 .79	15 22 16

Standard discrepancy scores based on the Payne & Jones (1957) equation were also calculated for studies that examined the efficacy of predictive IQ measures at postdicting WAIS-R IQ scores for normal subjects. These discrepancy scores are presented in Table 2. Discrepancy scores for the Carswell et al., (1993) study, which examined the ability of the NART to postdict WAIS-R IQ scores obtained five years earlier, were 24, 28, and 24 IQ points for VIQ, PIQ, and FSIQ, respectively.

Table 2

Standard Discrepancy Scores for Premorbid IQ Estimates Serving a Postdictive Function calculated using the Payne & Jones Equation

Author	IQ Estimate	IQ Measure	Time	IQ Test	Corr.	Discr Score
Carswell et al., (1993)	NART	WAIS-R	5 yrs	VIQ PIQ FSIQ	.54 .37 .53	24 28 24
Berry et al., (1993)	NART-R	WAIS-R	3.5 yrs	FSIQ	.70	19

The present study used data from the Carswell et al., (1993) study which examined the effectiveness of the NART at postdicting WAIS-R scores in a sample of 49 cognitively intact elderly subjects. Mean, standard deviation, and range values for all variables are presented in Table 3.

Stepwise multiple regression was employed to determine which independent variables including demographics(age, sex, education), NART errors, and age-scaled WAIS-R Vocabulary scores were most effective at postdicting WAIS-R VIQ as the dependent variable. Predictor variables were obtained following initial administration of the WAIS-R in 1983, and year of administration for each variable is reported in Table 3 with demographic data including sex, collected in 1988. Multiple regression analysis was performed using SYSTAT MGLH module.

Table 3

Mean, Standard Deviation and Range Values for all Variables

Variable	Yr.	Mean	Std. Dev.	Range Min.	Range Max.
WAIS-R VIQ	('83)	117.47	9.36	91.0	132.0
VOCABULARY	('85)	14.06	2.09	10.0	18.0
NART errors	('88)	13.82	5.26	5.0	31.0
AGE	('88)	71.29	7.69	55.0	91.0
EDUCATION	('88)	14.61	2.68	8.0	19.0

Table 4

Correlations between Predictor Variables and WAIS-R VIQ

Predictor Variables	r	r^2
VOCABULARY	.607 *	.37
NART errors	-.546 *	.30
SEX	-.521 *	.27
EDUCATION	.299 *	.09
AGE	-.265	

($p < .05$ *)

Pearson (r) correlations of predictor variables with WAIS-R VIQ are presented in Table 4. All correlations were significant at $p < .05$, except for the demographic variable of age. The remaining variables including Vocabulary, NART errors, sex and education accounted for 37%, 30%, 27% and 9% of the variance respectively, in WAIS-R VIQ.

Minimum tolerance for entry of variables into the stepwise regression model was set at $p = .01$. Only three independent variables contributed significantly to the postdiction of WAIS-R VIQ. The first variable that entered the model was Vocabulary and it was significant with $R = .607$, $R^2 = .37$, $F(1,47) = 27.38$, $p < .001$. At step 2, sex entered the model and was significant with $R = .759$, $R^2 = .58$, $F(2.46) = 31.23$. NART errors entered the model at step 3 at significance and produced $R = .81$, $R^2 = .66$, $F(3,45) = 28.60$. Age and education were not found to contribute significantly to the regression. Thus, Vocabulary, sex and NART errors combined, accounted for 66% of the variance in WAIS-R VIQs.

As stepwise regression identified three independent variables that contributed significantly to the postdiction of WAIS-R VIQs, the following regression equation was developed:

WAIS-R VIQ = 109.04 - 0.548(NART errors) + 1.997(age-scaled Vocabulary score) - 7.783(sex).

($R = .81$, $R^2 = .66$, $F(3,45) = 28.60$, $p < .001$)

Based on the Payne & Jones (1957) equation with zD = + 1.645 and $\underline{R}$ = .81, the standard discrepancy score for this new WAIS-R postdiction regression equation was calculated at 15 IQ points. Given the standard error of estimate was 5.671 for the new WAIS-R postdiction regression equation, the standard discrepancy was calculated at 9.4 IQ points (z =+ 1.645).

A procedure known as the Williams modification of the Hotelling test (Kenny, 1987) was used to determine if Vocabulary, which produced the highest correlation with WAIS-R VIQ on its own, ($\underline{R}$ = .607), accounted for proportionately more of the variance in WAIS-R VIQ than NART errors ($\underline{R}$ = -.546wor sex ($\underline{R}$ = -.521). Results indicated that the correlations for Vocabulary and NART errors did not differ significantly at a .01 level of significance ($\underline{t}$(46) = .987). However, the correlations for Vocabulary and sex differed significantly ($\underline{t}$(46) = 2.85) with Vocabulary accounting for proportionately more of the variance in WAIS-R VIQ than sex.

Of the variables in the WAIS-R VIQ postdiction equation, sex accounted for a significant proportion of the variance. Independently, sex accounted for 27% of the variance in WAIS-R VIQ. Further analyses revealed that males and females differed significantly on WAIS-R VIQ scores, $\underline{t}(21) = 4.614$, $\underline{p} < .001$, with males consistently outperforming females. The mean WAIS-R VIQ score for males was 122.8 while the mean WAIS-R VIQ score for females was 113.1. However, males and females did not differ significantly with regard to age $t(21) = .454$, or years of education $t(21) = .055$. A scatterplot reflecting the distribution of WAIS-R VIQ scores by sex, with males coded as 1 and females coded as 2, is presented in Figure 1.

