INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Prediction of Memory and Language Performance in Normal Elderly Canadians: Implications for the Assessment of Premorbid Cognition in Early Alzheimer's Disease

by

Lisa Marie Carswell
B.Sc., University of Western Ontario, 1991
M.Sc., University of Victoria, 1995

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

DOCTOR OF PHILOSOPHY

in the Department of Psychology

We accept this dissertation as conforming to the required standard

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

Dr. C. Mateer, Departmental Member (Department of Psychology)

Dr. M. Joschko, Departmental Member (Department of Psychology)

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

Dr. G. Iverson, External Examiner (Department of Psychiatry, University of British Columbia)

Supervisor: Dr. Roger E. Graves

ABSTRACT

The present study examined the concurrent validity of several proposed measures of premorbid IQ, present ability measures, and demographic variables at predicting intellectual, verbal memory, and language performance in a sample of 98 normal elderly Canadians (mean age = 71.9 years). Predictive regression equations were developed to estimate performance on criterion measures in each cognitive domain including general intellectual ability (i.e. Wechsler Adult Intelligence Scale-Revised Verbal IQ: WAIS-R VIQ), verbal memory (i.e. California Verbal Learning Test long delayed free recall: CVLTLDFR), and language domains (i.e. Boston Naming Test: BNT). These new regression equations utilized a combination of measures of premorbid VIQ and present ability measures to account for 63%, 32% and 54% of the variance in WAIS-R VIQ, CVLTLDFR, and BNT performance, respectively. The utility of these new equations for detecting impaired performance and cognitive decline in clinical samples was evaluated by calculation of sensitivity scores for each equation based on the method proposed by Graves, Carswell & Snow (in press). The results indicated that performance would have to decline by approximately 15 points for WAIS-R VIQ, 6 points for CVLTLDFR, and 6 points

for BNT scores, to be reliably detected 80% of the time. The implications of the sensitivity of each of these equations was discussed with regard to the clinical application of these equations for predicting premorbid cognition in early Alzheimer's disease. The current study was also the first study to develop predictive regression equations utilizing measures of premorbid VIQ and present ability measures to estimate verbal memory and language performance in a healthy elderly sample.

Examiners:

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

Dr. C. Mateer, Departmental Member (Department of Psychology)

Dr. M. Joschko, Departmental Member (Department of Psychology)

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

Dr. G. Iverson, External Examiner (Department of Psychiatry, University of British Columbia)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Dr. Roger Graves for his supervision on this project, and for his support and guidance throughout my graduate program at the University of Victoria.

Prediction of Memory and Language Performance in Normal Elderly Canadians: Implications for the Assessment of Premorbid Cognition in Early Alzheimer's Disease

A premorbid measure of cognitive ability is one which is purported to provide an estimate of premorbid functioning in a specific cognitive domain prior to the onset of disease process or injury. The utility of a premorbid measure of cognitive ability is that it not only identifies potential impairment if premorbid ability is estimated at a higher level than current cognitive functioning, but it offers an index of the extent of cognitive decline given the discrepancy between a patient's predicted cognitive performance and actual cognitive ability. To date, much of the research in the estimation of premorbid cognitive ability has focused on the prediction of performance in the domain of general intellectual ability.

PART I:

History of Premorbid IQ Prediction

Wechsler (1944) originally considered that preserved cognitive performance in the presence of organic impairment provided an estimate of a patient's premorbid level of functioning that could be useful in detecting/quantifying cognitive decline. Wechsler (1944) observed that subtests of

the Wechsler-Bellevue Test (WB; Wechsler, 1939) were similarly affected by age and organic impairment, and considered that if abilities that appeared resistant to the effects of aging were also resistant to the effects of organic brain damage, then these abilities would most accurately represent a patient's premorbid level of functioning regardless of current organic involvement. Wechsler (1944) subsequently developed a deterioration index in which 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). The original "Hold" subtests from the Wechsler-Bellevue Test included Information, Comprehension, Object Assembly, and Picture Completion, while the "Don't Hold" subtests included Digit Span, Arithmetic, Digit Symbol, and Block Design (Wechsler, 1944). With the development of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955), test combinations changed for Hold vs. Don't Hold subtests in which Vocabulary replaced Comprehension as a "Hold" subtest and Similarities substituted for Arithmetic as a "Don't Hold" subtest. Although controversy has surrounded the stability of Wechsler's four "Hold" tests when brain damage has been sustained (Klesges, Wilkening, & Golden, 1981), and the effectiveness of the deterioration indices has not been very

encouraging with regard to detecting cognitive decline in clinical samples (Meyer, 1961; Rabin, 1965), the Wechsler deterioration indices became widely known, and served to introduce the concept of employing tests of present ability for use in estimating premorbid IQ.

Criteria and Current Measures for Predicting Premorbid IQ

Crawford (1989) indicated that for a measure to qualify as a valid means of estimating premorbid IQ, it must not only demonstrate adequate reliability as a testing instrument and correlate highly with IQ in a normal population, but it must also remain unaffected by the consequences of neurological or psychiatric disorder. Efforts at quantifying premorbid IQ performance have largely focused on providing an estimate of Wechsler Adult Intelligence Scale/Revised IQ (WAIS; Wechsler, 1955/WAIS-R; Wechsler, 1981), currently the most popular measure of intellectual functioning (Harrison, Kaufman, Hickman, & Kaufman, 1988). Two main approaches have attempted to fulfill the above criteria for predicting premorbid IQ including 1) the use of demographic variables and 2) the use of postmorbid test results. The use of postmorbid test results, particularly reading ability, gained popularity in the estimation of premorbid IQ based on an initial observation by Nelson and McKenna (1975) that reading ability (oral pronunciation of words) appeared to remain

relatively preserved in elderly patients with dementing disorders.

Currently, the most commonly used approaches for estimating premorbid IQ in clinical practise include: 1) ***demographic based regression equations*** (Wilson, Rosenbaum, Brown, Rourke, Whitman, & Gisell, 1978; Barona, Reynolds, & Chastain, 1984; Barona & Chastain, 1986), 2) ***measures of present ability*** such as the Vocabulary subtest of the Wechsler scales, and reading measures including the National Adult Reading Test (NART; Nelson, 1982), the American version of the NART (AMNART; Schwartz & Saffran, 1987, as cited in Grober & Sliwinski, 1991), the New Adult Reading Test-Revised (NART-R; Blair & Spreen, 1989; also known as the North American Adult Reading Test - NAART), and the Reading subtest of the Wide Range Achievement Test-Revised (WRAT-R; Jastak & Wilkinson, 1984) and 3) ***a combination of demographic variables and present ability measures*** such as the NART Demographic Equation (NDE; Crawford, Stewart, Parker, Besson, & Cochrane, 1989), AMNART and demographics (Grober & Sliwinski, 1991), NART and demographics (WAIS-R norms; Willshire, Kinsella, & Prior, 1991), WRAT-R and demographics (Kareken, Gur, & Saykin, 1995), NART and WAIS-R Vocabulary subtest (Carswell, Graves, Snow, & Tierney, 1997), WAIS-R subtests and demographics (Krull, Scott, & Sherer, 1995; Vanderploeg & Schinka, 1995; Paolo, Ryan, & Tröster, 1997). Alternative measures recently proposed for

estimating premorbid IQ include the Cambridge Contextual Reading Test (Beardsall & Huppert, 1994), which is based on placing NART words in the context of meaningful sentences to facilitate semantic retrieval, and the Spot-the-Word Test (Baddeley, Emslie, & Nimmo-Smith, 1993), a lexical decision making task in which individuals are presented with two items and are required to state which of the pair is a real word (O'Carroll, 1995).

It is interesting to note that *personality variables* have also been considered in the prediction of premorbid IQ. Schlottmann and Johnsen (1991) attempted to determine premorbid intelligence through the use of a regression formula based on a scale of items reflecting interests, attitudes, values and personality characteristics. Schlottmann and Johnsen (1991) found that the resulting scale, the Intellectual Correlates Scale (ICS) contributed significantly to the prediction of WAIS-R IQs in a normative sample, and accounted for 42% of the variance in WAIS-R Full Scale IQs (FSIQs). Wrobel and Wrobel (1996) recently combined demographic (Barona et al., (1984) equations; Barona & Chastain (1986) equations) and personality variables (MMPI - Minnesota Multiphasic Personality Inventory; Hathaway & McKinley, 1943) to estimate WAIS-R IQs in patients with preexisting psychiatric disorders. Their goal was to develop predictive measures that would assist in detecting/quantifying neurological deterioration in patients

with preexisting psychiatric disorders.

Best Performance Method

An alternative approach to the estimation of premorbid ability has been the "best performance method", proposed by Lezak (1983, 1995). Lezak (1995) indicated that the "best performance method" relies on the level of best performance to serve as the best estimate of premorbid ability. Lezak (1995) considered that the level of best performance could be determined from test scores, observational data (i.e. behaviour), or historical data (i.e., premorbid achievement). Lezak (1995) suggested that once the highest level of functioning has been identified, it becomes the standard against which all other aspects of the patient's current performance are compared. Lezak (1995) however cautioned that the value of the "best performance method" depends on the judgement of the clinician in determining the basis and appropriateness of an estimate of premorbid ability when taking into account all available data. Mortensen, Gade, & Reinisch (1991) evaluated the utility of the "best performance method" at estimating WAIS IQs in both normal and clinical samples and reported that the "best performance method" (i.e., highest WAIS subtest score) resulted in the overestimation of WAIS IQs in both samples. Lezak (1995) however contended that Mortensen et al. (1991) misused the "best performance method" as they 1) selected

the highest WAIS subtest score to estimate WAIS IQs in a normal sample which would always result in an overestimation of WAIS IQ as the IQ score is essentially a mean score of subtest performance and 2) relied solely on the highest test score in a clinical sample without taking into account historical or behavioural data, or other test scores. Recently, some commonly used present ability measures of premorbid IQ have incorporated the "best performance method" into their approach to the prediction of premorbid intelligence (Vanderploeg, Schinka, & Axelrod, 1996; Paolo, Ryan, & Tröster, 1997; Scott, Krull, Williamson, Adams, & Iverson, 1997).

The Utility of Commonly Used Measures of Premorbid IQ at estimating WAIS/WAIS-R IQ scores

1. *Demographic Predictors*

Blair and 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) 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 (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.

Revision of the WAIS (WAIS-R; Wechsler, 1981) resulted in the development of new demographic regression equations to reflect performance on the WAIS-R (Barona, Reynolds, & Chastain, 1984). In addition to the original demographic variables utilized by Wilson et al. (1978), Barona et al. (1984) included geographic region and urban vs. rural residence as demographic predictors in their final equations. The variance accounted for between the final demographic variables and WAIS-R VIQ, PIQ, and FSIQ was 38%, 24%, and 36%, respectively. In 1986, Barona and Chastain attempted to improve the accuracy of the original WAIS-R demographic predictive equations and performed regression analyses on the WAIS-R standardization sample utilizing only Black and White participants between 20-74 years of age. This modification appeared to improve the predictive accuracy of the equations 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). However, Paolo and Ryan (1992) recently suggested that the 1984 equation may be slightly better at estimating WAIS-R IQs than the 1986 equation in a normal elderly sample, given that a greater proportion of IQ scores predicted by the 1984 equation fell within one standard error of obtained IQ scores, as compared to the

proportion of predicted IQs that fell within this range for the 1986 equation.

At present, there has been little attempt to employ or evaluate the utility of demographic predictors at estimating WAIS/WAIS-R IQ outside of North America (Crawford & Allan, 1997). Crawford, Stewart, Cochrane et al. (1989) used a sample of 151 healthy individuals to build regression equations to estimate premorbid WAIS IQ for the United Kingdom (UK). The regression models incorporated occupational classification, years of education, age and gender as predictors and accounted for 50%, 50%, and 30% of the variance in FSIQ, VIQ, and PIQ, respectively. Recently, Crawford and Allan (1997) developed demographic regression equations to predict WAIS-R IQ based on a sample of 200 healthy individuals that were considered to be representative of the adult UK population in terms of age, gender and occupational classification. Crawford and Allan (1997) found that occupation, education and age contributed significantly to the prediction of WAIS-R IQ and accounted for 53%, 53% and 32% of the variance in FSIQ, VIQ and PIQ, respectively. Crawford and Allan (1997) thus suggested that the demographic approach to the estimation of premorbid IQ has utility beyond the United States.

Cross-validation studies have revealed variable utility of demographic regression equations at estimating WAIS/WAIS-R IQs. Crawford (1989) has argued that appropriate cross-

validation studies of premorbid measures should be carried out with healthy, normal participants, as it is impossible to compare estimates of premorbid IQ with actual premorbid IQ in clinical samples, unless previous records exist. Studies with normal participants have revealed that demographic regression equations, including the Wilson et al. (1978) formula and the Barona et al. (1984) formula, have generally yielded similar mean predicted and mean obtained FSIQs for participants, but have tended to underestimate or overestimate IQ scores at the extremes of the WAIS/WAIS-R scales (Karzmark, Heaton, Grant, & Matthews, 1985; Goldstein, Gary, & Levin, 1986; Eppinger, Craig, Adams, & Parsons, 1987; Paolo & Ryan, 1992).

Paolo, Ryan, Tröster, and Hilmer (1996a) specifically examined the utility of the Barona et al. (1984) formula at estimating WAIS-R IQs by IQ range classification in a large normative group consisting of the WAIS-R (Wechsler, 1981) and elderly WAIS-R (Ryan, Paolo, & Brungardt, 1990) standardization samples. Paolo et al. (1996a) found that while the Barona et al. (1984) formula correctly estimated 70% of WAIS-R FSIQs for the normative sample within one standard error of estimate (SEe), significant prediction errors, defined as exceeding one SEe, occurred in over 50% of the normative group with IQs below 80 or above 120, and in one-third of the normative group with IQs between 80-89, and 110-119.

Helmes (1996) recently examined the utility of the Barona et al. (1984) formula at predicting IQ performance for 8663 normal elderly Canadians, ranging in age from 65-103 years, and recruited on a random sample basis from the Canadian Study of Health and Aging (CSHA; Canadian Study of Health and Aging Working Group, 1994). A modification for age was applied to the Barona formula to allow for the inclusion of very old elderly individuals (74 years+) in the study. Although direct cross-validation of results with actual WAIS-R IQs was not possible given that the sample had only been administered four subtests of the WAIS-R, results suggested that the demographic equations produced estimated IQs that closely approximated the expected mean of 100 IQ points for VIQ, PIQ, and FSIQ, for male and female participants up to their mid-late 80s in age. With increasing age, estimated IQs dropped by approximately 2-3 IQ points per five year age block in males and females. Although Helmes (1996) considered that the Barona formula may provide a reasonably accurate method for estimating IQ in older individuals, he cautioned that the use of demographic predictors leads to increased uncertainty regarding the accuracy of IQ estimation, and suggested that multiple methods of prediction may offer the best approach to determining premorbid psychometric intelligence.

Studies which have evaluated the utility of demographic IQ estimates at detecting cognitive decline in clinical

patients as compared to controls have generally found that while obtained IQs were significantly lower for neurologically impaired participants, there were no significant differences for estimated IQs (Barona et al. 1984 equation) between groups, suggesting that demographic equations could reveal deterioration from a higher functioning premorbid level in neurologically impaired participants (Eppinger et al. 1987; Paolo & Ryan, 1992; Paolo, Tröster, Ryan, & Koller, 1997). However, a comparative study which recently examined the efficacy of both the Wilson et al. (1978) formula and the Barona et al. (1984) formula in the prediction of concurrently obtained IQ levels in neurologically normal psychiatric patients and brain-damaged patients revealed that neither formula differed significantly in terms of classification of IQ range, with both formulas performing at essentially chance levels (Sweet, Moberg, & Tovian, 1990).

Demographically based regression equations have also recently been developed to estimate WAIS-R subtest scaled scores. Paolo, Ryan, Tröster, and Hilmer (1996b) combined the WAIS-R (Wechsler, 1981) and elderly WAIS-R (Ryan, Paolo, & Brungardt, 1990) standardization samples to develop regression equations to predict WAIS-R subtest scales scores utilizing demographic variables including age, education, gender, race, job, region and residence. Paolo et al. (1996b) found that less than 50% of the variance in the

eleven WAIS-R subtests was explained by demographic prediction equations, with variance accounted for ranging from 22% (i.e., Digit Span) to 48% (i.e., Digit Symbol). Paolo et al. (1996b) then examined the accuracy of the demographic equations at predicting scaled scores within ± 3 points (i.e., one standard deviation), and reported that the accuracy rate for demographic equations ranged from 76% (i.e., Arithmetic) to 89% (i.e., Vocabulary). Paolo et al. (1996b) considered that these findings suggested that demographic equations offered good estimates of actual subtest scaled scores in a normal sample. The clinical utility of these demographic equations at predicting subtest performance was also examined in a sample of 247 persons with confirmed brain damage/dysfunction. Paolo et al. (1996b) reported that significant mean differences emerged between estimated and actual scaled scores for persons with brain dysfunction and overall, one-fourth to one-third of persons with neurologic dysfunction evidenced possible deterioration in at least one subtest. However, Paolo et al. (1996b) cautioned that the judgement of a decline in WAIS-R subtest performance requires corroborating evidence (i.e., educational/occupational attainment) because the results of the study revealed that many normal persons had one or two subtests that evidenced possible decline, and the demographic equations tended to underestimate high subtest scores and overestimate low subtest scores.

Summary

In summary, demographically based regression equations have been developed to predict WAIS/WAIS-R IQs, and more recently WAIS-R subtest scaled scores, for North American and United Kingdom (i.e., WAIS/WAIS-R IQs only) populations. These measures remain attractive from a clinical perspective as estimators of premorbid IQ, as demographic variables are impervious to the effects of neurologic or psychiatric disorder. However, the above studies revealed that these measures demonstrated only moderate utility at estimating WAIS/WAIS-R IQs in normative groups, as demographic predictors accounted for at most 53% of the variance in Wechsler IQs (i.e., WAIS-R VIQ, UK sample; Crawford & Allan, 1997) and displayed significant errors in prediction for IQs below 80 or above 120. Given the questionable validity of demographic predictors at estimating WAIS/WAIS-R IQs in normal groups, it has been suggested that demographic variables not be used in isolation to detect cognitive decline in clinical samples.

2. *Measures of Present Ability*

Popular measures of present ability commonly used to estimate premorbid IQ include the Vocabulary subtest of the Wechsler scales (Crawford, 1989), and the National Adult Reading Test (NART; Nelson, 1982). As a testing instrument, the Vocabulary subtest has demonstrated high split-half

reliability and test-retest reliability (Matarazzo, Carmody, & Jacobs, 1980) across both the WAIS and WAIS-R, and has been identified as the single best measure of both verbal and general mental ability (Lezak, 1995), accounting for more than 70% of the variance in WAIS/WAIS-R FSIQs (Wechsler 1955, 1981). The clinical utility of the Vocabulary subtest as a predictive estimate of IQ has been questioned however, given that the Vocabulary subtest has not proven to be particularly resistant to the effects of neurologic or psychiatric dysfunction. Although various clinical groups have demonstrated compromised performance on the Wechsler Vocabulary subtest as compared to healthy controls, (Nelson & McKenna, 1975; Hart, Smith, & Swash, 1986; Crawford, Parker, & Besson, 1988; Crawford, Besson, Parker, Sutherland, & Keen, 1987; Sharpe & O'Carroll, 1991), some studies have suggested that it is worthwhile to consider the extent of decline of the Vocabulary subtest in relation to decline of other cognitive abilities in clinical groups. Such studies have noted that the Vocabulary subtest demonstrates greater preservation than other Wechsler subtests in the presence of cognitive decline (Whitehead, 1973; Crookes, 1974; Martin & Fedio, 1983), with some patient groups performing within the average range on the Vocabulary subtest when other cognitive skills are clearly compromised (Martin & Fedio, 1983; Mittenberg, Thompson, Schwartz, Ryan, & Levitt, 1991).

The National Adult Reading Test (NART) was the first measure developed to estimate premorbid IQ based on the observation that reading ability (accuracy of oral pronunciation) appeared to remain relatively well preserved in dementing individuals (Nelson & McKenna, 1975), and correlated highly with IQ in the normal population. The NART is a single-word oral reading test consisting 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 (i.e., ache, gauche) and in doing so measures previous familiarity with such words independent of present ability to analyze them as a complex visual stimulus (Wiens, Bryan, & Crossen, 1993). The NART was originally standardized in the United Kingdom (UK) against the WAIS and was recently re-standardized in the UK against the WAIS-R (Nelson & Willison, 1991).

The NART, as a testing measure, has demonstrated robust reliability, with internal reliability reported at .93 (Nelson & Willison, 1991), test-retest reliability reported at .98 (Crawford, Parker, Stewart, Besson & DeLacey, 1989), and inter-rater reliability ranging from .89 to .98 (O'Carroll, 1987; Crawford, Parker et al. 1989). Examination of the validity of the NART as a measure of intelligence in normal samples has however produced mixed results. Nelson (1982) found that the NART accounted for 55%, 59%, and 32%

of the variance in prorated WAIS FSIQ, VIQ, and PIQ scores, respectively, for the original standardization sample consisting of 120 neurologically normal participants. However, the reliability of this result is questionable as prorating of WAIS IQs has been reported as inadvisable (Wechsler, 1955). More recently, studies that have examined the utility of the NART at estimating WAIS/WAIS-R IQs in normal samples have reported that the NART has accounted for anywhere from 26% to 66% of the variance in WAIS/WAIS-R FSIQs, depending on the study (Crawford, Parker et al. 1989; Willshire et al. 1991; Sharpe & O'Carroll, 1991; Ryan & Paolo, 1992). Examination of the retrospective accuracy of the NART at predicting WAIS-R IQs obtained five years earlier in a normal elderly Canadian sample indicated that the NART only accounted for approximately 28% of the variance in WAIS-R FSIQs over five years (Carswell, Graves, Snow, & Tierney, 1997).