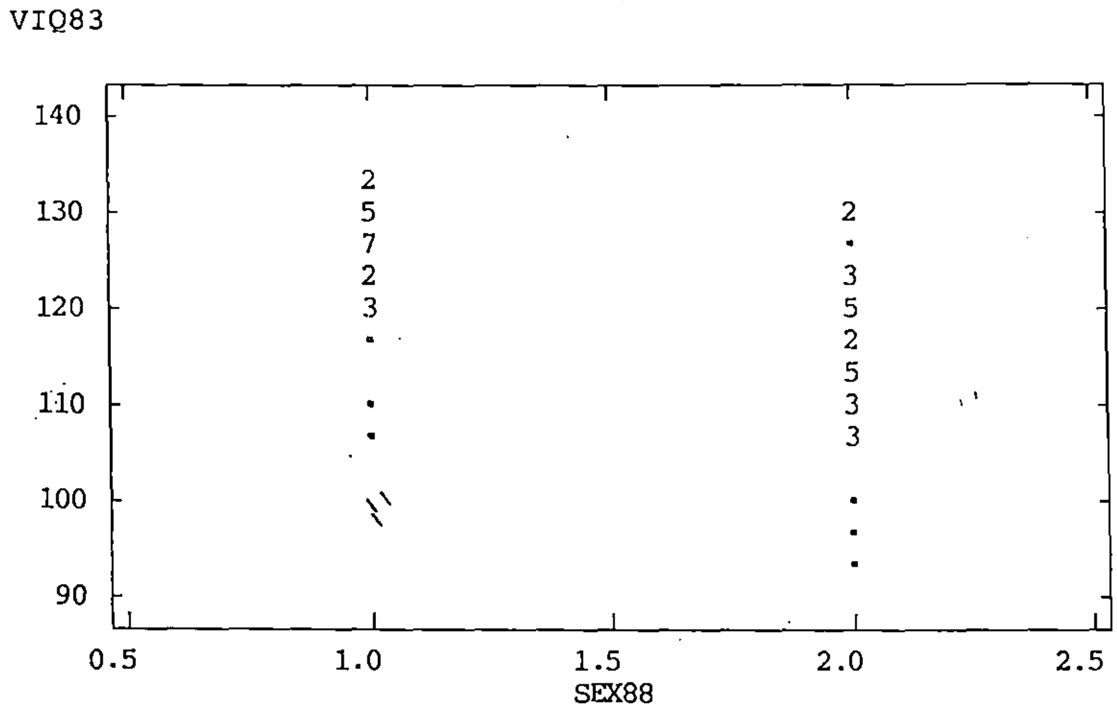

Figure 1. Scatterplot of WAIS-R VIQ scores by Sex (Males = 1, Females = 2)

As sex was heavily weighted in the regression equation and may have relected a potential sample bias, a second regression equation was developed utilizing only Vocabulary scores and NART errors. Vocabulary and NART errors alone contributed significantly to the postdiction of WAIS-R VIQs (R = .698), accounting for 49% of the variance in WAIS-R VIQs. Based on the above results, the following regression equation was developed:

WAIS-R VIQ = 97.066 + 2.101(age-scaled Vocabulary score) - 0.661(NART errors)

(R = .698, R^2 = .487, $F(2,46)$ = 21.841, p <.001)

The standard discrepancy score for this second WAIS-R postdiction regression equation, with zD = +1.645 and R = .698, was calculated at 19 IQ points. Given the standard error of estimate for this second WAIS-R postdiction regression equation was 6.848, the standard discrepancy was calculated at 11.3 IQ points (z = +1.645).

The accuracy of NART errors and demographics (age, education, sex), both alone and in combination, at postdicting WAIS-R VIQ was also examined. The combined equation of NART errors and demographic variables accounted for 54% of the variance in WAIS-R VIQ, $F(4,44)$ = 12.9, p < .001. Demographics alone accounted for 41% of the variance in WAIS-R VIQ, $F(3,45)$ = 10.3, p < .001, and NART errors alone accounted for 30% of the variance in WAIS-R VIQ, $F(1,47)$ = 19.99, p < .001. Thus, the combined equation predicted 13% more VIQ variance than the corresponding equation derived from demographic variables and 24% more variance than the corresponding NART equation. To determine whether this increase was statistically significant, F tests were performed on the change in the residual sum of squares when NART or demographic equations were compared to the combined equation. The proportion of VIQ variance accounted for by the combined equation differed significantly from the variance accounted for by the NART alone, $F(3,44) = 17.36$, $p < .001$. Similarly, the combined equation also accounted for a greater proportion of VIQ variance than the demographic equation, $F(1,44) = 52.07$, $p < .001$. The amount of WAIS-R VIQ variance accounted for by the three methods of postdiction, and corresponding discrepancy scores in IQ points as calculated by the Payne & Jones (1957) equation and the standard error of estimate, are presented in Table 5.

Table 5

VIQ Variance and Critical Discrepancy Scores using the NART and Demographics

			WAIS-R VIQ		
Variables	r	r^2	(P&J) Disc. Score	S.E. of Est.	S.E. Disc. Score
NART	.546	.30	23	7.924	13.1
SEX+AGE+ED.	.639	.41	21	7.439	12.3
COMBINATION	.736	.54	18	6.619	10.9

(P&J: Payne & Jones (1957) Equation)

The accuracy of the Ryan & Paolo (1992) WAIS-R regression equation at postdicting WAIS-R VIQ was also examined. This equation utilized NART errors alone. Subject scores based on this équation were regressed on WAIS-R VIQ. Results indicated that the Ryan et al., (1992) VIQ estimates were significant, ($F(1,47) = 20.43$, $p < .001$) and accounted for 30% of the variance, ($R = .550$) in obtained VIQs. The standard discrepancy score for the Ryan et al., (1992) WAIS-R regression equation in this sample of elderly normals was calculated at 23 IQ points.

Table 6

The Efficacy of IQ Estimates at Postdicting WAIS-R VIQ

		WAIS-R VIQ		
Variables	r	(P&J) Disc. Score	S.E. of Estimate	S.E. Disc. Score
VOCABULARY+NART+SEX	.810	15	5.671	9.4
NART+SEX+AGE+ED.	.736	18	6.619	10.9
VOCABULARY+NART	.700	19	6.848	11.3
SEX+AGE+ED.	.639	21	7.439	12.3
VOCABULARY	.607	22	7.520	12.4
NART	.546	23	7.924	13.1

(P&J: Payne & Jones (1957) Equation)

As the current study examined the efficacy of various IQ estimates, alone or in combination, at postdicting WAIS-R VIQ in the above sample of elderly normals, a summation reflecting the correlation of each measure with WAIS-R VIQ and corresponding discrepancy scores as calculated by the Payne & Jones (1957) equation and the standard error of estimate, is presented in Table 6. From Table 6, it is apparent that the initial regression equation with significant variables including the NART, Vocabulary, and sex, accounted for the greatest proportion of WAIS-R VIQ variance, and produced a restricted and potentially clinically useful discrepancy range for identifying cognitive impairment in a clinical sample.

Discussion

Initially, the current study examined the discrepancy between predictive IQ measures and obtained WAIS/WAIS-R IQ scores in normal populations using the Payne & Jones (1957) equation, to evaluate the potential clinical usefulness of the measures for detecting cognitive impairment. Of these studies, all involved administration of the predictive IQ measure and the WAIS or WAIS-R at the same testing period. From the results, the magnitude of the discrepancies between predicted and obtained FSIQs for the selected studies ranged from 13 IQ points (Vocabulary; Wechsler, 1955) to 26 IQ points (NART-R; Wiens et al., 1993). Thus, a patient's predicted and actual WAIS/WAIS-R FSIQ scores could deviate by a minimum of 13 IQ points or a maximum of 26 IQ points, depending on the premorbid estimate used, and still be considered to fall within the normal range. Similarly, the magnitude of the discrepancies between predicted and obtained VIQs for selected studies ranged from 11 IQ points (Vocabulary; Wechsler, 1955, 1981) to 23 IQ points (NART-R; Wiens et al., 1993).