Revisions of the NART were developed to improve upon limitations of the original British form of the NART for North American populations (Wiens et al. 1993). These revisions including the NART-R/NAART (Blair & Spreen, 1989), and the AMNART (Schwartz & Saffran, 1987, as cited in Grober & Sliwinski, 1991) have remained consistent with the original NART in that the revised lists consist of words that cannot be phonetically decoded, but have either replaced unfamiliar British words with American words of

comparable frequency, or have allowed for scoring of NART words by North American pronunciation standards. Blair and Spreen (1989) originally examined the utility of the NART-R at estimating WAIS-R IQs in a North American sample and found that the NART-R accounted for 69%, 16%, and 56% of the variance in WAIS-R VIQ, PIQ, and FSIQ, respectively. However, recent cross-validation of the NART-R has produced variable results. Weins et al. (1993) reported that the NART-R accounted for only 31%, 0.05% and 21% of the variance in WAIS-R VIQ, PIQ, and FSIQ, respectively in a sample of 302 normal participants, whereas, Corrigan and Berry (1992) found that the NART-R accounted for 66% of the variance in WAIS-R VIQ in a sample of 60 neurologically intact older adults. Retrospective examination of the NART-R at predicting WAIS-R IQs obtained 3.5 years earlier in a normal elderly North American sample revealed that NART-R accounted for 46%, 37%, and 49% of the variance in WAIS-R VIQs, PIQs, and FSIQs, respectively (Berry et al. 1994). Recent examination of the temporal stability of the NART-R over a one year period of time for a sample of 51 normal elderly individuals revealed a test-retest reliability of .92 with the NART-R accounting for 53% of the variance in WAIS-R FSIQs at both testing periods (Raguet, Campbell, Berry, Schmitt, & Smith, 1996).

The American version of the NART (AMNART), developed by Schwartz and Saffran (1987; as cited in Grober & Sliwinski

1991) accounted for 52%, 26%, and 52% of the variance in WAIS VIQ, PIQ, and FSIQ scores, respectively, based on their standardization sample of 109 normal adults which ranged in age from 40-89 years. Grober, Sliwinski, Schwartz, and Saffran (1989; as cited in Grober & Sliwinski, 1991) further examined the utility of the AMNART as compared to the NART for predicting WAIS IQ in an American sample, and found that approximately 87% of their sample of normal adults made fewer reading errors on the AMNART as compared to the NART, supporting the use of the AMNART as a predictive measure of IQ in an American sample. Boekamp, Strauss, and Adams (1995) examined the validity of the AMNART at estimating WAIS-R VIQ scores in healthy African-American and White elderly veterans. The results suggested that the AMNART was equally useful at estimating WAIS-R VIQs for both ethnic groups. However, Boekamp et al. (1995) indicated that caution should be used when estimating verbal intelligence for individuals with lower intellectual ability because the AMNART overestimated VIQ scores for these participants across ethnic groups.

A short version of the NART has also been developed for individuals with poor reading skills. Beardsall and Brayne (1990) considered that administration of the original 50-item full-length NART could provoke anxiety/distress in poor readers, and thus developed a regression equation to predict the score on the second half of the NART (i.e., last 25

words) from the words pronounced correctly on the first half of the NART (i.e., first 25 words: Short NART). Beardsall and Brayne (1990) reported that the total NART predicted score (i.e., sum of actual score on the first half of the test and the predicted score on the second half of the test) accounted for 86% of the variance in actual total NART scores in a sample of 122 healthy elderly women. Beardsall and Brayne (1990) examined the performance of the above participants on both parts of the NART and considered that administration of the Short NART was appropriate for individuals that scored from 0-20 points on the first half of the test. Scores of 0-11 on the Short NART were considered to represent the total correct score for the test as participants were not found to add to their score by completing the second half of the test. For scores of 12-20, the regression equation was used to predict the score on the second half of the test and summed with the Short NART score for a total correct NART score. The complete administration of the NART was recommended for scores of more than 20 on the first half of the test because a high degree of variability in performance on the second half of the NART was often seen for participants with scores in this range.

Crawford, Parker, Allan, Jack, and Morrison (1991) examined the utility of the Short NART at predicting full-length 50-item NART scores and estimating WAIS IQ scores in a large cross-validation sample of 674 healthy adult

participants. Crawford et al. (1991) indicated that the Short NART accounted for 73% of the variance in full-length NART scores. The Short NART also accounted for 64%, 69% and 29% of the variance in WAIS FSIQ, VIQ, and PIQ scores respectively, and compared favourably with full-length NART estimated IQs.

Recently, the Reading subtest of the Wide Range Achievement Test-Revised (WRAT-R) has also been considered as a potential predictor of premorbid IQ. The Reading subtest of the WRAT-R is a 74 item word list that an individual must read aloud and correctly pronounce. Unlike the NART, the Reading subtest of the WRAT-R was designed to measure reading achievement and thus consists of both regular and irregular words. Studies which have examined the utility of the WRAT-R Reading subtest for predicting WAIS-R IQ in samples of normal adults have found that WRAT-R Reading subtest scores have accounted for 20% to 36% of the variance in WAIS-R FSIQs (Wiens et al. 1993; Kareken et al. 1995; Cooper & Fraboni, 1988).

Summary

In summary, measures of present ability commonly used to estimate premorbid IQ include vocabulary (Wechsler Vocabulary subtest) and reading test measures (NART; NART-R/NAART; AMNART; Short NART; WRAT-R). These measures have been selected as potential estimators of premorbid IQ as

both vocabulary and oral reading skills are considered to remain relatively resistant to neurologic and psychiatric impairment. The results of the above studies showed that present ability measures also correlated highly with WAIS/WAIS-R IQ in normal populations, with the Wechsler Vocabulary subtest accounting for over 70% of the variance in WAIS/WAIS-R FSIQs and reading test measures (i.e., NART) accounting for up to 66% of the variance in WAIS/WAIS-R FSIQs. The utility of the NART and other reading measures as potential estimates of premorbid IQ in clinical samples will be discussed in detail in an upcoming section addressing the limitations of premorbid IQ measures.

3. ***Combined Predictive Equations***

Efforts to enhance the potential predictive value of measures of present ability at estimating premorbid IQ have typically resulted in the combination of present ability measures with demographic variables or other present ability predictors. Crawford, Stewart, Parker, Besson, and Cochrane (1989) were the first to develop a regression equation based on the combination of the NART with demographic variables (sex, social class, age) that accounted for 73%, 78%, and 39% of the variance in WAIS FSIQ, VIQ, and PIQ scores, respectively. Furthermore, the combined equation accounted for more of the variance in WAIS IQ scores than the NART or demographics alone. Recently, Willshire et al. (1991)

developed a regression equation based on the combination of NART errors and demographic variables (education, sex) to predict prorated WAIS-R IQs. Willshire et al. (1991) indicated that their NART/demographic equation predicted 56% of the variance in WAIS-R FSIQs, and also accounted for more of the variance in WAIS-R FSIQs than the NART or demographics alone. Similar results were obtained on cross-validation of the equation with a sample of 104 normal controls aged 20-64 years, as the NART/demographic equation accounted for 46% of the variance in WAIS-R FSIQs.

North American revisions of the NART have also been considered in combination with demographic variables in an effort to improve predictive accuracy. Grober and Sliwinski (1991) attempted to improve the predictive accuracy of the AMNART at estimating WAIS-R VIQ in a sample of non-demented elderly individuals, by combining AMNART errors and years of education in a predictive regression equation. Double cross-validation demonstrated that the model had high accuracy and stability in estimating current VIQ in non-demented elderly individuals.

Blair and Spreen (1989) originally examined the combined influence of demographic variables (age, sex, education, occupation) and NART-R scores at predicting WAIS-R IQs, but found that none of the demographic variables significantly improved NART-R prediction of WAIS-R IQs in their standardization sample. Blair and Spreen (1989)

considered that the limited variability of education and occupational status in their standardization sample may have reduced the predictive accuracy of demographic variables. More recently, Raguet et al. (1996) examined the combined predictive utility of demographic variables (i.e., Barona & Chastain 1986 estimates) and NART-R scores at estimating WAIS-R FSIQs in a sample of 51 healthy elderly participants and found that demographic variables significantly improved NART-R prediction of WAIS-R IQs. Raguet et al. (1996) reported that the NART-R/Barona & Chastain (1986) estimates accounted for 59% of the variance in WAIS-R FSIQs and demonstrated greater accuracy at predicting FSIQs than the NART-R or Barona & Chastain (1986) estimates alone.

Kareken, Gur, and Saykin (1995) attempted to improve the predictive potential of the WRAT-R Reading subtest as an estimate of premorbid IQ by combining it with demographic variables including race and parental education. Kareken et al. (1995) reported that the WRAT-R Reading subtest/demographic equation predicted 67%, 62%, and 72% of the variance in WAIS-R VIQ, PIQ, and FSIQ, respectively, and accounted for more of the variance in WAIS-R IQ scores than the reading subtest alone.

Renewed interest in the Wechsler subtests as potential predictors of IQ has resulted in the combination of Wechsler subtests with other measures of present ability and demographic variables. Nelson and McKenna (1975) built the

first regression equation utilizing Vocabulary age-scaled scores to predict prorated WAIS IQs. Carswell et al. (1997) created a regression equation that combined NART errors and age-scaled WAIS-R Vocabulary scores to predict/postdict WAIS-R VIQ scores obtained five years earlier in a normal elderly sample. Results indicated that the NART/Vocabulary regression equation predicted 49% of the variance in WAIS-R VIQs, and accounted for more of the variance in WAIS-R VIQs than Vocabulary or the NART alone. Krull, Scott, and Sherer (1995) employed various combinations of the Vocabulary subtest raw score, Picture Completion subtest raw score, and demographic variables (age, education, occupation, race) to predict IQ scores for the WAIS-R standardization sample. Three equations were created to predict WAIS-R VIQ, PIQ, and FSIQ scores, and accounted for 76%, 61%, and 76% of the variance in each IQ measure, respectively. The Krull et al. (1995) equations represented a method of premorbid IQ estimation referred to as the Oklahoma Premorbid Intelligence Estimation (OPIE). Vanderploeg and Schinka (1995) also developed regression formulas to predict IQ based on the WAIS-R standardization sample, utilizing WAIS-R subtests and demographic variables (age, race, sex, education, occupation). All of the eleven WAIS-R subtests were considered individually and combined with demographic variables to predict VIQ, PIQ, and FSIQ scores, resulting in the development of 33 regression formulas. Vanderploeg and

Schinka (1995) indicated that the new regression equations generally doubled the amount of IQ variance accounted for by demographic variables alone.

Recently, some combined predictive equations have incorporated Lezak's (1983, 1995) "best performance method" in their approach to premorbid IQ prediction. Scott, Krull, Williamson, Adams, and Iverson (1997) examined the utility of the Krull et al. (1995) WAIS-R subtest/demographic equations (i.e., OPIE) at estimating premorbid IQ in several neurological samples. Since the Krull et al. (1995) equations include present ability measures (i.e., WAIS-R Vocabulary subtest and Picture Completion subtest raw scores) that may be selectively compromised in clinical samples, Scott et al. (1997) created a decision rule to circumvent this problem which combined demographic variables with the best performance on either the Vocabulary or Picture Completion subtest (non-age-corrected scaled score) for a WAIS-R FSIQ best estimate. Scott et al. (1997) reported that the FSIQ best formula offered the best estimate of premorbid ability for patients with diffuse cerebral dysfunction or left lateralized lesions, whereas the Vocabulary/demographics equation was the best FSIQ prediction formula for patients with right hemisphere lesions.

Vanderploeg, Schinka, and Axelrod (1996) also recently employed Lezak's (1983, 1995) "best performance method" to

determine which of the 33 WAIS-R subtest/demographic formulas developed by Vanderploeg and Schinka (1995) might best predict premorbid functioning. Vanderploeg et al. (1996) applied two best performance approaches to the original regression equations based on the WAIS-R standardization sample and defined the BEST 11 approach as the highest predicted score from all eleven FSIQ, VIQ, and PIQ equations, whereas, the BEST 3 approach was defined as the highest predicted score from three FSIQ, VIQ, and PIQ equations that included a reliable subtest considered to be a relatively good "hold" measure (i.e., Information, Vocabulary, Picture Completion). Predicted scores from the BEST 11 and BEST 3 approaches were compared to actual WAIS-R IQ scores and the results revealed that the BEST 3 approach appeared to best parallel actual IQ scores in the WAIS-R standardization sample, although both the BEST 3 and BEST 11 approaches resulted in overestimation of actual IQ scores by about 5 and 9 points, respectively.

Paolo, Ryan, and Tröster (1997) examined the utility of the Vanderploeg and Schinka (1995) BEST 3 equations at estimating WAIS-R IQs in the elderly WAIS-R standardization sample (Ryan et al. 1990). Paolo et al. (1997) reported that the Vanderploeg and Schinka (1995) BEST 3 equations significantly underestimated the obtained IQs of healthy elderly individuals. Paolo et al. (1997) subsequently developed new regression equations (i.e., BEST 3 approach;

Vanderploeg et al. 1996) for predicting WAIS-R IQs in persons 75 years and older based on a sample of 130 healthy elderly persons with independent predictors including race, gender, socio-economic status (SES; based on a combination of education and occupation) and age-corrected WAIS-R subtest scores. These equations accounted for 23% to 77% of the variance in actual WAIS-R IQs in a cross-validation sample (N = 95) and were considered to be more appropriate than the Vanderploeg and Schinka (1995) equations for predicting premorbid IQ in persons 75 years and older (Paolo et al. 1997).

Summary

In summary, combined predictive equations utilizing demographic and present ability variables have consistently offered the best estimates of WAIS/WAIS-R IQs in normative samples, as compared to demographic or present ability measures alone. Such measures have accounted for as much as 76% of the variance in WAIS-R FSIQs in normal populations (WAIS-R subtest/demographics; Krull et al. 1995).

Limitations of Existing Premorbid IQ Measures

Although premorbid IQ measures have attempted to fulfill the criteria specified by Crawford (1989) to serve as valid means of estimating premorbid IQ, many are constrained by methodological and clinical limitations, some

of which have already been reviewed. The following discussion will focus largely on the limitations of reading measures at estimating premorbid IQ, given that these measures have recently gained popularity in clinical neuropsychological practise due to the relative ease of administration of such measures (O'Carroll, 1995). However, the limitations identified can be considered to apply to demographic predictors, present ability measures, and combined predictive equations, and will also be referred to in the discussion, if appropriate.

1. *Limitations of Premorbid IQ Measures with Normative Samples*

(a) Lack of Normative Data for Elderly Individuals

Few studies exist which have examined the predictive utility of measures of reading ability (i.e., NART) at estimating WAIS/WAIS-R IQs in samples consisting exclusively of normal elderly participants with a mean age of 50 years or more. Although this is not unusual given the general inadequacy of older age norms for the majority of commonly used neuropsychological measures (Naugle, Cullum, & Bigler, 1990), it is somewhat surprising given that reading measures were originally considered as potential predictors of premorbid IQ based on the observation that reading ability appeared to be relatively preserved in elderly dementing

individuals (Nelson & McKenna, 1975), suggesting that normal elderly individuals would serve as the most appropriate control group for such cases. A normative data base for elderly individuals would not only demonstrate the utility of reading measures at estimating WAIS/WAIS-R IQs in elderly individuals across a range of intellectual functioning, but would also offer a reliable means of determining the normal error range for predicting premorbid IQ using reading measures which would improve identification of cognitive decline in elderly individuals. Establishing normative data for elderly individuals performing at the extremes of the normal distribution with regard to intellectual functioning is particularly useful for high functioning individuals given that cognitive decline is often difficult to detect with such individuals as compromised performance may still fall within the average range due to a high premorbid level of functioning (Naugle et al. 1990).

Most studies, including both standardization and cross-validation investigations, have generally examined the utility of reading measures at predicting WAIS/WAIS-R IQs for normal adults, across a broad age range, with a mean age of less than 50 years. The NART was originally standardized on a sample of 120 neurologically normal individuals from the UK, that ranged in age from 20-70 years, with a mean sample age of 48 years (Nelson, 1982). Similarly, the NART-R standardization sample failed to adequately represent

elderly North American individuals as participants ranged in age from 18-49 years with a mean sample age of 27.4 years (Blair & Spreen, 1989). The American version of the NART (AMNART) however provided a more adequate representation of elderly American individuals as it was standardized on 109 normal adults that ranged in age from 40-89 years (Schwartz & Saffran, 1987; as cited in Grober & Sliwinski, 1991).

Cross-validation studies examining the predictive utility of reading measures have similarly been compromised by a lack of normative data with healthy elderly participants. Four studies have evaluated the utility of the NART at predicting WAIS-R IQ in samples consisting exclusively of normal elderly participants (Sharpe & O'Carroll, 1991; Willshire et al. 1991; Ryan & Paolo, 1992; Carswell et al. 1997). Sample sizes have ranged from 20-126 normal participants (Sharpe & O'Carroll, 1991; Ryan & Paolo, 1992, respectively), and the mean age of participants has ranged from 63.7 to 80.8 years (Willshire et al. 1991; Ryan & Paolo, 1992, respectively). A recent cross-validation of the Ryan and Paolo (1992) NART regression equations with a sample of 49 normal elderly individuals (mean age = 71 years) revealed similar predictive results suggesting that the NART regression equations may be potentially useful for estimating WAIS-R IQs in elderly samples (Carswell et al. 1997).

Fewer studies have investigated the utility of other

reading measures at estimating WAIS/WAIS-R IQs in normal elderly samples. Three American studies have examined the ability of the NART-R to predict WAIS-R IQ in normal elderly samples concurrently (Corrigan & Berry, 1992), and over 1 year and 3.5 year time periods, respectively (Raguet et al. 1996; Berry et al. 1994). Boekamp et al. (1995) examined the validity of the AMNART at estimating WAIS-R VIQ in healthy elderly African-American and White veterans. Studies which have examined the utility of the WRAT-R Reading subtest in predicting WAIS/WAIS-R IQs in normal adults are limited (Wiens et al. 1993; Cooper & Fraboni, 1989; Kareken et al. 1995), and although these studies generally examined a sample of normal participants who ranged in age from 19-58 years, none of the studies focused on an exclusively elderly sample. To date, there are also no published studies which have examined the utility of the Reading subtest of the most recent revision of the WRAT (WRAT-3; Wilkinson, 1993) at predicting WAIS-R IQ in normal elderly samples.

(b) ***Lack of Normative Data for Elderly Canadian Individuals***

As the NART was originally developed and standardized in the United Kingdom, the majority of studies investigating the utility of the NART at predicting WAIS/WAIS-R IQ scores have been based on United Kingdom participant samples. American, Canadian and Australian studies have also examined the utility of the NART at predicting WAIS/WAIS-R IQs in

their respective samples, although these studies have been far fewer in number. Overall, Canadian samples have been fairly limited as even standardization samples for North American revisions of the NART have included primarily American participants. The NART-R standardization data was based on a mixed Canadian/American sample (Blair & Spreen, 1989), while a pure American sample was used for standardization of the AMNART (Schwartz & Saffran, 1987; as cited in Grober & Sliwinski, 1991). Only two studies have examined the predictive utility of reading measures at estimating WAIS-R IQs in a sample of normal elderly Canadian individuals. Sharpe and O'Carroll (1991) evaluated the utility of the NART at estimating concurrent WAIS-R IQ scores in a normal elderly sample, while Carswell et al. (1997) examined the retrospective accuracy of the NART at predicting WAIS-R IQs obtained five years earlier for healthy elderly participants. Carswell et al. (1997) improved the accuracy of the NART at postdicting WAIS-R VIQs obtained five years earlier in the same normal sample, by developing a regression equation combining NART errors with WAIS-R Vocabulary scores. To date, no studies have examined the predictive utility of other measures of reading ability (i.e., NART-R) at estimating WAIS-R IQs in a normal elderly Canadian sample.

Recently, Helmes (1996) examined the utility of the Barona et al. (1984) equation at predicting premorbid

intelligence in a large sample of healthy elderly Canadians (N=8663) recruited as part of the Canadian Study of Health and Aging (CSHA; Canadian Study of Health and Aging Working Group, 1994). Helmes (1996) indicated that the demographic prediction equations showed estimated IQs closely approximating the expected mean of 100 IQ points up to the mid-late 80s in males and females. Cross-validation with actual WAIS-R IQs was not possible given that participants had only been administered four subtests of the WAIS-R. Naugle et al. (1990) cautioned that use of demographic procedures in estimating premorbid IQ among the elderly may be highly inaccurate due to individual variability in intellectual functioning among individuals matched with regard to demographic variables, and due to potentially lower correlations between IQ and educational level for elderly individuals as many older individuals discontinued their schooling prematurely not due to limited innate abilities, but due to external factors such as economic need, or military, or family responsibilities.

(c) Lack of Studies Examining the Comparative Utility of Premorbid IQ measures in Normal Elderly Samples

Since there are a limited number of studies which have examined the utility of measures of reading ability at predicting WAIS/WAIS-R IQs in normal elderly samples, it is not unexpected to find few studies which have compared the

utility of different premorbid IQ measures in the same sample of normal elderly participants. Grober, Sliwinski, Schwartz, and Saffran (1988; as cited in Grober & Sliwinski, 1991) administered both the AMNART and the NART to a sample of 40 normal American adults and found that 35 of them made fewer errors on the AMNART as compared to the NART. Wiens et al. (1993) examined the utility of the NART-R and WRAT-R Reading subtest at predicting WAIS-R IQs in a normal American sample ranging in age from 20-54 years, and reported that the NART-R and the WRAT-R Reading subtest accounted for essentially the same proportion of variance (20%) in WAIS-R FSIQS. Corrigan and Berry (1992) compared the utility of the Barona equations (Barona et al. 1984; Barona & Chastain, 1986) and the NART-R at predicting WAIS-R IQs in a neurologically intact sample of older adults. Corrigan and Berry (1992) found that the NART-R outperformed demographic estimation methods in the prediction of WAIS-R VIQ with the NART-R accounting for 66% of the variance in WAIS-R VIQ scores whereas, the Barona equations accounted for a maximum of 37% of the variance in WAIS-R VIQs. Raguet et al. (1996) also documented the superiority of the NART-R at predicting WAIS-R IQs as compared to other estimation methods. Raguet et al. (1996) reviewed the one year temporal stability of WAIS-R FSIQ and several predictors of premorbid IQ including the Barona and Chastain (1986) demographic equation, the NART-R, and the Intellectual Correlates Scale

(ICS; Schlottmann & Johnsen, 1991) in an exclusively elderly sample of 51 healthy American individuals. Results indicated that the NART-R accounted for proportionately more of the variance (53%) in WAIS-R FSIQs at both testing periods than the Barona and Chastain (1986) equation or the ICS score, with the Barona et al. (1986) equation and the ICS score only accounting for approximately 38% and 20% of the variance in WAIS-R FSIQs across the two testing trials, respectively.