As the standard discrepancies between predicted and obtained IQ scores were quite large for normal subjects, the clinical utility of the selected predictive measures for estimating IQ was questionable. Specifically, impairment in performance is often inferred if a patient's scores reveal a significant difference between predicted and obtained IQs. However, before impairment can be assumed, the standard discrepancy between the two measures must be considered. If the standard discrepancy between two measures is large in a normal population, discrepancies will not be clinically useful as a patient's performance may be impaired but still fall within the standard range. Clearly, if observed FSIQ must be 26 points less than predicted IQ for the difference to be statistically discrepant, then the IQ predictor may not be clinically useful. In other words, once IQ has dropped 26 points, the patient would likely be showing impairment in daily functioning, rendering the IQ predictions superfluous. Similarly, if VIQ must drop by 23 points to be considered statistically discrepant, the clinical utility of the predictive measure is also questionable.

It was originally considered that a restricted standard range of possibly ten points or less, for a normal population, would demonstrate adequate sensitivity for identifying cognitive impairment at the clinical level. Although the results revealed that in general, the standard discrepancies between predicted and obtained VIQs were quite large for selected studies (11-23 IQ points), seven studies produced VIQ discrepancies that ranged between 11-15 IQ points and demonstrated potential clinical utility. A further advantage of the Payne & Jones (1957) formula was that it allowed the sensitivity of different premorbid IQ predictors to be compared on a meaningful metric (IQ point decline). In comparison, a measure that produced a 15 point IQ discrepancy demonstrated greater sensitivity as a predictor of premorbid IQ and potentially greater clinical utility than a measure that produced a 23 IQ point discrepancy. In terms of predictive sensitivity, the Vocabulary subtest of the WAIS/WAIS-R (Wechsler, 1955, 1981), and the NDE (Crawford, Stewart, Parker et al., 1989) produced the smallest discrepancies between predicted and obtained VIQ scores in normal subjects at 11 and 12 IQ points, respectively.

However, the predictors which demonstrated potentially greater sensitivity in estimating premorbid IQ as compared to other measures, were constrained by the same methodological limitation which served to reduce the predictive sensitivity of all the measures in general. As all of the studies discussed above involved the administration of the predictive IQ measure and the WAIS/WAIS-R at the same testing period, the correlation between each predictive measure and the WAIS/WAIS-R was potentially elevated due to the fact that correlations between tests of similar abilities would be expected to be maximized if the two tests were administered at the same time. As the discrepancy scores reflected the correlation between the predictive measure and WAIS/WAIS-R VIQs, these scores could not be considered as sensitive or reliable as originally presumed. Therefore, even predictive measures that produced restricted discrepancy scores (11-15 IQ points) for normal subjects could not be considered as clinically reliable, as they demonstrated a tendency to overestimate the correlation between predicted and obtained WAIS/WAIS-R VIQ, and consequently reduced the magnitude of the discrepancy between predicted and obtained VIQs. Thus, based on the above results for the selected studies, it is suggested that caution be used in inferring cognitive impairment in a clinical sample, based on the discrepancy between a patient's predicted IQ score and actual WAIS/WAIS-R IQ.

As the above selected studies were constrained by a methodological limitation, the current study examined the magnitude of discrepancy scores for studies that attempted to postdict premorbid IQ in normal subjects by administering the pre/postdictive IQ measure at some point after the original WAIS/WAIS-R administration. This subsequently reduced the tendency to overestimate the correlation between pre/postdictive IQ and obtained WAIS/WAIS-R IQ scores, and produced a discrepancy range that was clinically more reliable. Two studies examined the postdictive ability of the NART and NART-R respectively, for normal subjects. Results indicated that discrepancy scores for the Carswell et al., (1993) study, which examined the ability of the NART to postdict WAIS-R IQ scores obtained five years earlier, were 24, 28 and 24 IQ points for VIQ, PIQ and FSIQ, respectively. Results from the Berry et al., (1993) study, which examined the ability of the NART-R to postdict WAIS-R IQ scores obtained 3.5 years earlier, produced a discrepancy score of 19 IQ points for FSIQ. While the Berry et al., (1993) study produced a discrepancy score that was somewhat smaller than the Carswell et al., (1993) study, it should be noted that the time interval between WAIS-R administration and NART administration was somewhat greater in the Carswell et al., (1993) study, a factor which would be expected to reduce the correlation still further, and subsequently increase the discrepancy score. As well, the Carswell et al., (1993) study examined the postdictive ability of the

NART, while the Berry et al., (1993) study examined the postdictive ability of the NART-R.

Although the Carswell et al., (1993) and Berry et al., (1993) studies increased the clinical reliability of selected premorbid IQ measures by examining their postdictive function, the results clearly indicated that the postdictive ability of the NART and NART-R as estimates of WAIS-R FSIQ was fairly modest ($r = .53$ and $r = .70$, respectively). Consequently, the standard discrepancies between postdicted and obtained IQ were large enough for normal subjects that the clinical utility of the measures was limited. The current study attempted to reduce the magnitude of the discrepancy in scores between postdicted and obtained IQ by developing a regression equation to account for a greater proportion of variance in WAIS-R VIQ, using postdictive data from the Carswell et al., (1993) study including Vocabulary scores, NART errors, and demographic data. Results indicated that the new regression equation produced a high correlation ($R = .81$) between significant postdictive IQ estimates (Vocabulary, NART errors, sex) and WAIS-R VIQs. This equation accounted for 66% of the variance in WAIS-R VIQs, and produced a 15 point discrepancy score based on the Payne & Jones (1957) equation between postdictive IQ estimates and obtained WAIS-R VIQs. The discrepancy score based on the standard error of estimate was calculated at 9.4 IQ points. The Vocabulary subtest score was the best single postdictor accounting for 37% of the variance in WAIS-R VIQ, with NART errors accounting for an additional 30%, and sex a further 27%.