Studies which have developed regression equations for predicting premorbid IQ utilizing different predictors including demographic variables and measures of present ability, have generally compared the utility of each premorbid IQ measure alone and in combination with other predictors in the same sample (NDE/Crawford et al. 1989; AMNART and demographics/Grober & Sliwinski, 1991; NART and demographics/Willshire et al. 1991; NART-R and demographics (Barona & Chastain 1986 equation)/Raguet et al. 1996; WRAT-R and demographics/Kareken et al. 1995; NART and WAIS-R Vocabulary/Carswell et al. 1997; WAIS-R subtests and demographics/Krull et al. 1995; Vanderploeg & Schinka, 1995; Paolo et al. 1997). These studies have typically reported that the combined predictive equation demonstrates greater utility in estimating WAIS/WAIS-R IQ scores than either of the predictive measures, independently. Five of these studies have examined the utility of their respective

premorbid regression equations at predicting WAIS-R IQs in exclusively elderly samples (Grober & Sliwinski, 1991; Willshire et al. 1991; Raguet et al. 1996; Carswell et al. 1997; Paolo et al. 1997).

Only two studies have compared the predictive utility of their respective premorbid regression equations to the Barona et al. (1984) formula for estimating WAIS-R IQs in exclusively normal elderly samples. Grober and Sliwinski (1991) reported that the predictive power of their regression equation which combined AMNART errors with education accounted for proportionately more of the variance in WAIS-R VIQs than the Barona (1984) formula with healthy elderly participants. Paolo et al. (1997) also compared the predictive utility of their regression equations which combined WAIS-R age-corrected subtest scores and demographics to the Barona (1984) formula and provided results that indicated that their equations accounted for more of the variance in WAIS-R IQs than the Barona (1984) formula in two samples of healthy elderly participants.

Other studies have compared the predictive utility of their respective premorbid regression equations to the Barona et al. (1984) formula for estimating WAIS-R IQs in normal samples that include, but are not exclusively elderly, or in clinical samples which combine healthy and impaired elderly individuals. Vanderploeg et al. (1996) recently compared the highest predicted IQ scores from their

BEST 11 and BEST 3 WAIS-R subtest/demographic based regression equations with the Barona (1984) formula for the WAIS-R standardization sample (i.e., ages 16-74 years) and reported that the BEST 11 and BEST 3 predicted scores resulted in significantly higher correlations with actual WAIS-R IQ scores than the Barona (1984) formula. Paolo, Tröster, Ryan & Koller (1997) examined the predictive utility of the NART and the Barona (1984) formula at estimating short form (i.e., seven subtest abbreviation) WAIS-R IQs in a combined sample of healthy elderly controls and persons with Alzheimer's disease. The results revealed that the NART accounted for 37% of the variance in short form WAIS-R VIQs whereas, the Barona (1984) formula only accounted for 18% of the variance in short form WAIS-R VIQs for the mixed normal/clinical sample.

A potential drawback in evaluating the comparative utility of different premorbid IQ measures is that not all measures have been based on Wechsler IQ, and those that have may either predict WAIS or WAIS-R IQ. For example, alternative measures of present ability which have recently demonstrated potential utility at predicting premorbid IQ have included the Cambridge Contextual Reading Test (CCRT; Beardsall & Huppert, 1994) and the Spot-the-Word test (Baddeley et al. 1993). The CCRT is based on placing NART words in the context of meaningful sentences to facilitate semantic retrieval, and the Spot-the-Word test is a lexical

decision task in which an individual is presented with two items and is required to state which of the pair is a real word (O'Carroll, 1995). Both measures have been administered to normal elderly individuals in the UK, although neither measure has been based on estimating WAIS/WAIS-R IQs. The predictive utility of the Spot-the-Word test in estimating premorbid intellectual functioning was considered favourable due to high correlations of the test with measures of verbal intelligence including the Mill Hill Vocabulary Scale ($r=.86$) and the NART ($r=.87$). The CCRT was considered to provide a more accurate estimate of reading ability and IQ in elderly and demented individuals than the NART, given that the pronunciation of NART words significantly improved when placed in the context of sentences, for both normal and demented readers (Beardsall & Huppert, 1994; Beardsall & Huppert, 1997). Beardsall and Huppert (1994) noted that the improvement in pronunciation of NART words was most marked for demented participants and for poor or average readers as compared with skilled readers. Conway and O'Carroll (1997) also found that the CCRT was most useful for reducing overall pronunciation errors of NART words for more severely cognitively impaired Alzheimer's patients as compared to less cognitively impaired patients, and agreed with Beardsall and Huppert (1994) that the format of the NART could lead to underestimation of reading ability and consequent underestimation of premorbid IQ in certain

patient groups.

Although some predictive estimates of IQ are not based on Wechsler IQ, and others continue to predict WAIS IQ, most have undergone the appropriate revisions to predict WAIS-R IQ. This has allowed for reasonable comparative evaluation of different premorbid IQ measures. However, given the recent release of the newest version of the Wechsler scales (WAIS-III; Wechsler, 1997), existing measures of premorbid IQ will once again have to update their normative data base to determine the utility of each measure in estimating WAIS-III IQs.

Summary

In summary, limitations of premorbid IQ measures with normative samples include a general lack of normative data for elderly individuals, with few studies examining the utility of premorbid IQ measures for elderly Canadians. Review of studies which examined the comparative utility of premorbid IQ measures at estimating WAIS/WAIS-R IQs in normal elderly samples showed that reading test measures outperformed demographic predictors, and combined predictive equations outperformed reading test or demographic predictors alone. Of the combined predictive equations, the WAIS-R subtest/demographic equations developed by Paolo et al. (1997) offered some of the best estimates of WAIS-R IQ in a normal elderly sample. Other potentially useful

measures of premorbid IQ that have not yet been normed on the WAIS/WAIS-R but have been examined in normal elderly samples were also discussed, including the Spot-the-Word test and the Cambridge Contextual Reading Test.

2. *Methodological Constraints of Premorbid IQ Measures*

(a) Lack of Clinical Sensitivity of Premorbid IQ Measures

Throughout the literature pertaining to the assessment of premorbid IQ, and thus reviewed above, the approach used most frequently for reporting the concurrent validity of a potential predictive IQ measure at estimating WAIS/WAIS-R IQ has been the correlation between the predictive IQ measure and actual WAIS/WAIS-R IQ scores, and the subsequent amount of variance accounted for by the predictive IQ measure in actual WAIS/WAIS-R IQ scores. This approach has been widely adopted as reflective of the "goodness" of WAIS/WAIS-R IQ prediction of a potential IQ measure, with measures accounting for a greater proportion of the variance in actual WAIS/WAIS-R IQ scores considered to be potentially more clinically useful as estimates of premorbid IQ. However, this approach provides, at best, only limited information as to the validity of such predictive IQ measures at estimating WAIS/WAIS-R IQ in normal samples, without any indication as to the sensitivity of such

measures at providing a reliable estimate of premorbid IQ or detecting cognitive decline in clinical samples.

Recently, Graves, Carswell, and Snow (in press) evaluated the sensitivity of various measures of premorbid IQ for detecting cognitive decline in clinical samples. As cognitive decline in clinical samples is often inferred if there is a significant discrepancy in scores between predicted IQ performance and actual WAIS/WAIS-R IQ scores, Graves et al. (in press) were interested in determining the expected discrepancy between various measures of premorbid IQ and WAIS/WAIS-R IQ for normal participants. Given the expected discrepancy between scores for normal controls based on the underlying psychometric properties of the tests considered, clinical impairment could be reliably inferred if the discrepancy in a patient's performance fell <u>outside</u> the range of expected discrepancy for normal controls. Graves et al. (in press) reviewed 22 studies that reported results for various methods of predicting WAIS/WAIS-R IQs for normal individuals including such methods as demographic regression equations and reading test estimates. The potential clinical utility of these methods was evaluated by using psychometric theory to calculate the magnitude of cognitive decline that could be reliably detected. The standard error of estimate was the statistic used to determine the sensitivity of premorbid IQ measures for detecting cognitive decline in clinical samples. Results

indicated that for a cognitive decline in WAIS-R VIQ to be detected 80% of the time, the decline would have to be at least 25 VIQ points for demographic prediction methods, and at least 20 VIQ points for reading test prediction methods. Premorbid IQ measures which demonstrated an expected discrepancy of more than 20 VIQ points between predicted and actual VIQ scores for normal controls were considered to be of limited clinical utility given that a patient would show obvious impairment if his/her score fell outside this range, rendering the IQ prediction superfluous. As few predictors fell at or under this 20 VIQ point index of cognitive decline sensitivity, the analysis highlighted the limited sensitivity of many of these predictive IQ methods for reliably detecting cognitive decline in clinical samples.

(b) Retrospective Accuracy of Premorbid IQ Measures

A second methodological issue that requires consideration is the relative lack of information regarding the retrospective accuracy of premorbid IQ measures. Most studies that have examined the validity of premorbid IQ measures at estimating WAIS/WAIS-R IQ scores in normal samples have administered the predictive IQ measure and the WAIS/WAIS-R at the same testing period. Graves et al. (in press) considered this procedure to be appropriate as the clinical application of these predictive methods generally involves concurrent testing. However, Graves et al. (in

press) indicated that information regarding the retrospective accuracy of a premorbid IQ measure (i.e., how well the predictive method can postdict an actual IQ score obtained several years prior to the estimator score) would help establish the validity of the assumption that estimation of current IQ is, in fact, equivalent to the postdiction of previous IQ.

Few studies have examined the postdictive ability of premorbid IQ measures. Berry et al. (1994) examined the ability of the NART-R to postdict WAIS-R IQ scores obtained 3.5 years earlier in a normal elderly sample. Berry et al. (1994) reported that the NART-R accounted for 46% of the variance ($r = 0.68$) in WAIS-R VIQ scores and produced a standard error of estimate (SEe) of 8.9 VIQ points. The resulting cognitive decline sensitivity score based on the SEe method described by Graves et al. (in press) (i.e., SEe multiplied by 2.487) was 22.1 VIQ points for the postdiction of VIQ.

Raguet et al. (1996) examined the retrospective accuracy of the Barona and Chastain (1986) demographics equation and the NART-R at postdicting WAIS-R IQ scores obtained one-year earlier in a normal elderly sample. Raguet et al. (1996) reported that the Barona and Chastain (1986) formula accounted for 37% of the variance ($r = 0.61$) in WAIS-R FSIQs, whereas the NART-R accounted for 49% of the variance in WAIS-R FSIQs ($r = 0.70$). Raguet et al. (1996)

reported that the SEe for the Barona and Chastain (1986) formula and the NART-R were 10.8 and 9.7 FSIQ points, respectively. The resulting cognitive decline sensitivity scores based on the SEe method described by Graves et al. (in press) were 26.9 FSIQ points for the Barona and Chastain (1986) formula and 24.1 FSIQ points for the NART-R.

Carswell et al. (1997) examined the ability of the NART to postdict WAIS-R VIQs obtained five years earlier in a normal elderly sample and found that the NART accounted for 29% of the variance ($r = 0.54$) in WAIS-R VIQs. Efforts to improve the postdictive accuracy of the NART resulted in the development of a regression equation utilizing NART errors and WAIS-R Vocabulary scores to postdict WAIS-R VIQs obtained five years earlier in the same elderly sample. Results indicated that the new NART/Vocabulary regression equation accounted for 49% of the variance ($r = 0.70$) in WAIS-R VIQs and produced a standard error of estimate of 6.85 VIQ points. The corresponding cognitive decline sensitivity was calculated at 17.0 VIQ points (Graves et al. in press).

Carswell et al. (1997) also undertook a postdictive cross-validation of the Ryan and Paolo (1992) NART regression equations by examining the utility of these equations at providing a postdictive estimate of WAIS-R VIQ obtained five years earlier in the previously mentioned normal elderly sample. Results indicated that the Ryan and

Paolo (1992) NART regression equation accounted for 30% ($r = 0.55$) of the variance in WAIS-R VIQ scores and produced a standard error of estimate of 7.92 VIQ points. The corresponding cognitive decline sensitivity was calculated at 19.7 VIQ points (Graves et al. in press).

Smith, Bohac, Ivnik, and Malec (1997) evaluated the postdictive ability of Grober and Sliwinski's (1991) AMNART/education equation at estimating Mayo normed WAIS-R VIQs obtained approximately 3.7 years earlier in a normal elderly sample. Mayo normed WAIS-R VIQs were age-corrected WAIS-R VIQ scores for normal elderly participants from the Mayo Older American Normative Studies (MOANS; Ivnik et al. 1992). Smith et al. (1997) reported that while the predicted VIQ scores based on the AMNART/education equation correlated well with prior obtained Mayo Verbal IQ scores ($r = 0.70$), the predicted scores tended to overestimate obtained Mayo VIQs by an average of 5.9 points. Smith et al. (1997) considered that the discrepancy may have been due to the fact that the Grober and Sliwinski (1991) AMNART/education equation was developed based on conventionally normed WAIS-R VIQ values, whereas Mayo VIQ scores were used in their study. Smith et al. (1997) subsequently developed a new regression equation to estimate Mayo Verbal IQs based on AMNART errors and education which accounted for 49% of the variance ($r = 0.70$) in obtained Mayo VIQs and produced a standard error of estimate of 6.9 VIQ points. The resulting

cognitive decline sensitivity based on the SEe method proposed by Graves et al. (in press) was calculated at 17.2 Mayo VIQ points. The above studies suggested that only a few estimates of premorbid IQ showed sufficient retrospective accuracy to reliably detect a 20 point decline in WAIS-R IQ or Mayo normed WAIS-R IQ.

(c) Prediction of Extreme versus Median IQ Scores

A significant limitation of many methods of premorbid IQ prediction is the restricted range of scores produced by such measures, resulting in the underestimation of performance for individuals with high premorbid abilities, and overestimation of performance for persons with lower premorbid ability. Many studies that have examined the utility of premorbid IQ measures at estimating WAIS/WAIS-R IQ scores in normative samples have identified serious prediction errors in estimating IQ scores outside of the median range (Barona et al. 1984; Blair & Spreen, 1991; Wiens et al. 1993; Kareken et al. 1995). Paolo, Ryan, Tröster, and Hilmer (1996a) specifically examined the utility of the Barona et al. (1984) formula at estimating WAIS-R IQs by IQ range classification in a large normative group (i.e., WAIS-R (Wechsler, 1981) and elderly WAIS-R (Ryan et al. 1990) standardization samples. Results indicated that significant errors in prediction, defined as exceeding one standard error of estimate, occurred in over

50% of the normative group with IQs below 80 or above 120, and in one-third of the normative group with IQs between 80-89, and 110-119. Such results led to the recommendation by Graves et al. (in press) that current methods of premorbid IQ prediction should not be used for persons suspected to fall outside of the 80-120 IQ range premorbidly. However, it is encouraging to note that Scott et al. (1997) recently identified one predictive method that demonstrates potential utility outside of this restricted range. Scott et al. (1997) examined the range in predicted WAIS-R FSIQ scores in normative samples for the Krull et al. (1995) equations (i.e., OPIE), the NART-R (Wiens et al. 1993), the Barona et al. (1984) demographic equation (Eppinger et al. 1987) and the "best performance method", (Mortensen et al. 1991). Scott et al. (1997) reported that the OPIE produced a less restricted range of predicted WAIS-R FSIQ scores (i.e., 73.3 - 125.3) than other estimation methods without systematic over-or-underestimation of WAIS-R FSIQs.

Summary

In summary, the clinical utility of premorbid IQ measures is restricted by methodological contraints. Graves et al. (in press) indicated that the sensitivity of premorbid IQ measures for detecting cognitive decline in clinical samples is limited, given that a decline of at least 20 and 25 VIQ points is required for reading test and

demographic prediction methods respectively, to reliably detect a decline in WAIS-R VIQ 80% of the time. The above studies also revealed that few estimates of premorbid IQ showed sufficient postdictive/retrospective accuracy to reliably detect a 20 point decline in WAIS-R IQ. Finally, premorbid IQ measures appear to be restricted to estimating median IQ scores, because significant errors in prediction have been documented for IQ scores below 80 or above 120.

3. *Limitations of Premorbid IQ Measures with Clinical Samples*

For a present ability measure to demonstrate utility in predicting premorbid IQ in a clinical sample, it must remain resistant to the effects of neurologic and psychiatric dysfunction. The NART was originally developed to predict premorbid functioning in a dementing population based on the observation that reading ability remained relatively preserved in dementing disorders (Nelson, 1982). To determine if in fact the NART remains a valid estimate of premorbid IQ in a dementing population, performance on the NART and WAIS/WAIS-R of dementing 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 dementing population as compared to controls, 2) NART predicted IQ is significantly higher than WAIS/WAIS-R obtained IQ in the

dementing 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 the dementing disorder and effective in providing an accurate measure of premorbid IQ (Crawford, 1989). Such steps must be taken to insure that the pronunciation of irregular words genuinely remains unaffected by dementing disorders. For the NART to be considered useful as a measure of premorbid IQ in other clinical conditions (i.e., traumatic brain injury, schizophrenia), the above observations/comparisons must be made to determine if reading ability on the NART remains unaffected by such conditions. O'Carroll (1995) has indicated that the NART has unfortunately been used in a "bandwagon" approach for estimating premorbid IQ across a variety of different clinical conditions. These studies have rarely determined whether reading performance remains preserved in their respective clinical conditions. The following discussion will examine the utility of the NART and other measures of present ability at predicting premorbid IQ in clinical samples.

(a) Alzheimer's Disease

Nelson and O'Connell (1978) initially reported that patients with cortical atrophy did not differ significantly from healthy controls on NART performance. Subsequent studies have similarly revealed no difference in NART

performance between Alzheimer's patients and normal controls at initial testing (Nebes, Martin, & Horn, 1984; O'Carroll & Gilleard, 1986; Crawford et al. 1989; Sharpe & O'Carroll, 1991), or over a one year period of time (O'Carroll, Baikie, & Whittick, 1987). Cummings, Houlihan, and Hill (1986) found no relation between NART performance and dementia severity for patients with Alzheimer's disease. Grober and Sliwinski (1991) also reported that estimated VIQ scores for a sample of mildly demented participants based on a regression equation combining AMNART errors and years of education, did not differ significantly from predicted VIQ scores for nondemented individuals.

However, the NART's resistance to the effects of cognitive decline associated with Alzheimer's disease has recently been challenged. Various studies have revealed that NART performance in Alzheimer's patients is ***significantly poorer than that of controls*** (Fromm, Holland, Nebes, & Oakley, 1991; Brayne & Beardsall, 1990; Hart et al. 1986), ***becomes progressively poorer over time*** (Fromm et al. 1991; Paque & Warrington, 1995), ***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; Patterson, Graham, & Hodges, 1994), and may even be ***suspect in the earliest stages of the disease*** (Storandt, Stone, & LaBarge, 1995). Performance of Alzheimer's patients

on revised reading measures including the AMNART and the NART-R has also been reported as significantly poorer than demographically matched controls, with performance deteriorating with increasing dementia severity (Taylor et al. 1996; Maddrey, Cullum, Weiner, & Filley, 1996).

Recently, Taylor et al. (1996) considered that estimation of premorbid IQ could be adjusted for dementia severity by quantifying the relationship between deterioration in AMNART performance and a measure of global cognitive functioning/impairment (i.e., MMSE - Mini Mental State Examination; Folstein, Folstein & McHugh, 1975). Taylor et al. (1996) documented declining AMNART performance in Alzheimer's patients over a three-year period and developed a correction formula for estimating WAIS-R VIQ that controlled for dementia severity by applying an MMSE-based correction factor to the Grober and Sliwinski (1991) AMNART/education equation. Taylor et al. (1996) reported that use of the correction factor produced VIQ estimates for Alzheimer's patients that did not differ significantly from demographically matched controls over three annual assessments. Although Taylor et al. (1996) managed to control for dementia severity in AMNART performance, the need to account for dementia severity in the estimation of premorbid IQ caused Taylor et al. (1996) to question the validity of the AMNART as an accurate estimate of premorbid IQ in Alzheimer's disease. Other studies have similarly

questioned the validity and stability of the NART and NART-R as estimates of premorbid IQ in Alzheimer's disease, and suggest that such measures may, at best, only offer reasonably accurate estimates of premorbid IQ in the early stages of dementia (Stebbins, Wilson et al. 1990; Patterson et al. 1994).

(b) Other Neurological Disorders

The utility of the NART at predicting premorbid IQ has been examined in various neurological populations including patients with Korsakoff's syndrome, Parkinson's disease, Huntington's disease, multi-infarct dementia, head trauma, and malignant brain tumors. Korsakoff's patients have generally performed significantly more poorly on the NART as compared to normal controls, have produced lower NART predicted IQs than demographically predicted IQs, have made more NART errors than predicted by demographic variables, and have demonstrated NART performance which has correlated with severity of memory impairment (Crawford et al. 1988; O'Carroll, Mofoot, Ebmeier, & Goodwin, 1992). Such results have suggested that use of the NART is invalid in patients with Korsakoff's syndrome (O'Carroll, 1995). Other patient groups that have also performed significantly more poorly on the NART relative to healthy controls have included Huntington's patients (Crawford et al. 1988; Blackmore, Crawford, & Simpson, 1994; as cited in O'Carroll, 1995), and

long term survivors of malignant glioma who had received whole brain prophylactic irradiation (Ebmeier et al. 1993; as cited in O'Carroll, 1995). Blackmore et al. (1994) suggested that demographic variables may be more useful than the NART at predicting premorbid IQ in Huntington's patients, and Ebmeier et al. (1993) concluded that the NART should be used with caution in survivors of malignant brain tumors, particularly involving the left temporal lobe.

Investigations with other neurological samples have suggested that the NART may demonstrate potential utility at estimating premorbid IQ in such samples. Boyd et al. (1991) found no significant difference in NART performance between Parkinson's patients and normal controls. Similarly, Crawford et al. (1988) indicated that patients with multi-infarct dementia and head-injured patients did not differ from normal controls on NART performance.

Bucks, Scott, Pearsall, and Ashworth (1996) examined the utility of the Short NART at estimating full-length NART error scores in patients from a memory disorders clinic. Bucks et al. (1996) found that the discrepancy between full-length NART error scores and Short NART predicted error scores was 4 points or more for 59% of the sample. Bucks et al. (1996) considered that such discrepancies were clinically and statistically unacceptable and recommended that the Short NART should not be used to estimate premorbid IQ in clinical practise without further modification.

Studies that have examined the predictive utility of other present ability measures at estimating premorbid IQ in clinical samples have produced mixed results. Johnstone, Hexum, and Ashkanazi (1995) considered that the WRAT-R Reading subtest was useful in predicting premorbid IQ in a sample of head-injured patients whereas, Johnstone and Wilhelm (1996) suggested that caution should be used when employing the WRAT-R/WRAT-3 Reading subtest for estimating premorbid IQ in a mixed neurological sample given the significant variability in reading score decline/improvement over time. Kupke (1996) examined the utility of the NART-R at predicting a measure of global intellectual capacity derived from the Armed Services Vocational Battery administered approximately 10 years prior in a sample of 100 individuals currently infected with HIV. Kupke (1996) reported a significant correlation of .72 between scores and considered that the NART-R demonstrated moderate accuracy in predicting premorbid cognitive status obtained approximately 10 years prior in an asymptomatic HIV sample. Beardsall and Huppert (1997) examined performance on the Spot-the-Word (STW) test in demented and normal groups and found that performance on the STW was preserved in patients with minimal dementia, but declined in patients with mild/moderate dementia as compared to normal controls. Scott et al. (1997) evaluated the utility of the Krull et al. (1995) equations (i.e., OPIE) at predicting WAIS-R FSIQs in

clinical samples including patients with dementia, traumatic brain injury, cerebral vascular accident, neoplasm, epilepsy, and a medical control chronic pain group. Scott et al. (1997) reported that the OPIE produced a less restricted range of scores than other currently available methods of estimating premorbid IQ, without systematic over-or-underestimation of predicted FSIQs, as the mean for each clinical group approached 100 FSIQ points.

Recently, a number of studies have attempted to compare the utility of various predictive measures at estimating premorbid IQ in clinical samples. Such comparisons have included the NART and Barona et al. (1984) equations in patients with Alzheimer's disease (Paolo, Tröster, Ryan, & Koller, 1997); Vanderploeg & Schinka's (1995) WAIS-R subtest/demographic equations BEST 3 approach and the Barona et al., (1984) equations in a heterogeneous neurological sample (Vanderploeg et al. 1996); Paolo et al. (1997) WAIS-R age-corrected subtest/demographic equations and the Barona et al. (1984) equations in a heterogeneous neurological sample; the NART and WRAT-R Reading subtest in a sample of patients with known or suspected neurological disorders (Glosser, Libon, & Friedman, 1993); and the NAART (i.e., NART-R) and the WRAT-R Reading subtest in a heterogeneous neurological sample (Johnstone, Callahan, Kapila, & Bouman, 1996). Unfortunately, the results of these studies have been difficult to interpret as there exists little consistency

between studies in terms of the means of comparing the clinical utility of premorbid IQ measures. Furthermore, many of the above studies failed to provide the necessary comparison of the performance of the premorbid IQ measure in both clinical and normative samples (as discussed earlier) to determine if the predictive measure is indeed resistant to the effects of the neurological disorder and provides an accurate measure of premorbid IQ (Crawford, 1989).

(c) Psychiatric Disorders

The utility of the NART as an estimate of premorbid IQ has also been evaluated in patients with schizophrenia, psychosis and depression. O'Carroll, Walker et al. (1992) found that there were no differences in performance for persons with acute schizophrenia or acute psychosis when compared to normal individuals after controlling for demographic differences. However, Young et al. (1991) found that patients with schizophrenia demonstrated significantly lower NART scores than controls. Crawford et al. (1992) also found that NART scores were significantly lower for patients with schizophrenia in a long stay ward, but found no significant difference in NART performance for patients with schizophrenia who were community residents as compared to normal controls. Crawford, Besson, Parker, Sutherland, and Keen (1987) found that there was no significant difference in NART performance for patients with depression as compared

to normal controls. Austin et al. (1992); also reported that patients with depression did not differ from normal controls on NART performance. Thus, as with neurological populations, the validity of the NART seems variable both within and between psychiatric conditions.

Summary

In summary, examination of the utility of the NART and other present ability measures as estimates of premorbid IQ in clinical samples has revealed that the NART and most other measures of present ability are not resistant to the effects of neurologic and psychiatric disorder as originally presumed. In Alzheimer's disease specifically, NART performance has been shown to become progressively poorer over time and is negatively affected by the severity of the dementing process and any language problems associated with the disorder. As such, it has been suggested that reading estimates of premorbid IQ may, at best, only offer reasonably accurate estimates of premorbid IQ in the early stages of Alzheimer's disease.

PART II:

The Utility of Predicting Premorbid IQ in Alzheimer's Disease

Given the above methodological limitations of premorbid

IQ measures in normative samples, and the clinical constraints among patients with Alzheimer's disease (AD), the utility of predicting premorbid IQ for patients with AD is questionable. Furthermore, predicting premorbid IQ may not be the most useful way to detect cognitive decline in Alzheimer's disease. Since most measures of present ability are likely to offer their most accurate estimate of premorbid functioning in the early stages of AD given the pervasiveness of the clinical course of the dementing disorder, estimation of premorbid functioning in cognitive domains sensitive to decline in the early stages of AD may provide a more useful index of cognitive impairment. Some studies have suggested that IQ may be more resistant to cognitive decline at the initial stages of a disease process than measures in other cognitive domains (Schlosser & Ivison, 1989). Thus, if IQ is initially more resistant to decline than other cognitive abilities in the early stages of AD, and if a minimum decline of 20 VIQ points (i.e., reading tests; Graves et al. in press) is required between a predicted and obtained VIQ score to reliably detect decline at a clinical level, then it is likely that the sensitivity of IQ at detecting cognitive impairment in the early stages of AD is limited. Predicting premorbid performance in cognitive domains sensitive to early Alzheimer's disease would allow for identification of decline in these domains at an early stage in the disease process. Such

quantification of decline in more sensitive cognitive domains may aid in diagnosis in the early stages of Alzheimer's disease when cognitive impairment may not yet be clinically obvious.

Cognitive Domains Sensitive to Decline in Early Alzheimer's Disease

1. *Verbal Memory*

Severe impairment in recent memory is a distinctive hallmark of Alzheimer's disease and is often regarded as the most important early manifestation of the disorder (Welsh, Butters, Hughes, Mohs, & Heyman, 1992). Early studies which examined the preclinical profile of Alzheimer's disease and the utility of various psychometric measures at differentiating mild senile dementia of the Alzheimer's type from normal aging, identified ***memory functioning*** as a highly sensitive indicator of Alzheimer's disease (Storandt, Botwinick, Danzinger, Berg, & Hughes, 1984; Storandt & Hill, 1989). Recently, interest has shifted to identifying which aspect or component of memory functioning demonstrates the greatest sensitivity to decline in early AD, and is most useful for detecting impairment in mild AD.

Two aspects of verbal memory, delayed recall and acquisition, have been found to accurately distinguish between mild Alzheimer's disease and normal aging. However,

controversy exists as to which verbal memory component demonstrates the greatest discriminative sensitivity in the classification of probable AD. Welsh, Butters, Hughes, Mohs & Heyman (1991) initially examined the performance of patients with early AD on a variety of memory tasks included in the CERAD (Consortium to Establish a Registry for Alzheimer's Disease) neuropsychological test battery to determine which memory measure best differentiated patients with early AD from cognitively normal elderly individuals. Welsh et al. (1991) reported that the ***delayed recall component*** of a three trial list learning task was the best overall discriminatory measure between groups. Other tests of memory including immediate recall, intrusion errors, and recognition memory demonstrated poor overall discriminability for their sample. Robinson-Whelen and Storandt (1992) however challenged the discriminative sensitivity of delayed recall in early AD as they found that delayed recall scores were not reliable at distinguishing between their sample of demented patients and normal controls on a measure of paragraph recall, once the scores had been controlled for initial learning. Similarly, Becker, Boller, Saxton, and McGonigle-Gibson (1987) reported that the rate of forgetting of verbal material did not differ significantly between AD patients and normal controls in their study. Although the memory measures used in the above studies were not directly comparable given that Welsh et al.

(1991) utilized a multiple trial free recall procedure, and Robinson-Whelen and Storandt (1992) administered a measure of paragraph recall with one learning trial, the role of delayed recall in the detection of early dementia was called into question.

Consequently, Petersen, Smith, Ivnik, Kokmen, and Tangalos (1994) investigated several aspects of memory function in AD and normal aging to determine which indices of performance were most sensitive at detecting early impairment. Petersen et al. (1994) evaluated 106 pairs of patients with probable AD and matched controls using a logistic regression model that included measures of memory function, verbal and nonverbal intelligence, attention, and language. Results indicated that a measure of acquisition, ***immediate free recall***, (especially with semantic cuing), was the most sensitive at separating the two groups. Several other studies have also supported the sensitivity of immediate recall at detecting decline in early AD, given the pronounced changes in the acquisition component of memory at the onset of dementia (Jacobs et al. 1995; Mitrushina et al. 1994; Almkvist & Backman, 1993; Tuokko, Vernon-Wilkinson, Weir, & Beattie, 1991). Such studies have generally examined immediate recall performance on multiple trial list learning measures such as the Selective Reminding Test (SRT; Buschke, 1973) and the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964). Petersen et al. (1994) addressed the disparity

between their results and those of the Welsh et al. (1991) study and suggested that the initial identification of delayed recall as a more sensitive indicator to early decline in AD than immediate recall may have been based on differences in the memory measures utilized in their respective studies. Petersen et al. (1994) indicated that the memory task used in the Welsh et al. (1991) study involved only three learning trials and suggested that this may not have been sufficient to detect true differences in acquisition performance between AD patients and controls. They indicated that measures which extend the learning task to five or six trials have demonstrated differences in acquisition performance in that normal controls continue to acquire additional information as the task exceeds the immediate memory span, whereas impaired individuals fail to exceed their immediate memory span with acquisition reaching a plateau after a few learning trials.

Recently, research in the investigation of cognitive markers sensitive to early decline in Alzheimer's disease has shifted from cross-sectional comparisons of persons with mild AD and normal controls to longitudinal prospective studies in which healthy elderly individuals, some with the risk of familial Alzheimer's disease, have undergone serial assessments over several years to identify the earliest neuropsychological features of the disease. Most studies have identified verbal memory decline as the earliest

neuropsychological sign that may predate symptoms of Alzheimer's disease by several years. Fox, Warrington, Seiffer, Agnew, and Rossor (1998) followed 63 asymptomatic individuals at risk of autosomal dominant familial Alzheimer's disease over a six-year period, and found that for individuals who subsequently went on to develop Alzheimer's disease, all scored lower on verbal memory testing (i.e., recognition for words) when first assessed (i.e., 1 to 5 years before they became symptomatic) as compared with the group that remained well. Fox et al. (1998) emphasized that most (i.e., 8) of the 10 individuals who went on to develop Alzheimer's disease performed within the normal range of the standardized population on verbal memory tests at the initial assessment when considered on an individual basis. It was not until these individuals were considered as a group that lower memory scores were statistically significant. Bondi et al. (1994) examined the performance of 56 nondemented elderly individuals (approximately half with a positive family history of progressive dementia) on three consecutive annual administrations of the California Verbal Learning Test, a multiple trial list learning measure (CVLT; Delis, Kramer, Kaplan, & Ober, 1987). Bondi et al. (1994) reported that normal controls with a positive family history of progressive dementia performed significantly worse than individuals with a negative family history for dementia on

11 of 16 CVLT measures at the first evaluation including learning, retention, and error measures. Bondi et al. (1994) indicated that five individuals subsequently went on to develop dementia of the Alzheimer's type, four of which had a positive family history of progressive dementia. Comparison of the performance of these preclinical individuals to normal controls at the first annual evaluation on the CVLT and additional memory measures revealed that the five preclinical individuals performed significantly worse than controls on immediate and delayed recall indices of the CVLT, SRT (Buschke, 1973), and the Logical Memory subtest (i.e., story/paragraph recall) of the Wechsler Memory Scale-Revised (WMS-R; Wechsler, 1987).

Studies which have serially assessed healthy elderly individuals without a documented history of familial Alzheimer's disease have also found evidence of verbal memory impairment predating clinical evidence of AD. Howieson et al. (1997) followed healthy individuals 80 years and older over a three-year period and reported that the baseline measure that contributed most to the classification of individuals who subsequently developed AD from those who remained well was the delayed recall task of the Logical Memory subtest of the WMS-R. Longitudinal followup over 13 years of elderly individuals in the Framingham cohort also revealed that verbal recall scores (i.e., percent retained over a delay on the Logical Memory subtest of the Wechsler

Memory Scale - WMS; Wechsler, 1945) were lower at initial examination in participants who later developed dementia as compared with those who remained well (Linn et al. 1995). Masur, Sliwinski, Lipton, Blau, and Crystal (1994) assessed neuropsychological performance in 317 initially nondemented elderly persons between 75 and 85 years of age and followed them for at least four years as part of the Bronx Aging Study. Masur et al. (1994) found that four measures of cognitive function from baseline assessment (i.e., delayed recall of the SRT; recall from the Fuld Object Memory Evaluation (Fuld, Masur, Blau, Crystal, & Aronson, 1990); Digit Symbol subtest from the WAIS (Wechsler, 1955); and a verbal fluency score) could classify persons who subsequently developed dementia over four years with 85% sensitivity and 95% specificity. Overall, the results of the above studies suggested that most aspects of verbal memory (i.e., immediate and delayed recall, recognition) are sensitive to decline in early Alzheimer's disease and this decline may predate clinical symptoms by several years.

2. ***Language***

Welsh et al. (1992) were interested in determining whether any nonmemory functions including naming, fluency, and praxis, added to the detection of early Alzheimer's disease, or were sensitive to the later progression of the illness. Welsh et al. (1992) examined the performance of 147

patients with Alzheimer's disease (AD) stratified according to disease severity and demographically matched with normal controls, and reported that ***confrontational naming*** (i.e., 60-Item Boston Naming Test (BNT); Kaplan, Goodglass, & Weintraub, 1983) was the only nonmemory measure that assisted in discriminating between patients with mild dementia and normal controls. Decline in the domain of language functioning is consistent with the clinical course of AD, and recent studies have supported the finding that naming problems are characteristic of preclinical AD and are independently associated with the later diagnosis of AD (Jacobs et al. 1995; Mitrushina, Uchiyama, & Satz, 1995). Howieson et al. (1997) reported that the BNT was the only nonmemory measure that distinguished individuals who subsequently became demented from those who maintained their cognitive status in a three year longitudinal study designed to identify preclinical markers of Alzheimer's disease in healthy older persons. Examination of cognitive profiles in normal aging has also identified three clusters of impaired ability including 1) verbal learning and memory, 2) naming and 3) visuoconstructive/visuospatial ability which Mitrushina et al. (1995) considered to be representative of the preclinical stages of the dementing process, suggesting that memory problems may not be the only manifestation of the disease process in early AD.

Examination of language performance (i.e., naming) may

also be particularly useful for identifying cognitive impairment in female patients with mild Alzheimer's disease. Studies examining gender differences in language performance of AD patients have revealed that women perform significantly more poorly than men on language measures such as the BNT (Ripich, Petrill, Whitehouse, & Ziol, 1995; Henderson & Buckwalter, 1994). Longitudinal examination of language performance in women and men with AD has also revealed that language abilities of women are more severely impaired at all stages of the dementing process than the language abilities of men, including the early stages, whereas, the language performance of normal elderly individuals does not differ significantly by gender over time (Ripich et al. 1995). Therefore, examination of both memory and language performance, especially in women with possible AD, may significantly improve the detection of decline and the diagnosis of AD in the earliest stages of the disease process.

Prediction of Premorbid Performance in Verbal Memory and Language Domains

Prediction of premorbid performance of a verbal memory measure (i.e., delayed recall or immediate recall of a multiple trial list learning task) or a language test (i.e., confrontation naming test) may be potentially accomplished via three different approaches. The first approach would be

to examine the utility of existing measures of premorbid ability (i.e., IQ) discussed previously, at estimating verbal memory performance and language performance in a normal elderly sample. A second approach would be to determine which present abilities remain relatively preserved in the early stages of Alzheimer's disease that would be useful in predicting verbal memory performance and language performance. These measures would also have to demonstrate high correlations with memory and language functioning in normal elderly controls prior to cross-validation with a clinical sample of patients with early Alzheimer's disease. A third approach would be to determine if demographic variables (i.e., age, gender, education) would be useful in predicting verbal memory and language performance. An advantage of demographic predictors is that they remain resistant to the effects of neurologic or psychiatric disease.

1. ***Predicting Performance in Cognitive Domains other than General Intellectual Ability***

Predictors of premorbid IQ, particularly the NART, have also been used to predict performance in cognitive domains other than general intellectual ability. Schlosser and Ivison (1989) examined the utility of the NART, the Schonnell Graded Word Reading Test (SGWRT) and demographic variables at predicting general memory performance on the

Wechsler Memory Scale (WMS; total raw score) for a sample of 65 healthy elderly participants aged 65-89 years. They reported that the NART alone accounted for 45% of the variance ($r = 0.67$) in WMS total raw score, and further improved the utility of the NART by combining it with SGWRT errors and age in a regression equation accounting for 55% of the variance ($r = 0.74$) in WMS total raw score. The discrepancy between predicted and obtained WMS scores that reliably accounted for 95% of the normal sample was 13 points based on the distribution of discrepancy scores for the sample, with a maximum discrepancy range of 21 points between predicted and obtained WMS scores for the normal sample (Schlosser & Ivison, 1989). Cross-validation of the WMS regression equation with 16 Alzheimer's patients revealed that all patients achieved discrepancies between predicted and obtained WMS total raw score of greater than 13 points, with only two patients (12.5% of the clinical sample) achieving discrepancy scores that fell at or within the 21 point maximum of the normal error range. Schlosser and Ivison (1989) considered that predicting memory performance was more sensitive at detecting decline in a dementing sample than predicting IQ performance, given that only 12.5% of their clinical sample achieved discrepancy scores that fell within the normal error range, whereas, Nelson and McKenna (1975) reported that 53% of their dementing patients achieved discrepancies between predicted

and obtained IQ performance that fell within the normal error range.

The NART has also been used to estimate performance on a measure of verbal fluency (VF; FAS Test), in the domain of language ability. Crawford, Moore, and Cameron (1992) administered the NART and VF test, which requires participants to orally produce words with designated letters (F,A,S) in a set time period (60 sec/letter), to a sample of 142 individuals free of neurological or psychiatric disorder. Crawford et al. (1992) reported a highly significant correlation between NART and VF, and developed a regression equation based on NART errors that predicted 45% of the variance ($r = 0.67$) in the VF score. They indicated that a discrepancy between predicted and obtained VF scores of more than 12 points was exhibited by less than 10% of the control sample. Cross-validation of the VF equation with a heterogeneous neurological sample revealed a highly significant difference in favour of NART estimated VF performance over current VF performance.

Although no other studies have developed regression equations to predict functioning in other cognitive domains based on measures of premorbid IQ, various studies have examined the correlations between measures of premorbid IQ and other cognitive abilities, or have used performance on the premorbid IQ measure to index decline in other cognitive domains. Storandt, Stone, and LaBarge (1995) recently

examined AMNART performance in patient groups with very mild and mild dementia of the Alzheimer's type matched with healthy elderly controls. All participants underwent a two hour psychometric battery that examined a broad range of psychological functions including memory, language skills, intellectual ability, psychomotor performance, and visuospatial abilities (Storandt & Hill, 1989). Correlations between AMNART scores and measures of paragraph learning (Logical Memory; WMS), confrontational naming (BNT), and verbal fluency (Letters S,P) were .37, .40, and .59, respectively, for healthy elderly individuals.

Johnstone, Hexum, and Ashkanazi (1995) examined the extent of decline in various cognitive domains including intelligence, memory, attention, speed of processing and cognitive flexibility for a sample of patients that had sustained a traumatic brain injury. Performance on the Reading subtest of the WRAT-R served as an estimate of premorbid ability <u>across</u> cognitive domains, and extent of decline in each cognitive domain was calculated by subtracting WRAT-R (z) scores from cognitive test (z) scores to determine a (z) difference score for each cognitive ability. The results suggested that intelligence declined the least after traumatic brain injury, followed by attention, memory, speed of processing and cognitive flexibility. Unfortunately, a control group was not utilized in this study and Johnstone et al. (1995) failed to provide

the correlations between the WRAT-R reading subtest and various measures of cognitive ability which would have not only demonstrated the potential utility of the WRAT-R Reading subtest at predicting performance in other cognitive domains, but would have determined whether the WRAT-R Reading subtest could in fact be used as an estimate of premorbid ability across cognitive domains in a traumatic brain injury sample.

Although there are a limited number of studies which have examined the predictive utility of premorbid IQ measures at estimating performance in cognitive domains other than general intellectual ability, and only two studies which have developed regression equations utilizing the NART to predict general memory and verbal fluency performance, respectively, it is suggested that measures of premorbid IQ may demonstrate potential utility in predicting performance in verbal memory and language domains in normal elderly participants. To date, no studies have developed regression equations to predict verbal memory or language performance based on measures of premorbid IQ. Furthermore, no measures of premorbid IQ other than the NART have been used to predict performance in a cognitive domain other than general intellectual ability.