As Vocabulary alone produced the highest correlation with WAIS-R VIQ, it was further determined if it accounted for proportionately more of the variance in WAIS-R VIQ than the NART alone, or sex alone. Results indicated that Vocabulary accounted for significantly more of the variance in WAIS-R VIQ than sex, but accounted for proportionately the same amount of variance in WAIS-R VIQ as NART errors. This result was somewhat unexpected as it was originally hypothesized that the Vocabulary subtest would account for proportionately more of the variance in WAIS-R VIQ than the NART alone, or significant demographic variables alone (sex). This a priori hypothesis was based on the fact that the Vocabulary subtest produced a higher correlation with WAIS-R VIQ ($r=.90$; Wechsler, 1981), than the NART ($r=.78$; Ryan et al., 1992) or demographic variables ($r=.69$; Barona et al., 1986). However, the above correlations were achieved with the administration of the predictive IQ measures and the WAIS-R at the same testing period. From the current study, it is evident that the NART is comparable to Vocabulary as a postdictive WAIS-R VIQ estimate, although Vocabulary accounted for proportionately more of the variance in WAIS-R VIQ than the demographic variable of sex.

In terms of practical utility, this new regression equation composed of three significant postdictive estimates (Vocabulary, NART errors, sex), demonstrated greater potential usefulness at postdicting premorbid IQ than either of the estimates on their own. This equation accounted for approximately twice the variance in WAIS-R VIQ as Vocabulary alone (66% vs. 37%), the NART alone (66% vs. 30%), or sex alone (66% vs. 27%). As this equation attempted to postdict WAIS-R VIQs, with measures administered <u>after</u> the original WAIS-R VIQ administration, it also reduced the tendency to overestimate premorbid IQ and increased the clinical reliability of the postdictive estimate. Further, it produced a restricted discrepancy range of 15 IQ points based on the Payne & Jones (1957) equation and 9.4 IQ points based on the standard error of estimate, which demonstrated potential clinical usefulness as an indicator of cognitive decline. This new equation was able to achieve the target of a standard discrepancy of 10 IQ points or less, and it reduced the magnitude of the original discrepancy in scores between postdicted and obtained VIQs in the Carswell et al., (1993) study. Thus, it produced a result that demonstrated potentially greater clinical utility than any previous postdiction method. However, before this new equation may indeed be considered clinically useful, it must first be cross-validated with both normal and clinical populations to determine if it is resistant to the effects of cognitive impairment.

A potential drawback of this new regression equation was that it was developed from a relatively small sample size (N=49), which could influence the generalizability of the results. Not only was the sample quite elderly in age with a mean age of 71.29 (S.D. = 7.69), but the sample was also above average in intelligence with a mean year 1 WAIS-R VIQ score of 117.47 (S.D. = 9.36). Nelson (1982) suggested that as the NART is limited by a ceiling effect, the NART may not be as good at predicting the premorbid IQ of individuals whose intelligence was above average. Further, sex alone accounted for 27% of the variance in WAIS-R VIQs, with males consistently outperforming females in the absence of any significant between sex differences in age or years of education. The mean difference between groups for WAIS-R VIQ scores was 9.7 IQ points. Although sex differences have been documented with regard to WAIS-R VIQ performance, none have been as prominent. Statistically significant between sex differences have been reported in terms of WAIS-R VIQ scores with males consistently outperforming females (Matarazzo, Bornstein, McDermott & Noonan, 1986; Inglis & Lawson, 1984; Kaufman, Kaufman-Packer, Mclean & Reynolds, 1991), although the clinical significance of these results has been questioned, as these differences have only accounted for 0.8 - 3.6 IQ points in WAIS-R VIQ scores. However, recently, Ilai & Willerman (1989) reported a significant between sex difference for WAIS-R VIQs with males outperforming females by an average of 5.6 IQ points in a sample of 206 young adults.

Kaufman et al., (1991) further examined the effect of age and gender on WAIS-R VIQ performance using the original WAIS-R standardization sample composed of 1480 adults aged 20-74 years. Although they found that sex and age were significant as main effects with males consistently outperforming females in terms of VIQ performance, and older subjects producing marginally higher VIQ scores than younger subjects, the age by sex interaction was not significant for VIQ performance. It is interesting to note however, that the greatest discrepancy between VIQ scores for males and females (3.6 IQ points) was found in the most elderly sample of subjects, aged 70-74 years. Further, Willshire et al., (1991), recently produced a regression equation combining NART errors and demographics for an elderly Australian population of 49 subjects, aged 55-69 years, in which sex was heavily weighted. Thus, future cross-validation of the current regression equation is clearly indicated to determine if this sex difference in WAIS-R VIQ is specific to the above sample, or is representative of a normal elderly population.

As sex contributed significantly to the postdiction of WAIS-R VIQs in the current elderly sample, and could not be confirmed as a true effect without future cross-validation, a second regression equation was developed using only

Vocabulary scores and NART errors. This equation also produced a fairly high correlation between postdictive IQ estimates and WAIS-R VIQs (R=.70), accounting for 49% of the variance in WAIS-R VIQs. Although this second equation did not account for as much of the variance in WAIS-R VIQs as did the original equation which included sex (49% vs. 66%), it accounted for proportionately more of the variance in WAIS-R VIQs than Vocabulary alone (49% vs. 37%), or the NART alone (49% vs. 30%). It also produced a postdictive estimate that may be more immediately useful in a clinical setting, and may be potentially more applicable to a wider age range of patients. However, given the variability in standard discrepancy scores for this second equation calculated at 19 IQ points for the Payne & Jones (1957) equation and 11.3 IQ points for the standard error of estimate, it is suggested that caution be used in inferring cognitive impairment based on a discrepancy in a patient's performance.

The accuracy of NART errors and demographics (age, education, sex) at postdicting WAIS-R VIQ, both alone and in combination, was also examined. Previous research indicated that the NART and demographic data in combination would account for more of the variance in WAIS-R VIQ than the NART or demographics alone. Specifically, Crawford, Stewart, Parker et al., (1989) developed a regression equation based on the combination of the NART with demographic variables (sex, social class, age) and found that the combined equation (NDE) accounted for more of the variance in WAIS IQ scores than the NART or demographics alone. Similarly, the current study found that the combined equation of NART errors and demographic variables (age, education, sex) accounted for 54% of the variance in WAIS-R VIQs, while demographic variables alone accounted for 41% of VIQ variance, and NART errors alone accounted for 30% of VIQ variance. The proportion of VIQ variance accounted for by the combined equation was also significantly greater than the variance accounted for by the demographics or NART alone. Although the NDE equation was developed from both premorbid IQ estimates and the WAIS administered at the same testing period (Crawford, Stewart, Parker et al., 1989), the results of the current study were consistent with the results of the prior study. Thus, it is apparent that, even when examining the postdictive function of IQ estimates, a greater proportion of IQ variance may be accounted for by using a combination of postdictive estimates, rather than evaluating the effectiveness of each IQ estimate alone.

In terms of clinical utility, the combined equation of postdictive estimates (NART+SEX+AGE+ED) demonstrated greater potential applicability, as it accounted for a greater proportion of VIQ variance than each IQ estimate alone. Results indicated that the combined equation produced a standard discrepancy score of 18 IQ points as calculated by the Payne & Jones (1957) equation, and 10.9 IQ points as calculated by the standard error of estimate. Demographics alone (sex, age, education) produced discrepancy scores as calculated by the Payne & Jones (1957) equation and the standard error of estimate, of 21 and 12.3 IQ points respectively, while the NART alone produced discrepancy scores of 23 and 13.1 IQ points, respectively. Although these estimates produced standard discrepancy scores that demonstrated potential clinical utility as indicators of cognitive impairment (as calculated by the standard error of estimate), the combined equation clearly produced a more restricted standard range. Thus, a combination of postdictive IQ estimates not only accounted for a greater proportion of VIQ variance than each IQ estimate alone, but subsequently increased the potential clinical utility of the estimates for detecting cognitive decline.

The current study also examined the accuracy of the Ryan & Paolo (1992) WAIS-R regression equation, which only utilized NART errors, at postdicting VIQ in the above sample of elderly normals. To date, this was the first cross-validation of this equation. The results indicated that the Ryan et al., (1992) estimates were significant and accounted for 30% of the variance in obtained WAIS-R VIQs. The standard discrepancy score for this regression equation was calculated at 23 IQ points for the above sample.

The Ryan et al., (1992) equation was also constrained by the same methodological limitation that influenced the studies discussed earlier, in that both the NART and the WAIS-R had been administered at the same testing period. Consequently, the original correlation between the results of the regression equation and obtained WAIS-R VIQ scores (r=.78; Ryan et al., 1992) was elevated in comparison to the proportion of WAIS-R VIQ variance accounted for by the equation in a postdictive capacity (r=.55). Similarly, the standard discrepancy for normal subjects calculated at 23 IQ points using the Payne & Jones (1957) equation, was too large to be considered clinically useful as an indicator of cognitive decline. Thus, the current cross-validation of the Ryan et al., (1992) equation with normal elderly subjects, was not as successful as the original study suggested.

Overall, in terms of postdictive efficacy, results indicated that the initial regression equation with significant variables including the NART, Vocabulary and sex, accounted for the greatest proportion of WAIS-R VIQ variance in the above sample of elderly normals. This measure subsequently produced the most restricted and potentially clinically useful discrepancy range (15 IQ points/Payne & Jones (1957); 9.4 IQ points/SE of Estimate) of all other postdictive estimates considered. As sex was heavily weighted in the initial regression equation and may have reflected a potential sample bias, a second regression equation was developed utilizing only Vocabulary scores and NART errors. This equation demonstrated greater potential generalizability than other estimates utilizing the demographic variable of sex (NART+SEX+AGE+ED; SEX+AGE+ED.), and accounted for a greater proportion of WAIS-R VIQ variance than Vocabulary alone or the NART alone. Thus, the two regression equations developed in the current study demonstrated the greatest postdictive efficacy and potential clinical utility of other postdictive estimates, either when sex was included as a variable, or when it was eliminated from the postdictive measure.

The current study also utilized two methods for calculating the standard discrepancy between pre/postdictive measures and obtained WAIS/WAIS-R IQ. These methods included the Payne & Jones (1957) equation, and the standard error of estimate based on the calculation of 95% confidence levels. As the Payne & Jones (1957) equation relied only on the correlation coefficient between the pre/postdictive measure and WAIS/WAIS-R IQ score, it demonstrated greater applicability to both standardization and cross-validation samples. Whereas, when the standard error of estimate was available, calculation of 95% confidence levels resulted in discrepancy scores that were more restricted and demonstrated greater potential clinical utility than discrepancy scores derived from the Payne & Jones (1957) equation. However, regardless of the method utilized for calculating standard discrepancy scores, results indicated that the two new regression equations developed in the current study demonstrated greater potential clinical utility than other existing estimates of premorbid IQ.

In conclusion, the current study successfully achieved its goal of identifying limitations of existing premorbid IQ estimates, and determining the most accurate way of postdicting WAIS-R VIQ scores in normal subjects, to make predicting premorbid IQ more reliable in clinical samples. The Payne & Jones (1957) equation provided a new approach to evaluate quantitatively the clinical sensitivity of several measures of premorbid IQ, and subsequently revealed that the standard discrepancies for many existing premorbid IQ estimates were too large to provide a clinically useful estimate of cognitive impairment. These estimates were further constrained by the fact that their predictive ability was somewhat inflated given that both predictive IQ scores and WAIS/WAIS-R IQ scores had been obtained at the same testing period. As the existing premorbid IQ estimates demonstrated limited practical utility, the current study developed two new regression equations to postdict WAIS-R VIQs in a sample of elderly normals. These new equations demonstrated greater potential clinical utility than other existing estimates of premorbid IQ, as they were both psychometrically and clinically reliable. However, before these equations can be applied clinically, they must first be cross-validated with both normal and clinical populations to determine if they are resistant to the effects of cognitive impairment. Nevertheless, it should be noted that to date, there is no comparable study that has generated regression equations to postdict WAIS-R VIQs in a normal sample, to make predicting premorbid IQ more reliable in a clinical sample.

References

Allen, R.M. (1947). The test performance of the brain injured. Journal of Clinical Psychology, 3, 225-230.

Allen, R.M. (1948). A note on the use of the Wechsler-Bellevue Scale Mental Deterioration Index with brain injured patients. Journal of Clinical Psychology, 4, 88-89.

Barona, A. & Chastain, R.L. (1986). An improved estimate of premorbid IQ for blacks and whites on the WAIS-R. International Journal of Clinical Neuropsychology, 8, 169-173.

Barona, A., Reynolds, C. & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS-R. Journal of Consulting and Clinical Psychology, 52, 885-887.

Bauer, C., Schlottmann, R., Kane, R. & Johnsen, D. (1984). An evaluation of the Digit Symbol component of the Russell, Neuringer and Goldstein Average Impairment Rating. Journal of Consulting and Clinical Psychology, 52, 317-318.

Berry, D.T.R., Carpenter, G.S., Campbell, D.A., Schmitt, F.A., Helton, K. & Lipke-Molby, T. (August, 1993). The New Adult Reading Test - Revised: Accuracy in estimating WAIS-R FSIQs obtained 3.5 years earlier. Paper presented at the American Psychological Association, Toronto, Canada.