2. *Verbal Memory*

(a) Predicting Verbal Memory Performance Utilizing Measures of Present Ability

Predicting verbal memory performance utilizing measures of present ability requires identification of measures that remain relatively preserved in the early stages of Alzheimer's disease (AD) and demonstrate potential utility in predicting verbal memory performance on a multiple trial list learning task. Identification of measures preserved in the early stages of AD is not easy given that few studies have compared the cognitive profiles of patients with early AD to normal elderly controls, with most studies focusing on impaired rather than preserved performance. An added limitation is that the criteria for classifying patients at various stages of AD (mild vs. moderate etc.) varies across studies, reducing the comparative utility of such investigations. Even in light of the above constraints, there has been some suggestion that verbal memory span procedures and serial position recall effects (i.e., recency effects) may demonstrate relative preservation in early AD (Parks, Zec, & Wilson, 1993). The following discussion will consider the utility of such measures as potential predictors of verbal memory performance.

(i) Verbal Memory Span

Verbal memory span procedures generally require the immediate recall or repetition of word, letter, or digit sequences. The ***forward digit span*** subtest of the WAIS/WAIS-R is a verbal memory span procedure commonly employed in neuropsychological practise that requires the oral repetition of digit sequences of increasing length. Forward verbal span has been traditionally regarded as a reliable index of short-term memory capacity (Miller, 1956) and has recently been viewed as a measure of the capacity of a verbal memory system (referred to as the phonological loop) within Baddeley's (1992) model of working memory that maintains information for complex verbal manipulations (Cherry, Buckwalter, & Henderson, 1996). Baddeley (1992) described working memory as a brain system that governs the simultaneous storage and processing of information necessary in complex cognitive tasks, and is divided into three subcomponents including 1) the *central executive* which serves as an attentional-controlling system, 2) the *visuospatial sketch pad* which manipulates visual images and 3) the *phonological loop* which stores and rehearses speech-based information. Preservation of forward verbal span in early Alzheimer's disease has been attributed to a relatively intact phonological loop (Carlesimo, Fadda, Lorusso, & Caltagirone, 1994). Several studies have documented a normal forward digit span for early AD

patients, in which WAIS/WAIS-R forward digit span scores did not differ significantly from controls (Small, Herlitz, Fratiglioni, Almkvist, & Bäckman, 1997; Lafleche & Albert, 1995; Carlesimo et al. 1994; Flicker, Ferris, & Reisberg, 1993; Lines et al. 1991; Orsini, Trojano, Chiacchio, & Grossi, 1988; Storandt, Botwinick, & Danzinger, 1986). Storandt et al. (1986) considered that forward digit span may also be useful in determining stage or severity of AD, as longitudinal investigation of cognitive functioning in patients with early AD demonstrated that forward verbal span did not deteriorate until the more advanced stages of AD. Morris (1984, 1986, 1987) identified additional components of the phonological loop other than span that have been found to remain relatively intact in early AD including normal phonological similarity and word length effects, and normal contribution of articulatory rehearsal to memory span. Studies which have demonstrated reduced digit span forward for patients with AD as compared to normal controls have generally tested patients in the more advanced stages of AD, or have combined patient groups with varying severity (Miller, 1971; Kaszniak, Garron, & Fox, 1979; Kopelman, 1985).

An alternative verbal span task that may demonstrate potential utility at predicting verbal memory performance on a word list learning task is a measure of ***sentence repetition***. Repetition itself remains relatively intact

until the more advanced stages of AD (Cummings, Benson, Hill, & Read, 1985), and Nebes et al. (1984) indicated that immediate recall of word strings may be improved for both normal and dementing individuals by placing the words within the format of a normal English sentence. If context does in fact improve word list recall, sentence repetition may offer a more accurate measure of forward verbal span than WAIS/WAIS-R forward digit span. A measure of sentence repetition is included in the Neurosensory Centre Comprehensive Examination for Aphasia (NCCEA; Spreen & Benton, 1969) that requires the oral repetition of sentences of increasing length and complexity. Murdoch, Chevery, Wilks, & Boyle (1981) indicated that the group mean score for patients with AD fell above the 40th percentile and outside the "aphasia range" on the sentence repetition component of the NCCEA.

Determining the utility of forward verbal span at predicting verbal memory performance on a multiple trial list learning task is difficult to ascertain given the relative lack of data examining the relationship between such measures in normal elderly samples. Only one study has examined the utility of forward digit span (WAIS) and forward word span at predicting free recall performance on a multiple trial list learning measure in a normal elderly sample. Parkinson, Lindholm, and Inman (1982) presented a 12 word list over 12 test trials to 26 normal elderly

individuals and found that forward digit span and forward word span accounted for 48% and 30% of the variance in free recall performance, respectively, supporting the utility of forward verbal span procedures at predicting verbal memory performance in a normal elderly sample.

(ii) Serial Position Recall Effects

Primacy and recency effects refer to the relatively better recall amongst normal controls for words at the beginning (primacy) and end (recency) of a list as compared to the middle of the list (Parks et al. 1993). The primacy effect is considered to be associated with the operation of long-term memory and reflects the greater rehearsal that information at the beginning of a list receives, whereas the recency effect is considered attributable to information at the end of the list that is maintained in short-term memory (Capitani, Della Sala, Logie, & Spinnler, 1992). Various scoring systems have been proposed for calculating primacy and recency effects. When a list is approximately 15 words in length, recall of the first 3-5 items is generally regarded as the primacy effect, whereas, recall of the last 3-5 items is generally regarded as the recency effect (Glanzer & Cunitz, 1966; Tulving & Patterson, 1968). More complicated scoring procedures have been developed for calculating primacy and recency effects that take into account interference that may arise between presented items

and items recalled (Waugh & Norman, 1965; Tulving & Colotla, 1970). Some multiple trial list learning measures including the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987) calculate serial position recall effects as part of their scoring procedure. The CVLT is a 16 item word list with items from one of four semantic categories including fruits, herbs/spices, tools and clothing that is presented over five learning trials to obtain a measure of total immediate recall (Trials 1-5). On the CVLT, the primacy and recency regions are defined as the first and last four words on the list, respectively, with the remaining eight words considered to be the middle region. Three indices are calculated to reflect the percentage of total words correct on trials 1-5 that came from each of the three list regions (primacy, middle, recency). The degree to which an examinee's recall corresponds to the serial recall pattern displayed by the normative sample can be determined by examining the pattern of standard scores on the three indices.

The ***recency component*** of free recall has been considered by many to be the test least sensitive to short-term memory impairment (Morris, 1994; Glanzer & Cunitz, 1966; Waugh & Norman, 1965). Studies that have examined the recency component of free recall in patients with early AD have either found no significant difference in recency effect between patients and controls (Incalzi, Capparella,

Gemma, Marra, & Carbonin, 1995; Spinnler, Della Sala, Bandera, & Baddeley, 1988; Wilson, Bacon, Fox, & Kaszniak, 1983; Moscovitch, 1982), or have reported only a mild reduction in the recency effect of patients with early AD as compared to controls (Miller, 1973; Pepin & Eslinger, 1989; Capitani et al. 1992). Research with the CVLT specifically, revealed no difference in the recency effect (i.e., mean recall in recency) between patients with mild to moderate Alzheimer's disease and normal controls (Simon, Leach, Winocur, & Moscovitch, 1994). Pepin and Eslinger (1989) also examined the pattern of free recall performance in patients with mild, moderate, and severe Alzheimer's disease and indicated that the patients with mild AD demonstrated a pattern of free recall performance that was qualitatively similar to that observed in normals with evidence of both primacy and recency effects. The recency effect has also been shown to be resistant to decline in other clinical conditions including head injury (Brooks, 1975), amnesia, (Baddeley & Warrington, 1970) and Parkinson's disease (Della Sala, Pasetti, & Sempio, 1987). Baddeley (1986) considered that the resistance of the recency effect may be largely attributable to its relative automaticity as an ordinal retrieval strategy.

Establishing the utility of the recency component at predicting verbal memory performance on a multiple trial list learning measure for a normal elderly sample remains

difficult given the relative lack of data examining the relationship between such measures. Although Parkinson et al. (1982) were primarily interested in determining the utility of measures of forward verbal span at predicting free recall performance, they also examined the utility of each of the span measures at estimating primacy and recency effects. Parkinson et al. (1982) reported that forward digit span and word span produced significant correlations with primary memory in their normal elderly sample and accounted for 34% and 16% of the variance in recency effects, respectively. As forward memory span procedures have demonstrated utility in predicting free recall performance, and memory span procedures (i.e., digit span) demonstrate some degree of shared variance with recency effects, it is possible that recency effects alone may demonstrate utility in predicting free recall performance. However, further investigation with a normal elderly sample will be necessary to determine/confirm the utility of recency effects at predicting verbal memory performance on a multiple trial list learning task.

(b) Predicting Verbal Memory Performance Utilizing Demographic Variables

Several studies have examined the effects of demographic variables including age, gender, and education on verbal memory performance. Age and gender have routinely

been reported to correlate significantly with acquisition and delayed recall components of multiple trial list learning measures, with performance declining with advancing age, and with females outperforming males (Paolo, Tröster, & Ryan, 1997b; Welsh et al. 1994; Wiederholt et al. 1993; Trahan & Quintana, 1990; Geffen, Moar, O'Hanlon, Clark, & Geffen, 1990; Kramer, Delis, & Daniel, 1988). The influence of educational level however, has been less consistent across studies. When education effects have been reported, they have generally been restricted to the acquisition component of verbal learning measures, with performance improving with advancing education (Paolo et al. 1997b; Geffen et al. 1990).

Given the relationship between demographic variables and verbal memory performance, two studies have examined the utility of demographic predictors at estimating verbal memory scores. Unverzagt, Hui, Farlow, Hall, and Hendrie (1998) developed a demographic regression equation utilizing age, education and gender that accounted for approximately 48% of the variance in the total immediate free recall score of the CERAD Word List Learning test in an elderly African-American sample. Williams (1997) however found that the combination of race, gender, education and occupation only accounted for 13% of the variance in the Verbal Memory Index of the Memory Assessment Scales (MAS; Williams, 1992). As the results of the above studies are somewhat variable,

further research is needed to determine the utility of demographic variables at estimating verbal memory performance in healthy elderly samples.

3. *Language*

(a) Predicting Language Performance Utilizing Measures of Present Ability

Predicting language performance utilizing measures of present ability requires identification of measures that remain relatively preserved in the early stages of Alzheimer's disease and demonstrate potential utility in predicting language performance on a measure sensitive to decline in early AD such as the Boston Naming Test (BNT). Difficulty with word finding, especially naming, appears early in the course of AD (Parks et al. 1993), and Welsh et al. (1992) found that confrontational naming, as measured by BNT, was the only nonmemory measure that was sensitive enough to discriminate between patients with early AD and normal controls. Naming performance on the BNT has also been shown to be a good predictor of future decline in Alzheimer's disease, and is sensitive to gender as women with AD demonstrate greater decline in naming performance on the BNT than men with AD (Boller et al. 1991). Other components of language including speech fluency, auditory comprehension, articulation, and repetition, have been found

to remain relatively intact in the early stages of Alzheimer's disease (Cummings et al. 1985). The Vocabulary subtest of the Wechsler scales has also demonstrated relative preservation in the early stages of Alzheimer's disease (Sullivan, Sagar, Gabrieli et al. 1989), as has simple object identification (Flicker, Ferris, Crook, & Bartus, 1987). The following discussion will consider the utility of the WAIS/WAIS-R Vocabulary subtest and a simple object identification task as potential predictors of naming performance on the BNT. The Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981) is a receptive vocabulary measure that will also be considered in the prediction of BNT performance.

(i) Wechsler Vocabulary

The Vocabulary subtest of the Wechsler scales is the most commonly used test of vocabulary in clinical practise, and has long been recognized as an excellent guide to general mental ability (Lezak, 1995). Wechsler (1955) originally classified Vocabulary as a "Hold" subtest that demonstrated resistance to age and organic impairment, and could potentially provide an estimate of a patient's premorbid level of functioning. Although it is clear now that Vocabulary performance is compromised in various clinical conditions (Crawford et al. 1988; Crawford, Besson, Parker, Sutherland & Keen, 1987; Klesges et al. 1981), and

in the more advanced stages of Alzheimer's disease (Flicker et al. 1993; Sharpe & O'Carroll, 1991; O'Carroll et al. 1987; Hart et al. 1986; Houlihan, Abrahams, LaRue, & Jarvik, 1985), fewer studies have examined Vocabulary performance in patients with early AD. Some studies reviewed have reported no difference in Wechsler Vocabulary subtest performance for patients with early AD and normal controls (Mittenberg, Thompson, Schwartz, Ryan, & Levitt, 1991; Sullivan et al. 1989), whereas, other studies have reported minimal decline in Vocabulary performance for patients with early AD as compared to normal controls, with Vocabulary performance still falling within the average range (Flicker et al. 1987; Martin & Fedio, 1983). Such studies would suggest that Wechsler Vocabulary subtest performance may be relatively preserved, at least in the early stages of AD.

Only two studies have examined the predictive utility of the Wechsler Vocabulary subtest at estimating performance on the BNT. Thompson and Heaton (1989) originally reported a highly significant correlation between the WAIS-R Vocabulary subtest and BNT scores of .79 in a clinical sample characterized by heterogeneous neurological conditions. More recently, Tombaugh and Hubley (1997) examined the relationship between the WAIS-R Vocabulary subtest and BNT scores in a sample of cognitively intact Canadian adults aged 25 to 88 years and found that BNT scores correlated .53 with WAIS-R Vocabulary raw scores. Although the results of

the above studies are favourable, further investigation is necessary to determine if the predictive utility of the Wechsler Vocabulary subtest is maintained in normal elderly samples.

(ii) Peabody Picture Vocabulary Test-Revised

An alternate vocabulary test that may be useful in the prediction of BNT performance in early Alzheimer's disease is the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981). The PPVT-R is a receptive vocabulary measure that requires a participant to choose one of four items displayed on a card that corresponds with the word spoken by the examiner (Spreen & Strauss, 1998). As a verbal response is not required to complete the test (i.e., the individual may point to the item), the PPVT-R has been considered to be useful in assessing vocabulary skills in patients with expressive language deficits (Maxwell & Wise, 1984). Unfortunately, there are no studies to date that have examined the utility of the PPVT-R as a vocabulary test in early Alzheimer's disease, or whether the PPVT-R is resistant to decline in early Alzheimer's disease.

Only limited data exists regarding the relationship between the PPVT-R and other cognitive measures in adult samples. Various studies have reported a significant relationship between the PPVT-R and measures of verbal intelligence (i.e., WAIS/WAIS-R VIQ) and vocabulary (i.e.,

Wechsler Vocabulary subtest) in both normal and clinical adult samples (Maxwell & Wise, 1984; Carvajal, Shaffer, & Weaver, 1989; Stevenson, 1986; Prout & Schwartz, 1984; Mangiaracina & Simon, 1986). To date, no studies have examined the predictive utility of the PPVT-R at estimating performance on the BNT in a normal adult or elderly sample. One study however, examined the relationship between the PPVT-R and BNT in normal children aged 6 to 12 years and reported a significant correlation of .54 between PPVT-R and BNT scores (Halperin, Healey, Zeitchik, Ludman, & Weinstein, 1989). Although data with children cannot be generalized to an adult or elderly population, the above result is encouraging regarding the utility of the PPVT-R at estimating BNT scores. Additionally, the PPVT-R has been shown to demonstrate a significant relationship with the Wechsler Vocabulary subtest, which itself has been shown to correlate with BNT scores in adult samples (Thompson & Heaton, 1989; Tombaugh & Hubley, 1997). However, further investigation is warranted to determine the predictive utility of the PPVT-R at estimating BNT scores in normal elderly samples.

(iii) Object Identification

Flicker, Ferris, Crook, and Bartus (1987) described an object identification task in which a representation of a 25 room house was presented on a video monitor screen, with

each room filled with images of household items. In the object identification task the name of one of the objects appeared above the house on each of the 12 trials, and the individual was instructed to point to the object. Flicker et al. (1987) initially administered this task to a sample of young normals, elderly normals, and patients with mild and moderate Alzheimer's disease and indicated that there was no significant difference in the performance of this task between normal controls and patients with early AD. Only the patients in the more advanced stages of AD demonstrated impairment on this task. Flicker et al. (1987) interpreted the results to suggest that patients with early AD are unimpaired in the visual recognition of a named object, given that the association between word concepts and their visual representation remains relatively intact, even though the association between word concepts and their verbal label is much weaker. These results are also consistent with the conclusion that the naming deficit in early dementia is not primarily dependent upon a perceptual impairment (Bayles & Tomoeda, 1983).

Flicker, Ferris, and Reisberg (1993) proceeded to examine cognitive functioning in normal aging and Alzheimer's disease over a two year period with the administration of a comprehensive neuropsychological test battery including the object identification task. The study population included 33 young neurologically normal persons,

50 elderly neurologically normal persons, and 86 persons with probable AD divided into two groups (early vs. advanced) according to level of impairment. Results revealed that at initial assessment there was no significant difference in performance on the object identification task for young normals, elderly normals, and patients with early AD, and at two year follow-up, the performance of patients with early AD did not differ significantly from baseline performance on the object identification task. Patients with advanced AD demonstrated impairment on the object identification task at initial testing, with performance declining progressively on the task over the two year period. The above results were consistent with the results of the original study by Flicker et al. (1987) suggesting that performance on the object identification task remains relatively intact in early AD.

Similar tasks have been developed which require the visual recognition of a named object. The word-picture matching subtest of the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983) requires the selection of a picture of an object, action, colour, or number that corresponds with a written word over 10 trials. However, further investigation is required to determine the predictive utility of object identification tasks such as the word-picture matching subtest of the BDAE at estimating naming performance on the BNT, as there do not appear to be

any studies which have examined the relationship between such measures in normal elderly samples.

(b) Predicting Language Performance Utilizing Demographic Variables

A number of studies have examined the effects of demographic variables on naming performance on the BNT. Significant effects have been reported for both education and age as BNT scores have been shown to increase with a higher level of education, and decrease with advancing age, particularly after age 70 (Tombaugh & Hubley, 1997; Welch, Doineau, Johnson, & King, 1996; Ross, Lichtenberg, & Christensen, 1995; Thompson & Heaton, 1989; Van Gorp, Satz, Kiersch, & Henry, 1986; LaBarge, Edwards, & Knesevich, 1986). Most studies however, have failed to demonstrate a significant relationship between gender and BNT performance (Tombaugh & Hubley, 1997; Ross et al. 1995; LaBarge et al. 1986). Although several studies have identified a significant relationship between demographic variables (i.e., age, education) and BNT scores, it should be noted that to date, there are no studies that have examined the utility of demographic predictors at estimating BNT scores in normal elderly samples.

Purpose

The purposes of the current study were: 1) to compare

the concurrent validity of several recently proposed measures of premorbid IQ (i.e., IQ equations and independent IQ predictors) at predicting WAIS-R IQ in a normal elderly Canadian sample, 2) to determine the concurrent validity of these same IQ measures at estimating performance in other cognitive domains thought to be particularly sensitive to decline in early Alzheimer's disease (i.e., verbal memory and language), in the same normal elderly sample, 3) to determine the concurrent validity of several additional present ability measures and demographic variables at estimating verbal memory and language performance in the same normal elderly sample, and 4) to evaluate the sensitivity of these predictive measures for detecting cognitive decline in a clinical sample, given the error of prediction of such measures in a normal elderly sample.

1. ***IQ Predicted by IQ Measures***

Examination of the predictive validity of several proposed measures of premorbid IQ at estimating WAIS-R IQ scores in a normal elderly Canadian sample provided an opportunity to: 1) obtain normative data for elderly Canadians on these measures, 2) cross-validate some of the more recently proposed measures of premorbid IQ with a healthy elderly sample, 3) compare the utility of various measures of premorbid IQ at estimating WAIS-R IQ in the same normal elderly sample, and 4) identify and/or develop the

most useful predictive measure/equation for estimating WAIS-R IQ for elderly Canadians, and detecting cognitive decline in clinical samples. Within the domain of general intellectual ability, WAIS-R Verbal IQ (VIQ) was selected as the criterion measure for intelligence prediction, since previous studies have shown that measures of present ability are most accurate at estimating VIQ. Several existing VIQ equations were selected for comparative investigation in the estimation of WAIS-R VIQ including: the Krull et al. (1995) equation (WAIS-R Vocabulary & demographics); two equations by Carswell et al. (1997), (i.e., NART & WAIS-R Vocabulary; NART alone); the Blair & Spreen (1989) equation (NART-R); the Ryan & Paolo (1992) equation (NART); the Nelson & Willison (1991) equation (NART); and the Barona et al. (1984) equation (Demographics). The current study also examined the utility of independent VIQ predictors (i.e., NART errors, NART-R errors, WAIS-R Vocabulary (age-scaled score), Spot-the-Word (STW) subtest score, and WRAT-3 Reading subtest raw score) and demographic variables (age, sex, education) at predicting WAIS-R VIQs in a normal elderly Canadian sample.

2. *Memory and Language Predicted by VIQ Predictors*

The second part of the present study involved the examination of the concurrent validity of the above VIQ measures (i.e., VIQ equations and independent VIQ

predictors) at estimating memory and language performance in the same normal elderly Canadian sample.

In the domain of verbal memory, the long delay free recall trial of the California Verbal Learning Test (CVLT) was selected as the criterion measure for memory prediction. Although there has been some debate in the literature with regard to the aspect of verbal memory that is most sensitive to decline in early Alzheimer's disease (i.e., acquisition versus delayed recall), the current study selected the delayed recall component of the CVLT as the criterion variable for psychometric reasons primarily, to guard against the likelihood of contamination of the criterion variable with a present ability memory predictor obtained from the same memory measure (i.e., CVLT Recency score). The CVLT was selected as the verbal memory measure as it is a multiple trial list learning measure that has been reported to correlate highly with other memory measures (Delis et al. 1987), and offers advantages over other multiple trial list learning measures including normative data for males and females up to 80 years of age, and calculation of the recency effect within the scoring profile.

In the domain of language functioning, the Boston Naming Test was selected as the criterion measure for language prediction as confrontational naming has been identified as the most sensitive non-memory measure for detecting decline in early AD.

3. ***Memory and Language Predicted by Measures of Present Ability and Demographics***

The third part of the study involved the examination of the concurrent validity of several measures of present ability and demographic variables (i.e., age, sex, education) at predicting memory and language performance in the same normal elderly Canadian sample.