Bershoff, D.N. (1970). The revised deterioration formula for the Wechsler Adult Intelligence Scale: A test of validity. Journal of Clinical Psychology, 26, 71-73.

Blair, J.R. & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. The Clinical Neuropsychologist, 3, 129-136.

Bolter, J., Gouvier, W., Veneklasen, J. & Long, C. (1982). Using demographic information to predict premorbid IQ: A test of clinical validity with head trauma patients. Clinical Neuropsychology, 4, 171-174.

Boyd, J.L., Cruickshank, C.A., Kenn, C.W., Madeley, P., Mindham, R.H.S., Oswald, A.C., Smith, R.J. & Spokes, E.G.S. (1991). Cognitive impairment and dementia in Parkinson's disease: a controlled study. Psychological Medicine, 21, 911-921.

Brayne, C. & Beardsall, L. (1990). Estimation of verbal intelligence in an elderly community: An epidemiological study using the NART. British Journal of Clinical Psychology, 29, 217-223.

Brinkman, S. & Braun, P. (1984). Classification of dementia patients by a WAIS profile related to central cholinergic deficiencies. Journal of Clinical Neuropsychology, 6, 393-400.

Carswell, L.M., Snow, W.G. & Tierney, M.C. (February, 1993). The NART as a postdictor of IQs in bright, normal elderly subjects. Paper presented at the 21st Annual

Meeting of the International Neuropsychological Society, Galveston, Texas.

Crawford, J.R. (1989). Estimation of premorbid intelligence: A review of recent developments. In: J.R. Crawford & D.M. Parker (Ed.). Developments in Clinical and Experimental Neuropsychology. London: Plenum Press, 55-74.

Crawford, J., Allan, K., Besson, J., Cochrane, R. & Stewart, L. (1990). A comparison of the WAIS and WAIS-R in matched UK samples. British Journal of Clinical Psychology, 29, 105-109.

Crawford, J.R., Besson, J.A.O., Bremner, M., Ebmeier, K.P., Cochrane, R.H.B. & Kirkwood, K. (1992). Estimation of premorbid intelligence in schizophrenia. British Journal of Psychiatry, 161, 69-74.

Crawford, J.R., Besson, J.A.O., Parker, D.M., Sutherland, K.M. & Keen, P.L. (1987). Estimation of premorbid intellectual status in depression. British Journal of Clinical Psychology, 26, 313-314.

Crawford, J.R., Cochrane, R.H.B., Besson, J.A.O., Parker, D.M. & Stewart, L.E. (1990). Premorbid IQ estimates obtained by combining the NART and demographic variables: construct validity. Personality and Individual Differences, 11, 209-210.

Crawford, J.R., Nelson, H.E., Blackmore, L., Cochrane, R.H.B. & Allan, K.M. (1990). Estimating premorbid intelligence by combining the NART and demographic variables: An examination of the NART standardisation sample and supplementary equations. Personality and Individual Differences, 11, 1153-1157.

Crawford, J.R., Parker, D.M. & Besson, J.A.O. (1988). Estimation of premorbid intelligence in organic conditions. British Journal of Psychiatry, 153, 178-181.

Crawford, J.R., Parker, D.M., Stewart, L.E., Besson, J.A.O. & De Lacey, G. (1989). Prediction of WAIS IQ with the National Adult Reading Test: Cross-validation and extension. British Journal of Clinical Psychology, 28, 267-273.

Crawford, J.R., Stewart, L.E., Cochrane, R.H.B., Parker, D.M. & Besson, J.A.O. (1989). Construct validity of the National Adult Reading Test: A factor analytic study. Personality and Individual Differences, 10, 585-587.

Crawford, J.R., Stewart, L.E., Parker, D.M., Besson, J.A.O. & Cochrane, R.H.B. (1989). Estimation of premorbid intelligence: Combining psychometric and demographic approaches improves predictive accuracy. Personality and Individual Differences, 10, 793-796.

Crookes, T.G. (1974). Indices of early dementia on WAIS. Psychological Reports, 34, 734.

Eppinger, M.G., Craig, P.L., Adams, R.L. & Parsons, O.A. (1987). The WAIS-R index for estimating premorbid intelligence: Cross-validation and clinical utility. Journal of Consulting and Clinical Psychology, 55, 86-90.

Filley, C.M., Kobayashi, J. & Heaton, R.K. (1987). Wechsler intelligence scale profiles, the cholinergic system and Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 9, 180-186.

Fromm, D., Holland, A.L., Nebes, R.D. & Oakley, M.A. (1991). A longitudinal study of word-reading ability in Alzheimer's disease: Evidence from the National Adult Reading Test. Cortex, 27, 367-376.

Fuld, P.A. (1984). Test profile of cholinergic dysfunction and of Alzheimer type dementia. Journal of Clinical Neuropsychology, 6, 380-392.

Gfeller, J.D. & Rankin, E.J. (1991). The WAIS-R profile as a cognitive marker of Alzheimer's disease: A misguided venture? Journal of Clinical and Experimental Neuropsychology, 13, 629-636.

Goldstein, F.C., Gary, H.E. & Levin, H.S. (1986). Assessment of the accuracy of regression equations proposed for estimating premorbid intellectual functioning on the Wechsler Adult Intelligence Scale. Journal of Clinical and Experimental Neuropsychology, 8, 405-412.

Gonen, J.Y. (1970). The use of Wechsler's Deterioration Quotient in cases of diffuse and symmetrical cerebral atrophy. Journal of Clinical Psychology, 26, 174-177.

Gouvier, W., Bolter, J., Veneklasen, J. & Long, C. (1983). Predicting verbal and performance IQ from demographic data: Further findings with head trauma patients. Clinical Neuropsychology, 5, 119-121.

Grober, E. & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid intelligence in the elderly. Journal of Clinical and Experimental Neuropsychology, 13, 933-949.

Hart, S., Smith, C.M. & Swash, M. (1986). Assessing intellectual deterioration. British Journal of Clinical Psychology, 25, 119-124.

Ilai, D. & Willerman, L. (1989). Sex differences in WAIS-R item performance. Intelligence, 13, 225-234.

Inglis, J. & Lawson, J.S. (1984). Handedness, sex and intelligence. Cortex, 20, 447-451.

Johnsen, D. Schlottmann, R., Kane, R., Bauer, C. & Quintana, D. (1985). An evaluation of the Mahan method for assessing intellectual impairment. The International Journal of Clinical Neuropsychology, 7, 8-10.

Karzmark, P., Heaton, R.K., Grant, I. & Matthews, C. (1985). Use of demographic variables to predict Full Scale IQ: A replication and extension. Journal of Clinical and Experimental Neuropsychology, 7, 412-420.

Kaufman, A.S., Kaufman-Packer, J.L., McLean, J.E. & Reynolds, C.R. (1991). Is the pattern of intellectual growth and decline across the adult life span different for men and women? Journal of Clinical Psychology, 47, 801-812.

Kenny, D.A. (1987). Testing measures of association. In: D.A. Kenny (Ed.). Statistics for the Social and Behavioural Sciences. Boston: Little Brown and Company, 270-288.

Klesges, R., Fisher, L., Vasey, M. & Pheley, A. (1985). Predicting adult premorbid functioning levels: Another look. The International Journal of Clinical Neuropsychology, 7, 1-3.

Klesges, R., Sanchez, V. & Stanton, A. (1981). Cross-validation of an adult premorbid functioning index. Clinical Neuropsychology, 3, 13-15.

Klesges, R. & Troster, A.I. (1987). A review of premorbid indices of intellectual and neuropsychological functioning: What have we learned in the past five years?. The International Journal of Clinical Neuropsychology, 9, 1-10.

Klesges, R.C., Wilkening, G.N. & Golden, C.J. (1981). Premorbid indices of intelligence: a review. Clinical Neuropsychology, 3, 32-39.

Lezak, M. (1983). Neuropsychological assessment. (2nd Ed.). New York: Oxford University Press.

Livesay, J.R. (1986). Clinical utility of Wechsler's Deterioration Index in screening for behavioural impairment. Perceptual and Motor Skills, 63, 619-626.

Logsdon, R.G., Teri, L., Williams, D.E., Vitiello, M. & Prinz, P.N. (1989). The WAIS-R profile: A diagnostic tool for Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 11, 892-898.

Mahan, H. (1979). Measuring intellectual impairment with the WAIS: An improved approach to scoring. Clinical Neuropsychology, 1, 54-63.

Martin, A. & Fedio, P. (1983). Word production and comprehension in Alzheimer's Disease: The breakdown of semantic knowledge. Brain and Language, 19, 124-141.

Matarazzo, J.D. (1972). Wechsler's Measurement and Appraisal of Adult Intelligence. (5th Ed.). Baltimore: Williams & Wilkins.

Matarazzo, J.D., Bornstein, R.A., McDermott, P.A. & Noonan, J.V. (1986). Verbal IQ vs. Performance IQ difference scores in males and females from the WAIS-R standardization sample. Journal of Clinical Psychology, 42, 965-975.

Matarrazo, J.D., Carmody, J.P. & Jacobs, L.D. (1980). Test-retest reliability and stability of the WAIS: A literature review with implications for clinical practise. Journal of Clinical Neuropsychology, 2, 89-105.

McFie, J. (1975). Assessment of organic intellectual impairment. New York: Academic Press.

Meyer, V. (1961). Psychological effects of brain damage. In:

H.J. Eysenck (Ed.). Handbook of abnormal psychology. New York: Basic Books.

Nebes, R.D., Martin, D.C. & Horn, L.C. (1984). Sparing of semantic memory in Alzheimer's disease. Journal of Abnormal Psychology, 93, 321-330.

Nelson, H.E. (1982). National Adult Reading Test (NART). Test Manual. Windsor: NFER Nelson.

Nelson, H.E. & McKenna, P. (1975). The use of current reading ability in the assessment of dementia. British Journal of Social and Clinical Psychology, 14, 259-267.

O'Carroll, R.E. (1987). The inter-rater reliability of the National Adult Reading Test (NART): A pilot study. British Journal of Clinical Psychology, 26, 229-230.

O'Carroll, R.E. & Gilleard, C.J. (1986). Estimation of premorbid intelligence in dementia. British Journal of Clinical Psychology, 25, 157-158.

O'Carroll, R.E., Moffoot, A., Ebmeier, K.P. & Goodwin, G.M. (1992). Estimating pre-morbid intellectual ability in the Alcoholic Korsakoff syndrome. Psychological Medicine, 22, 903-909.

O'Carroll, R.E., Walker, M., Dunan, J., Murray, C., Blackwood, D., Ebmeier, K.P. & Goodwin, G.M. (1992). Selecting controls for schizophrenia research studies: the use of the national adult reading test (NART) is a measure of premorbid ability. Schizophrenia Research, 8, 137-141.

Paolo, A.M. & Ryan, J.J. (1992). Generalizability of two methods of estimating premorbid intelligence in the elderly. Archives of Clinical Neuropsychology, 7, 135-143.

Payne, R.W. & Jones, H.G. (1957). Statistics for the investigation of individual cases. Journal of Clinical Psychology, 13, 115-121.

Rabin, I.A. (1965). Diagnostic use of intelligence tests. In: B.B. Wolman (Ed.). Handbook of clinical psychology. New York: McGraw Hill.

Reitan, R.M. (1959). The comparative effects of brain damage on the Halstead Impairment Index and the Wechsler-Bellevue Scale. Journal of Clinical Psychology, 15, 281-285.

Russell, E.W. (1972). WAIS factor analysis with brain-damaged subjects using criterion variables. Journal of Consulting and Clinical Psychology, 39, 133-139.

Ryan, J.J. & Paolo, A.M. (1992). A screening procedure for estimating premorbid intelligence in the elderly. The Clinical Neuropsychologist, 6, 53-62.

Ryan, J.J. & Prifitera, A. (1990). The WAIS-R index for estimating premorbid intelligence: Accuracy in predicting short form IQ. International Journal of Clinical Neuropsychology, 12, 20-23.

Satz, P., Van Gorp, W.G., Soper, H.V. & Mitrushina, M. (1987). WAIS-R marker for dementia of the Alzheimer's type? An empirical and statistical induction test. Journal of Clinical and Experimental Neuropsychology, 9, 767-774.

Sharpe, K. & O'Carroll, R. (1991). Estimating premorbid intellectual level in dementia using the National Adult Reading Test: A Canadian study. British Journal of Clinical Psychology, 30, 381-384.

Starr, J.M., Whalley, L.J., Inch, S. & Shering, P.A. (1992). The quantification of the relative effects of age and NART-predicted IQ on cognitive function in healthy old people. International Journal of Geriatric Psychiatry, 7, 153-157.

Stebbins, G.T., Gilley, D.W., Wilson, R.S., Bernard, B.A. & Fox, J.H. (1990). Effects of language disturbances on premorbid estimates of IQ in mild dementia. The Clinical Neuropsychologist, 4, 64-68.