Present ability measures selected for prediction of delayed recall on the CVLT included Digits Forward (WAIS-R), Sentence Repetition (NCCEA), and the Recency Effect of the CVLT, due to their relative preservation in the early stages of AD.

Present ability measures selected for prediction of performance on the Boston Naming Test, due to their relative preservation in the early stages of AD, included the Vocabulary subtest of the WAIS-R, the Peabody Picture Vocabulary Test-Revised, and the word-picture matching task on the Boston Diagnostic Aphasia Examination. As the WAIS-R Vocabulary subtest was investigated as a present ability language measure in the prediction of BNT performance, it was not evaluated as an VIQ predictor in the estimation of BNT scores.

4. ***Clinical Sensitivity of Predictive Measures***

The final part of the study involved the evaluation of the clinical sensitivity of the above predictive measures at

estimating intellectual, verbal memory, and language performance in the same normal elderly sample. The advantage of this investigation is that it not only provided information with regard to the clinical utility of a measure for detecting cognitive decline based on the error of prediction within a normal elderly sample, but it also provided a quantitative basis for comparison of the concurrent validity of various predictors with respect to a specific criterion measure. Examination of the clinical sensitivity of potential predictors of intellectual, verbal memory, and language performance were included at each respective part of the study.

Nature and Goal of the Study

As there was only limited data in the literature regarding the above premorbid VIQ measures, present ability measures, and demographic variables at estimating performance in memory and language domains for normal elderly individuals, the present study remained largely investigative in an effort to determine the most useful measures for predicting memory and language performance in a normal elderly sample. Given the investigative nature of the study, no specific a priori hypotheses were proposed. The goal of the study was to identify the most useful measures for estimating memory and language performance in normal elderly individuals so that these measures could serve as

potential predictors of premorbid functioning in memory and language domains for patients with early Alzheimer's disease (AD). If decline could be detected at an early stage of AD with a reliable and sensitive estimate of premorbid ability in memory and language domains, then diagnosis and intervention could be possible much earlier in the disease process than currently available.

Method

Participants

The normative sample consisted of 98 community-dwelling volunteers ranging in age from 53 to 89 years (M = 71.9, S.D. = 7.2). The average education level was 13.6 years (S.D. = 3.1, Range = 6 to 22 years). Twenty-one percent of the participants were male (21 males, 77 females). The primary language of all participants was English. All participants were either living independently or in retirement homes in the community. Participants were recruited through poster advertisements at social/recreational organizations for seniors. Each volunteer received $10.00 for participating in the study. Volunteers were screened for a history of medical and psychiatric problems through a clinical interview, which included documentation of currently perscribed medications. Volunteers were included for participation in the study if

their medical history was compatible with five criteria for research with normal elderly participants as specified by Malec, Ivnik, and Smith (1993). These criteria included: 1) no evidence of active central nervous system or psychiatric conditions that would adversely affect cognition; 2) no complaint of cognitive difficulty during history-taking, and no examination findings indicating cognitive compromise; and 3) no psychiatric medication used in amounts expected to compromise cognition. Malec et al. (1993) further indicated that: 4) prior histories of disorders potentially affecting cognition (e.g., head injury, substance abuse) may be present, provided it is clear that the condition is no longer active and there is no residual cognitive deficit; and 5) chronic medical illness (e.g., diabetes, hypertension, cardiac problems) may be present, provided that the condition is not associated with compromised cognition. Participants were also screened for global cognitive functioning and depressive symptomatology and were excluded from the study if they scored lower than 26 on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975; Roper, Bieliauskas, & Peterson, 1996), or higher than 13 on the Geriatric Depression Scale (GDS; Brink et al. 1982).

Informed Consent

All procedures and methods for obtaining informed

consent from participants were approved by the Human Research Ethics Committee, University of Victoria.

Procedures and Measures

All participants were tested individually, either at their home, their respective social/recreational organization, or the Psychology Clinic at the University of Victoria. A detailed neuropsychological test battery was administered to each participant and included the following tests listed by cognitive domain:

1. ***Intelligence:*** *Wechsler Adult Intelligence Scale-Revised* (WAIS-R; Wechsler, 1981): Verbal IQ (VIQ); Ryan, Paolo, and Brungardt (1990) normative data used for persons 75 years and older.

2. ***Premorbid Intelligence:*** *(a) National Adult Reading Test* (NART; Nelson, 1982): total error score; *(b) National Adult Reading Test-Revised* (NART-R; Blair & Spreen, 1989): total error score; *(c) Reading subtest of the Wide Range Achievement Test-3* (WRAT-3, Blue Version; Wilkinson, 1993): total raw score; *(d) Spot-the-Word subtest of the Speed and Capacity of Language-Processing Test* (SCOLP; Baddeley, Emslie, & Nimmo-Smith, 1992): total raw score.

3. ***Verbal Memory:*** *(a) California Verbal Learning Test* (CVLT; Delis et al. 1987): total number of words recalled from the recency component (last four words) of the word list over five learning trials, and the total number of words recalled from the long delay free recall trial.

4. ***Language:*** *(a) 60-Item Boston Naming Test* (BNT; Kaplan, Goodglass, & Weintraub, 1983): total number of correct responses (spontaneous responses and correct responses to stimulus cues); *(b) Word-Picture Matching subtest of the Boston Diagnostic Aphasia Examination* (BDAE; Goodglass & Kaplan, 1983): total number of correct responses, and latency (in seconds) to complete the subtest; *(c) Sentence Repetition subtest of the Neurosensory Centre Comprehensive Examination for Aphasia* (NCCEA: Spreen & Benton, 1969): total corrected score; *(d) Peabody Picture Vocabulary Test-Revised* (PPVT-R, Form L: Dunn & Dunn, 1981): total raw score.

Several WAIS-R VIQ regression equations were also examined in the above sample for estimating intellectual, verbal memory, and language performance. Details for calculating WAIS-R VIQ scores based on these equations are presented in Appendix A.

Data Analysis

1. *Development of Predictive Regression Equations*

Regression equations were developed for measures that demonstrated the greatest concurrent validity for estimating intellectual, verbal memory, and language performance in the above normative sample. Stepwise multiple regression was employed to determine which measures contributed significantly to the prediction of intellectual, verbal memory, and language performance, with a conservative p value set for entry into the regression analysis ($p < .01$) to reduce the likelihood of a Type 1 error. However, a more liberal p value (i.e., $p < .05$) was adopted in circumstances in which only a limited number of predictors contributed to the estimation of a criterion measure, to improve the sensitivity of the prediction equation. Although this increased the possibility of a Type 1 error in such circumstances, cross-validation calculations accounted for this problem (as discussed in the fifth section of the data analysis), because cross-validation calculations corrected for the shrinkage in the accuracy of equations at predicting the criterion in a new sample.

2. *Sensitivity of Predictive Estimators for Detecting Cognitive Decline*

The sensitivity of new predictive regression equations

for detecting decline in their respective cognitive domains was evaluated according to two approaches outlined by Graves et al. (in press) including:

(a) The Standard Error of Estimate (SEe) method: This method utilized the SEe from the predictive regression equation, and was dependent on a normal distribution of discrepancy scores (obtained score minus predicted score). Two scores were calculated using this approach including the cut-off discrepancy score and the cognitive decline sensitivity score. The *cut-off discrepancy score* was considered to reflect the score necessary for 95% correct classification of normals (i.e., 95% specificity), and was calculated by multiplication of the SEe by 1.645 (z value; 95% (one-tail) confidence level). The *cognitive decline sensitivity score* was defined as the score reflecting the minimum cognitive decline that could be successfully detected 80% of the time (80% sensitivity), while using a cut-off that would correctly classify 95% of normals (95% specificity). The cognitive decline sensitivity score was calculated by multiplication of the SEe by 2.487. The SEe used for these calculations was the *cross-validation SEe* which incorporates a correction for the shrinkage expected on cross-validation (i.e., the SEe expected when the regression equation is used in a new sample). Darlington (1968) presented formulas for determining cross-validation

SEe. The reader is referred to the fifth section of the data analysis which discusses cross-validation of predictive regression equations.

(b) The Percentile Table method: This method required calculation of the complete percentile table for the discrepancies between obtained and predicted scores. This method was not dependent on a normal distribution of discrepancy scores. The cut-off discrepancy score and the cognitive decline sensitivity score were derived directly from the percentile table utilizing this method. The *cut-off discrepancy score* (95% specificity) was determined by locating the 5th percentile of the table (i.e., the negative discrepancy exceeded by 5% of normals). A negative discrepancy represented an overestimation of the obtained score by the predicted score. The cut-off sensitivity score (80% sensitivity) was determined by locating the 80th percentile of the table (i.e., the positive discrepancy exceeded by 20% of normals). A positive discrepancy represented an underestimation of the obtained score by the predicted score. The *cognitive decline sensitivity score* was then calculated by adding the two absolute discrepancy scores (i.e., the cut-off discrepancy score and the cut-off sensitivity score). A drawback of the percentile table method is that it does not include a correction for the shrinkage expected on cross-validation.

3. ***Normality of the Frequency Distribution***

Normality of the frequency distribution of discrepancy scores was determined by examining skewness and kurtosis values provided by each distribution. Normality was rejected if the ratio of skewness to its standard error was less than -2 or greater than +2 (SPSS 7.5 Windows; Norusis, 1996). However, Tabachnick and Fidell (1989) indicated that as standard errors for both skewness and kurtosis contain N (sample size), the null hypothesis is likely to be rejected in large samples when there are only minor deviations from normality. They indicated that in a large sample, a variable with significant skewness (or kurtosis) often does not deviate enough from normality to make a realistic difference in the analysis. Therefore, cut-off discrepancy scores and cognitive decline sensitivity scores for predictive equations in the current study were calculated by both the SEe method and the percentile table method outlined by Graves et al. (in press) to examine the impact of the distribution of discrepancy scores on the calculation of sensitivity scores in this sample.

4. ***Standard of Comparison***

The clinical utility of predictive regression equations was evaluated by comparing the cognitive decline sensitivity score of the predictive measure to the cognitive decline sensitivity score derived from test-retest reliability data

of the criterion variable with healthy participants. Graves et al. (in press) suggested that cognitive decline sensitivity scores calculated from the test-retest reliability data of the criterion variable could serve as a standard of comparison for the sensitivity of other methods for detecting cognitive decline, as sensitivity scores based on test-retest data are calculated from the ideal situation in which previous criterion scores are available. Cross-validation cognitive decline sensitivity scores based on test-retest data for each criterion variable (i.e., WAIS-R VIQ, CVLT long delay free recall score, BNT) were included in the current study as a standard of comparison to evaluate the sensitivity of predictive regression equations for detecting cognitive decline in clinical samples.

5. ***Cross-validation of Predictive Regression Equations***

Cross-validation calculations were performed for new predictive regression equations to determine the accuracy of each equation for predicting the criterion measure for a new case (i.e., participant not in the original sample; Graves, 1998). Two approaches discussed by Graves (1998) were used to examine the accuracy of prediction of each equation in a new sample including: 1) cross-validation of the multiple correlation (R; Cattin, 1978, 1980); and 2) cross-validation of the standard error of estimate (Darlington, 1968). Graves (1998) viewed cross-validation of the SEe as the more

accurate estimate of the actual magnitude of the overall error of prediction for each equation, as compared to the cross-validation R, given that it accounts for more of the possible error in the equation (i.e., error in the constant term). The cross-validation R, based on Cattin's (1978, 1980) approach, and the cross-validation SEe, based on Darlington's (1968) approach, were calculated for each equation using a computer program developed by Graves (1998). Cross-validation sensitivity scores (i.e., cut-off discrepancy score and cognitive decline sensitivity score) were provided for each equation, based on the cross-validation SEe.

Results

All statistics were calculated using SPSS for Windows Version 7.5 (Norusis, 1996).

VIQ Predicted by VIQ Measures

The first part of this study examined the ability of various estimation methods at predicting actual WAIS-R VIQ scores of the current normal elderly Canadian sample. As the normative data for the WAIS-R (Wechsler, 1981) does not extend beyond 74 years of age, it was originally planned to calculate VIQs for persons 75 years and older using normative data provided by Ryan et al. (1990). However, the

WAIS-R VIQs based on the Ryan et al. (1990) norms were found to correlate significantly with age in the current sample ($r = 0.299$, $p<.003$). This suggested that the Ryan et al. (1990) norms failed to adequately correct for age in this sample, as IQ is not considered to vary with age. Accordingly, WAIS-R VIQs based on the Wechsler (1981) normative data, with 70-74 year old norms applied to VIQ scores for persons aged 75+ years were considered to be more representative of VIQ scores for this sample than the scores obtained from the Ryan et al. (1990) norms since they did not correlate significantly with age ($r = 0.072$). Therefore, the following results employ WAIS-R VIQ scores based on the Wechsler (1981) norms as the criterion measure.

1. ***WAIS-R VIQ Normative Data***

WAIS-R VIQ statistics for the present sample ($N = 98$) were: $M = 112.65$; $S.D. = 9.17$; Range = 82.00 to 132.00. Normative data for WAIS-R Verbal subtests (age-scaled scores) are presented in Table 1.

2. ***Cross-validation of Existing VIQ Prediction Equations***

Descriptive statistics for VIQ prediction equations at estimating WAIS-R VIQs in the current normal elderly Canadian sample are presented in Table 2. The efficacy of existing VIQ equations at predicting WAIS-R VIQs in the current sample was determined by calculating measures of the

accuracy of prediction of each of the equations. Table 3 shows the correlation and statistics for the actual minus predicted VIQ scores. The mean error in prediction of actual WAIS-R VIQs per IQ range classification for existing VIQ prediction equations is presented in Table 4. The cognitive decline sensitivity scores for the original sample cross-validation results of the seven selected VIQ prediction equations together with the cross-validation results with the new sample are shown in Table 5. Also included in Table 5 are cross-validation standard error of estimate (SEe) and cognitive decline sensitivity values based on WAIS-R test-retest reliability data over a one-year period for healthy elderly Canadians aged 50 to 84 years ($r = 0.86$, S.D. = 9.9, $N = 101$; Snow, Tierney, Zorzitto, Fisher, & Reid, 1989). These latter scores were included to serve as a standard of comparison regarding the sensitivity of premorbid VIQ equations for detecting cognitive decline in clinical samples, given that these values were derived from the ideal situation in which actual previous VIQ scores were available (Graves et al. in press).

3. ***Development of New VIQ Prediction Equations***

The study then examined the utility of existing premorbid VIQ measures (i.e., NART, NART-R, WAIS-R Vocabulary), newly proposed premorbid VIQ measures (i.e., WRAT-3 reading subtest, STW), and demographic variables

(i.e., age, sex, education) at predicting WAIS-R VIQs in the above normal elderly Canadian sample. The purpose of this was to: 1) obtain normative data for VIQ measures for elderly Canadians, and 2) develop a new clinically useful WAIS-R VIQ predictive equation for detecting cognitve decline in clinical samples.

Table 1

Normative Data for WAIS-R Verbal Subtests

	Age-Scaled Scores			
			Range	
Subtest	Mean	S.D.	Min.	Max.
Information	12.51	2.17	5.00	16.00
Digit Span	11.92	2.74	6.00	19.00
Vocabulary	12.46	1.80	7.00	17.00
Arithmetic	11.91	2.67	6.00	17.00
Comprehension	11.36	1.76	6.00	15.00
Similarities	12.93	1.65	7.00	18.00

Note. S.D. = Standard Deviation; Min. = Minimum; Max. = Maximum.

Table 2

Descriptive Statistics for Existing VIQ Predictive Equations

Study	VIQ Preds. in Eq.	M	S.D.	Range Min.	Range Max.
Carswell et al. (1997)	NART + WAIS-R Voc.	114.3	7.5	89.9	129.3
Krull et al. (1995)	WAIS-R Voc. + Demog.	110.9	6.7	90.4	121.2
Blair & Spreen (1989)	NART-R	113.8	8.6	87.8	126.9
Nelson & Willison(1991)	NART	111.9	7.9	90.0	125.0
Ryan & Paolo (1992)	NART	116.6	8.1	93.9	130.1
Carswell et al. (1997)	NART	117.7	6.7	98.8	128.9
Barona et al. (1984)*	Demog.	109.9	7.8	89.1	121.5
Barona et al. (1984)	Demog.	109.7	7.9	88.6	120.5

Note. NART = National Adult Reading Test; WAIS-R Voc. = Wechsler Adult Intelligence Scale-Revised Vocabulary

subtest; VIQ = Verbal Intelligence Quotient; Demog. = Demographics; NART-R = New Adult Reading Test-Revised; Barona et al. (1984)* = Barona equation with Helmes (1996) age modification; Preds. = Predictors; Eq. = Equation; M = Mean, S.D. = Standard Deviation, Min. = Minimum; Max. = Maximum.

Table 3

Cross-validation Accuracy of Existing VIQ Prediction Equations

Study	VIQ Predictors in equation	r	Mean Er. (IQPt)	S.D. Er. (IQPt)
Carswell et al. (1997)	NART + WAIS-R Vocab.	0.766	-1.62	5.91
Krull et al. (1995)	WAIS-R Vocab. + Demographics	0.744	+1.74	6.13
Blair & Spreen (1989)	NART-R	0.694	-1.12	6.98
Nelson & Willison (1991)	NART	0.645	+0.70	7.29
Ryan & Paolo (1992)	NART	0.645	-3.98	7.33
Carswell et al. (1997)	NART	0.645	-5.09	7.06

Barona et al. (1984)*	Demographics	0.472	+2.68	8.81
Barona et al. (1984)	Demographics	0.470	+2.93	8.84

Note. All correlations were significant at $p < .01$; Mean errors with negative values (-) represented an overestimation of actual WAIS-R VIQs whereas, mean errors with positive values (+) represented an underestimation of actual WAIS-R VIQ scores; NART = National Adult Reading Test; NART-R = New Adult Reading Test-Revised; WAIS-R = Wechsler Adult Intelligence Scale-Revised; Vocab. = Vocabulary subtest; r = Multiple correlation; Mean Er. = Mean Error; S.D. Er. = Standard Deviation Errors; IQPt = IQ Points; Barona et al. (1984)* = Barona equation with Helmes (1996) age modification.

(a) Descriptive Statistics and Normative Scores for VIQ Predictors

Mean, standard deviation, and range values for all VIQ predictors are presented in Table 6. Demographic data for the above sample has been summarized in the Methods section. Normative data for VIQ predictors for elderly Canadians are presented in Table 7, according to age. Pearson (r) correlations of VIQ predictors with WAIS-R VIQ are presented in Table 8.

Table 4

Mean Error in WAIS-R VIQ Prediction for Existing VIQ Equations by IQ Range Classification

Study	VIQ Prs. in Eq.	IQ Range 80-89	90-109	110-119	120-129	130+
Carswell et al. (1997)	NART + W-R Voc.	-10.7	-4.9	-1.3	+3.7	+1.9
Krull et al. (1995)	W-R Voc. + Demog.	-10.5	-1.9	+1.9	+7.7	+9.5
Blair & Spreen (1989)	NART-R	-7.9	-3.9	-1.4	+3.7	+6.4
Nelson & Willison (1991)	NART	-9.5	-3.3	+1.3	+6.2	+7.7
Ryan & Paolo (1992)	NART	-13.9	-7.8	-3.5	+1.5	+2.9
Carswell et al. (1997)	NART	-18.3	-9.8	-4.4	+1.2	+3.7
Barona et al. (1984)*	Demog.	-17.9	-1.5	+3.0	+9.5	+13.9
Barona et al. (1984)	Demog.	-17.9	-1.3	+3.3	+9.6	+14.3

Note. Mean error values are expressed in IQ points; Mean errors with negative values (-) represented an overestimation of actual WAIS-R VIQs whereas, mean errors with positive values (+) represented an underestimation of actual WAIS-R VIQ scores; WAIS-R VIQ = Wechsler Adult Inteliigence Scale-Revised Verbal Intelligence Quotient; Prs. = Predictors; Eq. = Equation; W-R Voc. = WAIS-R Vocabulary subtest; NART = National Adult Reading Test; NART-R = New Adult Reading Test-Revised; Demog = Demographics; Barona et al. (1984)* = Barona equation with Helmes (1996) age modification.

Examination of Table 8 identified six predictor variables (NART errors, NART-R errors, WAIS-R Vocabulary age-scaled score (VAS), WRAT-3 Reading raw score, STW score, education) that correlated significantly with WAIS-R VIQ ($p < .01$). Further analyses showed that WAIS-R VIQ scores increased with advancing education with a significant difference in WAIS-R VIQ scores for persons with 13 years+ education as compared to persons with less than 13 years of education ($t(96) = 5.27$, $p < .001$). Accordingly, for the subsequent regression, education was coded as -1 for education < 13 years, and +1 for education $\geq$ 13 years.

Table 5

Cross-validation of VIQ Equations: Cognitive Decline Sensitivity

Study	VIQ Predictors in equation	Org. SEe	Obs. SqMSe	Org. CDS	Obs. CDS
Cross-validation Standard of Comparison					
Snow et al. (1989)	WAIS-R (1 yr previous)	5.13		12.8	
Carswell et al. (1997)	NART + WAIS-R Vocab.	7.05	6.10	17.5	15.2
Scott et al. (1997)*	WAIS-R Vocab. + Demographics	6.32	6.34	15.7	15.8
Blair & Spreen (1989)	NART-R	6.68	7.03	16.6	17.5
Nelson & Willison (1991)	NART	7.30*	7.28	18.2*	18.1
Ryan & Paolo (1992)	NART	7.70*	8.31	19.1*	20.7
Carswell et al. (1997)	NART	8.09	8.67	20.1	21.6
Barona et al. (1984)	Demographics	11.81	9.27	29.4	23.1

Note. All values are expressed in IQ points; Scott et al. (1997)* study contained data for cross-validation calculations of Krull et al. (1995) equations; Barona et al. (1984) equation with the Helmes (1996) age modification was not included in the table as there was no preexisting data for comparison; * values = original scores that could not be cross-validated due to a lack of supporting data in the original study; Org. SEe = Original study standard error of estimate; Obs. SqMSe = Observed square root of the mean square error; Org. CDS = Original study cognitive decline sensitivity score (2.487SEe); Obs. CDS = Observed cognitive decline sensitivity score (2.487SqMSe).