Stebbins, G.T., Wilson, R.S., Gilley, D.W., Bernard, B.A. & Fox, J.H. (1990). Use of the National Adult Reading Test to estimate premorbid IQ in dementia. The Clinical Neuropsychologist, 4, 18-24.

Sullivan, E.V., Sagar, H.J., Gabrieli, J.D.E., Corkin, S. & Growdon, J.H. (1989). Different cognitive profiles on standard behavioural tests in Parkinson's Disease and Alzheimer's Disease. Journal of Clinical and Experimental Neuropsychology, 11, 799-820.

Tuokko, H. & Crockett, D. (1987). Central cholinergic deficiency WAIS profiles in a nondemented aged sample. Journal of Clinical and Experimental Neuropsychology, 9, 225-227.

Vogt, A.T. & Heaton, R.K. (1977). Comparison of Wechsler Adult Intelligence Scale indices of cerebral dysfunction. Perceptual and Motor Skills, 45, 607-615.

Watson, C.G., Thomas, R.W., Anderson, D. & Felling, J. (1968). Differentiation of organics from schizophrenics at two chronicity levels by use of the Reitan-Halstead organic test battery. Journal of Consulting and Clinical Psychology, 32, 679-684.

Wechsler, D. (1939). The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins.

Wechsler, D. (1944). Measurement of Adult Intelligence (3rd Ed.). Baltimore: Williams & Wilkins.

Wechsler, D. (1955). Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corporation.

Wechsler, D. (1958). The Measurement and Appraisal of Adult Intelligence (4th Ed.). Baltimore: Williams & Wilkins.

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

Whitehead, A. (1973). The pattern of WAIS performance in elderly psychiatric patients. British Journal of Social and Clinical Psychology, 12, 435-436.

Wiens, A.N., Bryan, J.E. & Crossen, J.R. (1993). Estimating WAIS-R FSIQ from the National Adult Reading Test - Revised in normal subjects. The Clinical Neuropsychologist, 7, 70-84.

Willshire, D., Kinsella, G. & Prior, M. (1991). Estimating WAIS-R IQ from the National Adult Reading Test: A cross-validation. Journal of Clinical and Experimental Neuropsychology, 13, 204-216.

Wilson, R.S., Rosenbaum, G. & Brown, G. (1979). The problem of premorbid intelligence in psychological assessment. Journal of Clinical Neuropsychology, 1, 49-53.

Wilson, R.S., Rosenbaum, G., Brown, G., Rourke, D., Whitman, D. & Gisell, J. (1978). An index of premorbid intelligence. Journal of Consulting and Clinical Psychology, 46, 1554-1555.

Young, A.H., Blackwood, D.H.R., Roxborough, H., McQueen, J.K., Martin, M.J. & Kean, D. (1991). A magnetic resonance imaging study of schizophrenia: Brain structure and clinical symptoms. British Journal of Psychiatry, 158, 158-164.

Appendix

Raw Data Set

Sub. #	WAIS-R VIQ (83)	VOC (85)	NART Errors(88)	AGE (88)	ED. (88)	SEX (88)
1	118	15	20	60	14	2
2	124	14	16	84	12	1
3	105	14	11	72	18	2
4	119	12	18	76	16	1
5	108	14	13	78	12	2
6	91	10	12	81	13	2
7	110	13	16	74	9	1
8	118	10	9	66	16	1
9	129	17	16	70	13	1
10	113	16	13	63	15	2
11	120	16	11	74	18	2
12	99	10	23	79	16	2
13	127	15	14	75	13	1
14	124	16	9	78	10	1
15	124	16	10	76	18	1
16	106	13	13	82	15	2

17	122	16	13	81	19	2
18	132	14	9	75	15	1
19	117	14	16	75	12	1
20	128	14	8	79	17	2
21	113	12	17	70	13	2
22	126	18	18	70	16	1
23	109	11	17	73	8	2
24	113	16	6	91	17	2
25	128	14	5	62	13	2
26	120	14	7	64	14	2
27	105	14	21	79	16	2
28	125	16	13	73	16	1
29	120	16	12	62	18	1
30	109	11	24	68	10	2
31	131	18	9	73	16	1
32	117	14	12	68	16	2
33	119	17	11	65	16	2
34	125	16	11	73	17	2
35	118	15	16	55	16	2

36	124	14	8	69	14	1
37	116	15	21	63	10	2
38	122	14	15	66	14	1
39	113	14	21	77	12	2
40	94	10	31	83	12	2
41	106	11	20	65	13	1
42	113	12	16	61	13	2
43	129	12	8	77	18	1
44	116	12	12	69	12	1
45	124	14	17	62	17	1
46	128	14	7	60	16	1
47	117	15	11	64	16	2
48	127	15	10	61	18	1
49	115	16	11	72	18	2

Note. WAIS-R VIQ was considered in IQ points, the NART was evaluated in terms of the number of NART errors based on the United Kingdom pronunciation standards, the Vocabulary score was age-scaled, age and education were considered in years, and sex was dummy variable coded with males = 1 and females = 2.

VITA

Surname: Carswell Given Names: Lisa Marie

Place of Birth: Toronto,Canada Date of Birth: 26/06/68

Educational Institutions Addended:

University of Victoria	1991 to 1995
University of Western Ontario	1987 to 1991

Degrees Awarded:

B.Sc. (Honours) University of Western Ontario 1991

Honours and Awards:

In-Course Scholarship 1989-1990
University of Western Ontario

Dean's Honour List Scholar 1989-90; 90-91
University of Western Ontario

Graduate Teaching Fellowship 1991-1992
University of Victoria

Presentations:

Rollman, G.B. & Carswell, L.M. (1991 November). Stress-induced analgesia in humans? Paper presented at the 32nd. Annual Meeting of the Psychonomic Society, San Francisco, California.

Carswell, L.M., Snow, W.G. & Tierney, M.C. (1993 February). The NART as a postdictor of IQs in bright normal elderly subjects. Paper presented at the 21st Annual Meeting of the International Neuropsychological Society, Galveston, Texas.

Carswell, L.M., Graves, R.E., Snow, W.G. & Tierney, M.C. (1995 February). The effectiveness of premorbid IQ estimates at postdicting verbal IQ in normal elderly subjects. Paper presented at the 23rd. Annual Meeting of the International Neuropsychological Society, Seattle, Washington.

Carswell, L.M., Graves, R.E. & Snow, W.G. (1995 February). Quantitative evaluation of the sensitivity of premorbid IQ estimates. Paper presented at the 23rd. Annual Meeting of the International Neuropsychological Society, Seattle, Washington.