Table 6

Mean, Standard Deviation and Range Values for VIQ Predictors

			Range	
Variable	Mean	S.D.	Min.	Max.
NART errors	13.53	6.92	2.00	33.00
NART-R errors	16.77	9.68	2.00	46.00
WRAT-3 Rd (Raw Sc.)	51.15	3.66	38.00	57.00
STW	53.56	4.92	38.00	60.00
WAIS-R VAS	12.46	1.80	7.00	17.00

Note. WRAT-3 Rd (Raw Sc.) = Wide Range Achievement Test-

Revised Reading subtest (Raw Score); STW = Spot-the-Word subtest score; WAIS-R VAS = Wechsler Adult Intelligence Scale-Revised Vocabulary subtest age-scaled score; S.D.= Standard Deviation; Min. = Minimum, Max. = Maximum.

Table 7

Mean and Standard Deviation for VIQ Predictors by Age

	Age			
Variable	50-59 (N=5)	60-69 (N=32)	70-79 (N=47)	80-89 (N=14)
NART er.	15.0(10.9)	13.0(6.9)	14.1(6.9)	12.1(5.3)
NART-R er.	19.0(16.4)	16.1(9.5)	17.5(9.9)	14.9(6.5)
WRAT-3 Rd.	51.0(5.0)	51.6(3.7)	50.6(3.5)	51.7(3.3)
STW	52.0(6.9)	53.0(4.7)	53.6(5.1)	54.9(3.6)
WAIS-R VAS	12.4(3.3)	12.5(1.5)	12.4(1.8)	12.5(1.7)

Note. (N) = number of participants in each age group in sample; NART er. = NART errors; NART-R er. = NART-R errors; WRAT-3 Rd. = WRAT-3 Reading subtest raw score; STW = Spot-the-Word score; WAIS-R VAS = WAIS-R Vocabulary subtest age-scaled score.

(b) Regression Analysis

Stepwise multiple regression ($p < .01$ for entry) was

employed to determine which of the five significant predictor variables identified in Table 8 (i.e., NART-R errors, WAIS-R VAS, WRAT-3 Rd., STW, education: point biserial) were most effective at estimating WAIS-R VIQs in the above sample. NART errors were not included in the regression analysis as the NART-R, the North American version of this test, was more appropriate for consideration with the above Canadian sample, and, in fact, produced a higher correlation with WAIS-R VIQ than NART errors.

Table 8

Pearson (r) correlations of VIQ Predictors with WAIS-R VIQ

Variable	WAIS-R VIQ
NART errors	-0.65*
NART-R errors	-0.69*
WAIS-R VAS	0.74*
WRAT-3 Rd. (Raw score)	0.55*
STW	0.68*
Age	0.07
Sex	0.18
Education (continuous)	0.44*
Education (point biserial, $\leq$ 13 years)	0.47*

* $p < .01$

Only two variables, WAIS-R Vocabulary subtest age-scaled score (VAS) and STW scores, contributed significantly to the prediction of WAIS-R VIQs, accounting for 63% of the variance in WAIS-R VIQs. The regression equation based on these results was:

estimated WAIS-R VIQ = 43.038 + 2.601 (VAS) + 0.695 (STW)

The multiple correlation (R) for the above equation was 0.795. A scatterplot of WAIS-R VIQ predicted scores to actual/obtained WAIS-R VIQ scores is presented in Figure 1 for the new predictive VIQ regression equation.

The sensitivity of the new WAIS-R VIQ predictive equation for detecting cognitive (VIQ) decline was examined in the context of discrepancy values between obtained and predicted VIQ scores. The frequency distribution of discrepancy values between obtained and predicted VIQ scores is shown in Figure 2.

Given the normal distribution of WAIS-R VIQ discrepancy scores (skewness = 0.233, standard error = 0.244, (ratio = +0.95); kurtosis = -0.130, standard error = 0.483, (ratio = -0.27)), as illustrated in Figure 2, sensitivity scores for detecting cognitive decline (i.e., cut-off discrepancy score; cognitive decline sensitivity score) were calculated according to the two approaches outlined by Graves et al. (in press), i.e., the standard error of estimate (SEe) method and the percentile table approach.

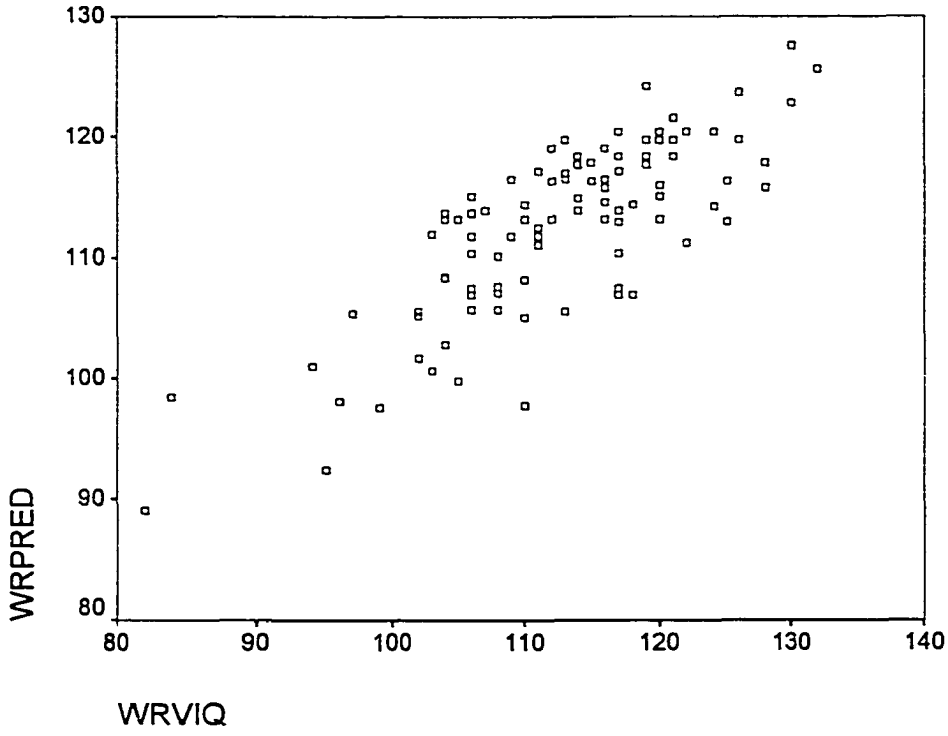

Figure 1. Scatterplot of predicted scores to WAIS-R VIQs

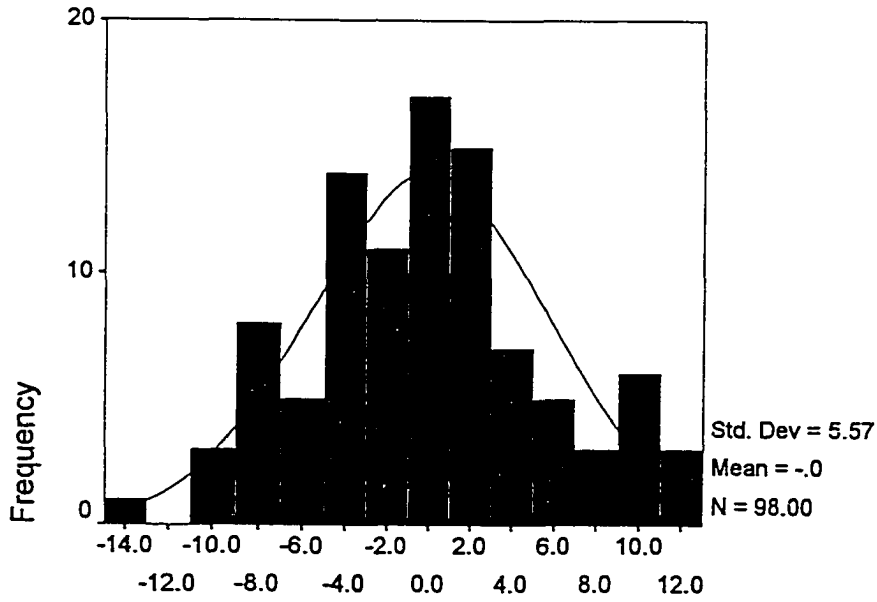

Figure 2. Frequency distribution of discrepancy values between obtained and predicted WAIS-R VIQ scores

The percentile table for the discrepancy values between obtained and predicted VIQ scores is presented in Table 9. The cut-off discrepancy score was identified as 9 VIQ points based on locating the 5th percentile of the table (i.e., the negative discrepancy exceeded by about 5% of normals). The cut-off sensitivity score was identified as 4 VIQ points based on locating the 80th percentile of the table (i.e., the positive discrepancy exceeded by about 20% of normals). Addition of these two absolute discrepancy scores produced a cognitive decline sensitivity score of 13 VIQ points.

Cross-validation calculations were performed for the new predictive regression equation to determine the accuracy of this equation for predicting a WAIS-R VIQ score for a participant not in the original sample. Correction of the observed multiple correlation ($R = 0.795$) with 5 available predictors considered in the regression analysis produced an estimated cross-validation multiple correlation of 0.772 ($R^2 = 0.595$), based on Cattin's (1978, 1980) approach. Calculation of the standard error of estimate (n.b., SEe = 5.6237 as reported by SPSS 7.5) utilizing Darlington's (1968) approach produced an estimated cross-validation SEe of 5.896. This value, which incorporates a correction for shrinkage, represents the expected accuracy of prediction of this equation in a <u>new</u> sample. Cross-validation sensitivity scores based on the estimated cross-validation SEe included a cut-off discrepancy score of 9.70 VIQ points and a

cognitive decline sensitivity score of 14.66 VIQ points.

Table 9

Percentile Table of Discrepancy Values between Obtained and Predicted WAIS-R VIQs (VAS and STW based)

Difference (IQ Points)	Est.VIQ Higher than Obt.VIQ*		Est.VIQ Lower than Obt.VIQ**	
	(N)	(%)	(N)	(%)
1	16	16.3	16	16.3
2	4	4.0	7	7.1
3	9	9.2	3	3.0
4	5	5.1	4	4.0
5	2	2.0	1	1.0
6	3	3.0	4	4.0
7	5	5.1	2	2.0
8	3	3.0	1	1.0
9	3	3.0	3	3.0
10	-	-	3	3.0
11	-	-	-	-
12	-	-	3	3.0
13	-	-	-	-
14	1	1.0	-	-

Note. *(negative discrepancy); **(positive discrepancy);

Est. = Estimated; Obt. = Obtained; (N) = number of scores; (%) = percentage of sample.

The accuracy of the new WAIS-R VIQ regression equation at predicting extreme versus median range VIQ scores was also examined. Mean and range values of discrepancy scores (i.e., obtained minus predicted VIQs) are presented in Table 10, by IQ range classification.

Table 10

Accuracy of Predicted WAIS-R VIQs (VAS and STW based) at Estimating Extreme versus Median Obtained WAIS-R VIQs

IQ Range	(N)	Discrepancy (IQ Points) Mean	Range
130+	3	+5.29	+2.44 to +7.12
120-129	18	+5.33	-0.67 to +12.23
110-119	44	+0.11	-7.07 to +12.28
90-109	31	-3.12	-9.69 to +5.19
80-89	2	-10.73	-14.42 to -7.04

Note. (N) represents the number of persons from the sample in each respective IQ range.

As there has been some concern expressed in the

literature regarding the ability of the WAIS-R Vocabulary subtest to resist decline in early Alzheimer's disease, a new predictive WAIS-R VIQ equation was developed with WAIS-R Vocabulary excluded from the regression analysis. Stepwise multiple regression ($p < .01$ for entry) was employed to determine which of the four predictor variables including NART-R errors, WRAT-3 Reading scores, STW scores and education were most effective at estimating WAIS-R VIQs. Two variables including NART-R errors and STW scores contributed significantly to the prediction of WAIS-R VIQs, accounting for 55% of the variance in WAIS-R VIQ scores. The regression equation based on these results was:

estimated WAIS-R VIQ = 81.452 - 0.390 (NART-R errors) + 0.705 (STW)

The multiple correlation (R) for the above equation in this sample was 0.738. A scatterplot of WAIS-R VIQ predicted scores to actual/obtained WAIS-R VIQ scores is presented in Figure 3.

The frequency distribution of discrepancy values between obtained and predicted WAIS-R VIQs based on the new regression equation is shown in Figure 4. Given the normal distribution of WAIS-R VIQ discrepancy scores (skewness = 0.326, standard error = 0.244, (ratio = +1.33); kurtosis = 0.074, standard error = 0.483, (ratio = +0.15)), both the SEe method and the percentile table method were employed in the calculation of sensitivity scores.

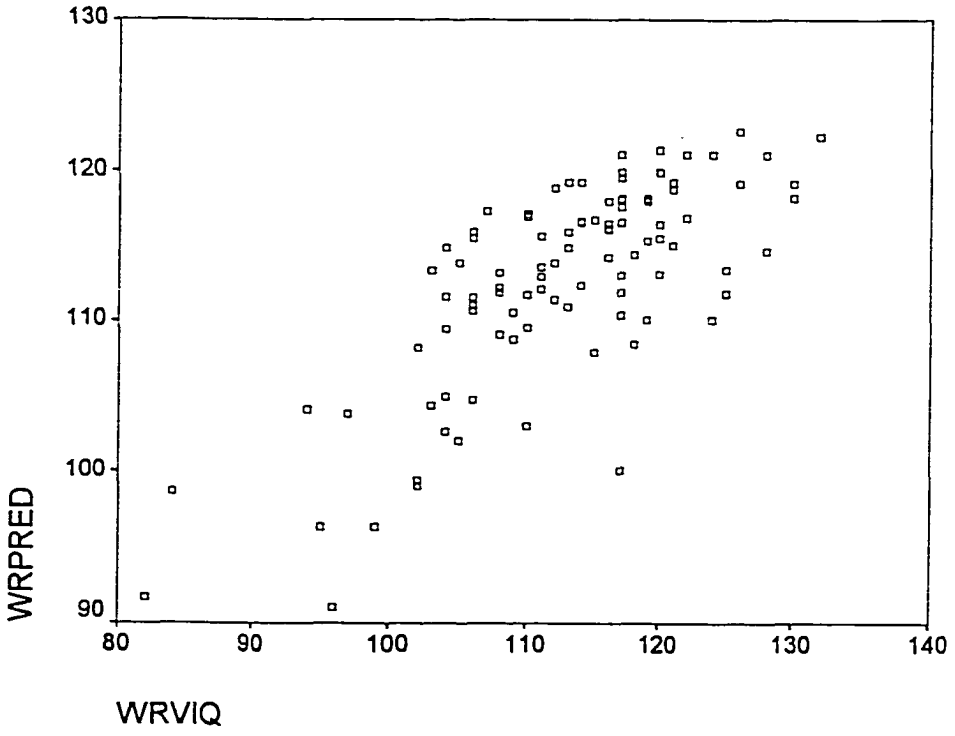

Figure 3. Scatterplot of predicted scores (NART-R, STW based to WAIS-R VIQs

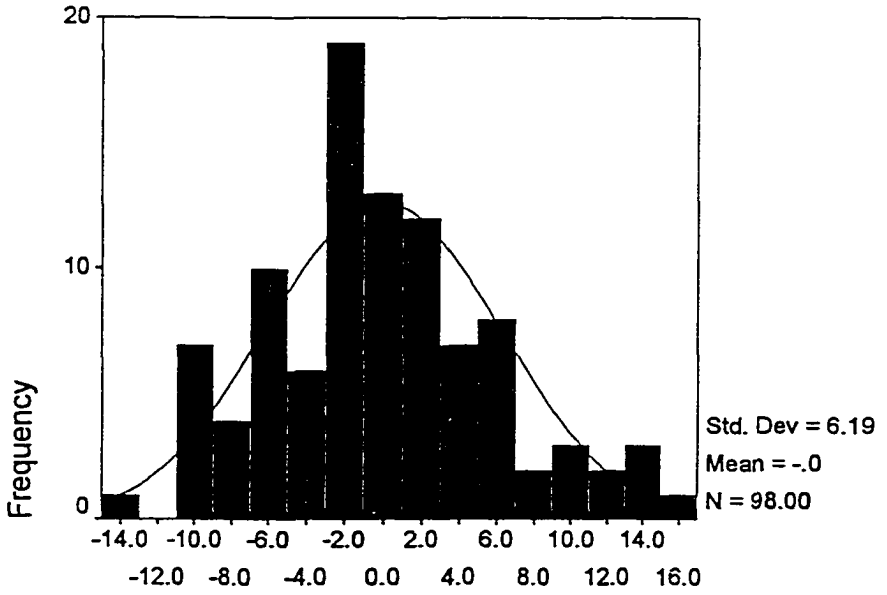

Figure 4. Frequency distribution of discrepancy values between obtained and predicted WAIS-R VIQs (NART-R, STW based)

The percentile table for the discrepancy values between obtained and predicted WAIS-R VIQs is presented in Table 11. The cognitive decline sensitivity score based on the table was considered to be 15 VIQ points, with a cut-off discrepancy score of 10 VIQ points, and a cut-off sensitivity score of 5 VIQ points.

Cross-validation calculations were performed for this new predictive regression equation to determine the accuracy of the equation for predicting the WAIS-R VIQ score of a participant not in the original sample. Correction of the multiple correlation ($R = 0.738$) with 4 available predictors considered in the regression analysis produced an estimated cross-validation multiple correlation of 0.714 ($R^2 = 0.510$) based on Cattin's (1978, 1980) approach. Calculation of the cross-validation standard error of estimate (n.b., SEe = 6.2556 as reported by SPSS 7.5) utilizing Darlington's (1968) approach produced an estimated cross-validation SEe of 6.488. Cross-validation sensitivity scores based on the estimated cross-validation SEe were a cut-off discrepancy score of 10.67 VIQ points, and a cognitive decline sensitivity score of 16.14 VIQ points.

The accuracy of this new WAIS-R VIQ regression equation at predicting extreme versus median range VIQ scores was also examined. Mean and range values of discrepancy scores (i.e., obtained minus predicted VIQs) are presented in Table 12, by IQ range classification.

Table 11

Percentile Table of Discrepancy Values between Obtained and Predicted WAIS-R VIQs (NART-R errors, STW based)

Difference (IQ Points)	Est. VIQ Higher than Obt. VIQ		Est. VIQ Lower than Obt. VIQ	
	(N)	(%)	(N)	(%)
1	18	18.4	13	13.3
2	6	6.1	7	7.1
3	2	2.0	5	5.1
4	4	4.0	2	2.0
5	6	6.1	3	3.0
6	4	4.0	5	5.1
7	3	3.0	1	1.0
8	1	1.0	1	1.0
9	2	2.0	2	2.0
10	5	5.1	1	1.0
11	-	-	2	2.0
13	-	-	3	3.0
14	1	1.0	-	-
16	-	-	1	1.0

Note. Est. = Estimated; Obt. = Obtained.

Table 12

Accuracy of Predicted WAIS-R VIQs (NART-R, STW based) at Estimating Extreme versus Median Obtained WAIS-R VIQs

IQ Range	(N)	Discrepancy (IQ Points) Mean	Range
130+	3	+10.74	+9.73 to +11.72
120-129	18	+5.62	-1.41 to +13.93
110-119	44	+0.15	-7.18 to +16.92
90-109	31	-3.78	-10.92 to +4.92
80-89	2	-12.19	-14.67 to -9.71

Note. (N) represents the number of persons from the sample in each respective IQ range.

As several predictive VIQ measures independently produced significant correlations with WAIS-R VIQ, cross-validation sensitivity scores were calculated for each of these measures (based on the SEe method) to determine their independent clinical utility. A summation of the efficacy of VIQ predictors at estimating WAIS-R VIQs and detecting cognitive decline is presented in Table 13. WAIS-R test-retest cross-validation sensitivity scores were included in Table 13 for comparison with the sensitivity of other methods for detecting cognitive decline. WAIS-R VIQ

regression equations for independent VIQ predictors are presented in Appendix B.

Table 13

Efficacy of VIQ predictors at Estimating WAIS-R VIQs

Predictor(s)	Cross-validation SEe (IQPT)	COD (IQPT)	CDS (IQPT)
Standard of Comparison			
WAIS-R (1 yr previous) Snow et al. (1989)	5.13	8.44	12.8
VAS/STW	5.90	9.70	14.7
VAS	6.27	10.32	15.6
NART-R errors/STW	6.49	10.67	16.1
NART-R errors	6.71	11.03	16.7
STW	6.79	11.16	16.9
NART errors	7.12	11.71	17.7
WRAT-3 Reading raw score	7.45	12.74	19.3
Education (point biserial)	8.21	13.5	20.4

Note. SEe = Standard error of estimate; COD = Cut-off discrepancy score; CDS = Cognitive decline sensitivity score; IQPT = IQ Points; yr = year.

Verbal Memory Predicted by VIQ Measures, Present Ability Measures and Demographic Variables

The second part of the study examined the concurrent validity of VIQ measures considered in part one, measures of present ability (i.e., WAIS-R Digits Forward: raw score; NCCEA Sentence Repetition: total corrected score; CVLT Recency score: total words recalled from the recency component of the word list over 5 learning trials) and demographic variables (i.e., age, sex, education) at estimating long delay free recall performance on the CVLT (total words recalled on the long delay free recall trial). One participant was not included in the following CVLT analyses due to questionable/impaired performance on the CVLT, for a total sample size of 97 participants.

1. ***CVLT Normative Data***

Statistics for the criterion measure, CVLT long delay free recall score (CVLTLDFR), for the present sample (N = 97) were: M = 11.11; S.D. = 2.84; Range = 6.00 to 16.00. Normative data for all CVLT scores by age are presented in Table 14.

2. ***VIQ Measures***

The efficacy of existing VIQ equations, newly developed VIQ equations (previous section), and independent VIQ predictors at estimating long delay free recall performance

on the CVLT (i.e., CVLTLDFR) was determined by examining Pearson (r) correlations of VIQ measures with CVLTLDFR scores, presented in Tables 15 and 16. The results of Tables 15 and 16 showed that none of the VIQ measures considered in the current study produced significant correlations with CVLTLDFR scores at $p < .01$, or $p < .05$. It is interesting to note however, that WAIS-R VIQ correlated significantly with CVLTLDFR at $p < .01$ ($r = 0.278$).

Table 14

CVLT Normative Data by Age

	Age (years)			
	< 70 (N = 37)		≥ 70 (N = 60)	
Score	M	S.D.	M	S.D.
Recall Measures (total words):				
T1-5	55.57	8.42	49.47	10.42
A1	6.67	1.62	6.03	1.78
A5	13.65	2.16	12.33	2.37
B	7.29	1.84	6.23	1.89
SDFR	11.86	2.86	10.00	2.62
SDCR	12.35	3.06	10.92	2.42
LDFR	12.11	2.85	10.50	2.67
LDCR	12.14	2.89	11.02	2.54

Learning Scores:				
SEM	2.15	0.86	1.91	0.81
SER	1.84	1.51	2.46	2.01
%PRIM	28.03	4.94	30.10	7.35
%MID	44.41	4.84	43.47	7.31
%REC	27.29	5.61	26.42	5.04
SLOPE	1.59	0.48	1.52	0.49
RECALL CONS.	87.67	7.29	82.40	8.76
Recall Errors:				
PERS	0.95	1.05	0.97	1.16
FRINTR	1.73	2.81	1.96	2.22
CRINTR	1.08	1.44	1.38	1.55
TOTINTR	2.81	4.01	3.35	3.49
Recognition Scores:				
RECHITS	15.14	1.00	14.45	1.67
DISCR	96.70	3.89	93.82	4.91
FALSEPOS	0.59	1.54	1.17	1.58
RESPBIAS	-0.05	0.24	-0.07	0.34
Additional Scores:				
Recency raw score	14.92	2.56	13.02	3.24

Note. T1-5 = Total words trials 1-5; A1 = List A trial 1; A5 = List A trial 5; B = List B; SDFR = Short delay free recall; SDCR = Short delay cued recall; LDFR = long delay free recall ; LDCR = long delay cued recall; SEM = Semantic clustering; SER = Serial clustering; %PRI = Percent primacy

region; %MID = Percent middle region; %REC = Percent recency region; RECALL CONS. = Recall consistency; PERS = Perseverations; FRINTR = Free recall intrusions; CRINTR = Cued recall intrusions; TOTINTR = Total intrusions; RECHITS = Recognition Hits; DISCR. = Discriminability; FALSEPOS = False positives; RESPBIAS = Response bias.

3. ***Present Ability Memory Measures and Demographics***

The study then examined the utility of present ability memory measures and demographic variables at predicting CVLTLDFR performance in the above normal elderly Canadian sample. Demographic data for the current sample (N = 97) included: age (M = 71.9, S.D. = 7.1, Range = 53 to 89 years); education (M = 13.6, S.D. = 3.1, Range = 6 to 22 years); gender distribution (20 males, 77 females). Mean, standard deviation and range values for all present ability memory measures are presented in Table 17.

Pearson (r) correlations of memory predictors (i.e., present ability memory measures and demographics) with CVLTLDFR performance are presented in Table 18. Examination of Table 18 identified four predictor variables (CVLT Recency raw score, WAIS-R Digits Forward raw score (DF), age, sex) that correlated significantly with CVLT long delay free recall performance ($p < .01$). Further analyses revealed that CVLTLDFR scores decreased with advancing age, with a significant difference in CVLTLDFR scores for persons aged

70 years+ as compared to persons less than 70 years of age ($t(95) = -2.809$, $p < .01$). A gender effect was also observed as males and females differed significantly on CVLTLDFR scores ($t(95) = -2.096$, p, $< .05$), with females outperforming males.

Table 15

Pearson (r) correlations of VIQ Equations with CVLTLDFR

Study	VIQ predictors in equation	CVLTLDFR
	Existing Equations	
Carswell et al. (1997)	NART + WAIS-R Vocabulary	0.095
Krull et al. (1995)	WAIS-R Vocabulary + Demographics	0.120
Blair & Spreen (1989)	NART-R	0.088
Carswell et al. (1997)	NART	0.073
Ryan & Paolo (1992)	NART	0.073
Nelson & Willison (1991)	NART	0.074
Barona et al. (1984)	Demographics	0.156

Barona et al. (1984)*	Demographics	0.137
	Newly Developed VIQ Equations	
Current study	WAIS-R Vocabulary + STW	0.118
Current study	NART-R + STW	0.108

Note. None of the VIQ equations produced significant correlations with CVLTLDFR scores at $p < .01$, or $p < .05$; CVLTLDFR = CVLT long delay free recall performance; Barona et al. (1984)* = Barona et al., (1984) equation with Helmes (1996) age modification.

Table 16

Pearson (r) correlations of VIQ Predictors with CVLTLDFR

IQ Predictor	CVLTLDFR
NART errors	-0.073
NART-R errors	-0.088
WAIS-R Vocabulary (age-scaled score)	0.100
STW	0.115
WRAT-3 Reading subtest (raw score)	0.053

Note. None of the VIQ predictors produced significant correlations with CVLTLDFR scores at $p < .01$, or .05.

Table 17

Mean, Standard Deviation and Range Values for all Present Ability Memory Measures

Variable	Mean	S.D.	Range Min.	Range Max.
DF (Raw sc.)	9.02	2.13	4.00	13.00
SR (Corr. sc.)	19.26	2.11	15.00	25.00
CVLT Rec. (Raw sc.)	13.74	3.13	5.00	20.00

Note. DF = WAIS-R Digits Forward; SR = NCCEA Sentence Repetition; CVLT Rec. = CVLT Recency score; Raw sc.= Raw score; Corr. sc.= Corrected score; S.D. = Standard Deviation; Min. = Minimum; Max. = Maximum.

Stepwise multiple regression ($p < .05$ for entry; a liberal p value was adopted to improve the sensitivity of the analysis) was employed to determine which of the four significant predictor variables identified in Table 18 (i.e., CVLT Recency raw score, WAIS-R Digits Forward raw score, age, sex) were most effective at estimating CVLTLDFR performance in the above normal elderly sample. Two variables including the CVLT Recency raw score and the WAIS-R Digits Forward raw score contributed significantly to the prediction of CVLTLDFR performance accounting for 32% of the

variance in CVLTLDFR scores. The regression equation based on these results was:

estimated CVLTLDFR = 2.114 + 0.468 (CVLT Recency raw score)
+0.284 (WAIS-R Digits Forward raw score)

The multiple correlation for the above equation was 0.567. A scatterplot of CVLTLDFR predicted values to actual/obtained CVLTLDFR scores is presented in Figure 5 for the new predictive memory equation.

Table 18

Pearson (r) correlations of Memory Predictors with CVLTLDFR

Variable	CVLTLDFR
Present Ability Memory Measures	
WAIS-R Digits Forward raw score	0.234*
Sentence Repetition corrected score	0.072
CVLT Recency raw score	0.526**
Demographics	
Age	-0.302**
Sex	-0.210*
Education	0.149

* p <.05, ** p <.01

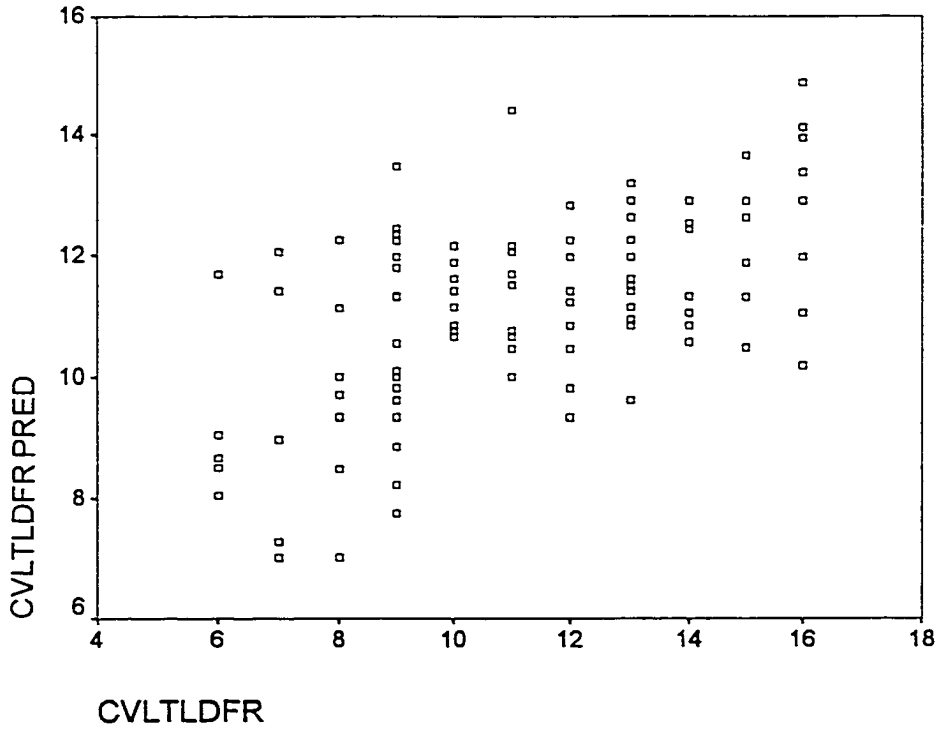

Figure 5. Scatterplot of predicted scores to CVLTLDFR scores

The sensitivity of the new CVLT long delay free recall equation for detecting cognitive (verbal memory/delayed recall) decline was examined in the context of discrepancy values between obtained and predicted CVLTLDFR scores. Figure 6 shows the frequency distribution of discrepancy values between obtained and predicted CVLT long delay free recall scores.

The frequency distribution of CVLTLDFR discrepancy scores was a normal distribution (skewness = -0.064, standard error = 0.245, (ratio = -0.26); kurtosis = -0.288, standard error = 0.485, (ratio = +0.59,)), which allowed for the application of the SEe method as well as the percentile

table method in the calculation of sensitivity scores.

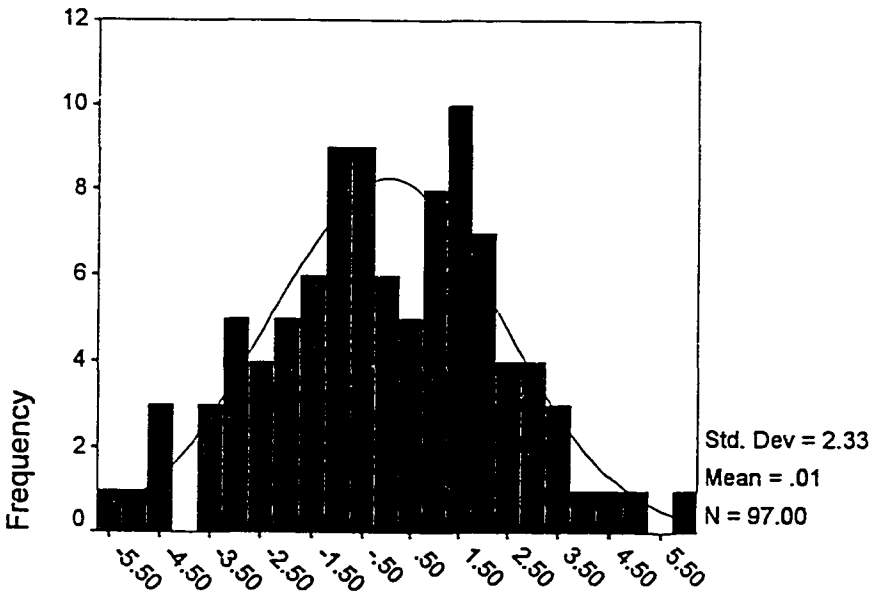

Figure 6. Frequency distribution of discrepancy values between obtained and predicted CVLTLDFR scores

The percentile table for the discrepancy values between obtained and predicted CVLT long delay free recall scores is presented in Table 19. The cognitive decline sensitivity score based on the percentile table was considered to be 6 points, with a cut-off discrepancy score of 4 points and a cut-off sensitivity score of 2 points.

Cross-validation calculations were performed for the new predictive regression equation to determine the accuracy of the equation for predicting CVLT long delay free recall performance for a participant not in the original sample. Correction of the multiple correlation (R = 0.567) with 4

predictors considered in the regression analysis produced an estimated cross-validation multiple correlation of 0.519 (R^2 = 0.270) based on Cattin's (1978, 1980) approach. Calculation of the cross-validation SEe (n.b., SEe = 2.3597 as reported by SPSS 7.5) utilizing Darlington's (1968) approach produced an estimated cross-validation SEe of 2.454 points. Cross-validation sensitivity scores based on the estimated cross-validation SEe included a cut-off discrepancy score of 4.04 points, and a cognitive decline sensitivity score of 6.10 points.

Table 19

Discrepancies between Obtained and Predicted CVLTLDFR Scores

Difference (CVLT points)	Est.CVLT Higher than Obt.CVLT		Est.CVLT Lower than Obt.CVLT	
	(N)	(%)	(N)	(%)
1	28	28.9	29	29.9
2	9	9.3	10	10.3
3	6	6.2	6	6.2
4	3	3.1	3	3.1
5	2	2.0	1	1.0

Note. Est. = Estimated; Obt. = Obtained.

Table 20

Efficacy of Memory Measures at Predicting CVLTLDFR Scores

Variable	Cross-validation SEe (CVLTPT)	COD (CVLTPT)	CDS (CVLTPT)
Cross-validation Standard of Comparison			
CVLTLDFR (1.3 yr previous) Paolo et al. (1997a)	1.86	3.07	4.64
CVLT Rec. + DF	2.45	4.04	6.10
CVLT Rec.	2.45	4.04	6.10
Age	2.75	4.52	6.84
DF	2.81	4.61	6.98
Sex	2.82	4.64	7.02

Note. SEe = Standard Error of Estimate; COD = Cut-off Discrepancy score; CDS = Cognitive Decline Sensitivity score; CVLTPT = CVLT Points; yr = year; CVLT Rec. = CVLT Recency raw score; DF = WAIS-R Digits Forward raw score.

Cut-off discrepancy scores and cognitive decline sensitivity scores (SEe method) were calculated for memory predictors that correlated significantly with CVLT long

delay free recall performance to determine their independent clinical utility. The efficacy of significant memory measures at predicting CVLT long delay free recall performance and detecting cognitive decline is presented in Table 20. Included in Table 20 are cross-validation cut-off discrepancy and cognitive decline sensitivity scores based on test-retest reliability data of the CVLT with healthy elderly participants aged 54 to 87 years over a 1.3-year period ($r = 0.76$; S.D. = 2.84; N = 151; Paolo, Tröster, & Ryan, 1997a). These scores were included in Table 20 for comparison with the sensitivity of other methods of detecting decline in the domain of verbal memory. CVLTLDFR regression equations for independent predictors are presented in Appendix C.

Language Predicted by VIQ Measures, Present Ability Measures and Demographic Variables

The third part of the study examined the concurrent validity of VIQ measures (i.e., existing VIQ equations, newly developed VIQ equations and independent VIQ predictors), measures of present ability (i.e., WAIS-R Vocabulary subtest: raw score; Peabody Picture Vocabulary Test-Revised (PPVT-R): raw score; Word-Picture Matching subtest of the BDAE (WPM): raw score, latency (seconds) to complete the task) and demographic variables (i.e., age, sex, education) at estimating confrontational naming

performance on the Boston Naming Test (BNT: total number of correct spontaneous responses and stimulus cue responses). The WAIS-R Vocabulary subtest was investigated as a present ability measure in this part of the study, and therefore was not included as an VIQ predictor in the estimation of BNT performance. The sample size for the following analyses included 98 participants.

1. ***BNT Normative Data***

Statistics for the criterion measure BNT, for the present sample (N = 98) were: M = 56.33; S.D. = 3.39; Range = 42.00 to 60.00. Normative data (percentiles) for BNT scores by education is presented in Table 21. Normative data for BNT scores is presented in terms of percentiles as recommended by Tombaugh & Hubley (1997).

2. ***VIQ Equations***

The efficacy of existing VIQ equations and newly developed VIQ equations at predicting BNT performance was first examined by calculating the correlations between VIQ prediction equations and BNT scores. As shown in Table 22, all of the VIQ prediction equations correlated significantly with BNT scores at $p < .01$. WAIS-R VIQs also correlated significantly with BNT scores at $p < .01$ ($r = 0.532$). Cross-validation cut-off discrepancy scores and cognitive decline sensitivity scores based on the SEe method were calculated

for VIQ equations that accounted for more than 10% of the variance in BNT scores (r >.33), as illustrated in Table 23. Included in Table 23 were cross-validation SEe, cut-off discrepancy and cognitive decline sensitivity scores based on test-retest reliability data of the BNT. Test-retest data was collected in a sample of 122 healthy elderly adults aged 57 to 85 years over a two-year period (r = 0.89, S.D. = 5.1; Mitrushina & Satz, 1991). These sensitivity scores were included in Table 23 for comparison with the sensitivity of VIQ equations for detecting a decline in BNT performance.

Table 21

BNT Norms Expressed as Percentiles for Education

	Education	
%iles	<13 years (N = 44)	≥13 years (N = 54)
90	59.00	60.00
75	58.00	59.00
50	55.00	58.00
25	53.00	55.00
10	49.00	53.00

Note. %iles = percentiles; N = number of individuals at each education level.

Table 22

Pearson (r) correlations of VIQ Equations with BNT

Study	VIQ predictors in equation	BNT
	Existing VIQ Equations	
Krull et al. (1995)	WAIS-R Vocab. + Demographics	0.661
Carswell et al. (1997)	WAIS-R Vocab. + NART	0.564
Blair & Spreen (1989)	NART-R	0.471
Carswell et al. (1997)	NART	0.439
Ryan & Paolo (1992)	NART	0.439
Nelson & Willison (1991)	NART	0.435
Barona et al. (1984)	Demographics	0.269
Barona et al. (1984)*	Demographics	0.260
	Newly Developed VIQ Equations	
Current study	WAIS-R Vocab. + STW	0.664
Current study	NART-R + STW	0.577

Note. All correlations significant at $p < .01$.; Barona et al. (1984)* = Barona equation with the Helmes (1996) age modification.

Table 23

Efficacy of VIQ Equations at Predicting BNT scores

Study	VIQ Predictors in equation	Cross-validation SEe (BNTPT)	Cross-validation COD (BNTPT)	Cross-validation CDS (BNTPT)
Cross-validation Standard of Comparison				
Mitrushina & Satz (1991)	BNT (2 yr previous)	2.35	3.87	5.86
Existing VIQ Equations				
Krull et al. (1995)	WAIS-R Vocab. + Demographics	2.81	4.63	6.99
Carswell et al. (1997)	WAIS-R Vocab. + NART	2.87	4.73	7.15
Blair & Spreen (1989)	NART-R	3.04	4.99	7.55
Carswell et al. (1997)	NART	3.09	5.09	7.69
Ryan & Paolo (1992)	NART	3.09	5.09	7.69
Nelson & Willison (1991)	NART	3.10	5.10	7.71

Newly Developed VIQ Equations

Current study	WAIS-R Vocab.+ STW	2.60	4.28	6.47
Current study	NART-R + STW	2.84	4.68	7.07

Note. SEe = Standard error of estimate; COD = Cut-off Discrepancy score; CDS = Cognitive Decline Sensitivity score; BNTPT = BNT Points; yr = years; WAIS-R Vocab.= WAIS-R Vocabulary subtest.

3. ***VIQ Predictors, Present Ability Language Measures and Demographics***

The study then examined the utility of independent VIQ predictors (i.e., NART errors, NART-R errors, STW, WRAT-3 Reading subtest raw score), present ability language measures and demographic variables at predicting BNT performance in the above normal elderly Canadian sample. Descriptive statistics for the above VIQ predictors have been presented in Table 6, and demographic data for the sample (N = 98) has been summarized in the Methods section. Mean, standard deviation and range values for all present ability language measures are presented in Table 24.

Pearson (r) correlations of language predictors (i.e., VIQ predictors, present ability measures, and demographics) with BNT performance are presented in Table 25. Examination of Table 25 identified 7 predictor variables (NART errors, NART-R errors, STW, WRAT-3 Reading subtest raw score, WAIS-R

Vocabulary subtest raw score, PPVT-R raw score and education) that correlated significantly with BNT scores ($p < .01$). Further analyses showed that BNT scores increased with advancing education with a significant difference in BNT scores for persons with 13 years+ education as compared to persons with less than 13 years of education ($t(96) = 3.01$, $p < .001$).

Table 24

Mean, Standard Deviation and Range Values for all Present Ability Language Measures

Variable	Mean	S.D.	Range Min.	Range Max.
Vocab. (Raw sc.)	55.10	7.40	31.00	67.00
PPVT-R (Raw sc.)	168.60	7.04	144.00	175.00
WPM (Raw sc.)	9.68	0.51	8.00	10.00
WPM-L (seconds)	31.02	8.97	19.03	85.51

Note. S.D. = Standard Deviation; Min. = Minimum; Max. = Maximum; Vocab. = WAIS-R Vocabulary; PPVT-R = Peabody Picture Vocabulary Test-Revised; WPM = BDAE Word-Picture Matching test; WPM-L = WPM - Latency; Raw sc. = Raw score.

Table 25

Pearson (r) Correlations of Language Predictors with BNT

Variable	BNT
VIQ Predictors	
NART errors	-0.439*
NART-R errors	-0.471*
STW	0.613*
WRAT-3 Reading subtest (raw score)	0.327*
Present Ability Measures	
WAIS-R Vocabulary (raw score)	0.635*
PPVT-R (raw score)	0.616*
WPM (raw score)	-0.029
WPM-L (seconds)	-0.135
Demographics	
Age	-0.149
Sex	0.119
Education (point biserial)	0.294*

* $p < .01$; Education point biserial = +1 if education $\geq$ 13 years and, -1 if education < 13 years.

Stepwise multiple regression ($p < .01$ for entry) was employed to determine which of six significant predictor variables identified in Table 25 (i.e., NART-R errors, STW, WRAT-3 Reading, WAIS-R Vocabulary, PPVT-R, and education) were most effective at estimating BNT scores in the above normal elderly Canadian sample. NART errors were excluded from the regression analysis given the preferred use of the NART-R in a North American sample, and the higher correlation of NART-R errors with BNT scores in this sample. Three variables including WAIS-R Vocabulary raw score, STW, and WRAT-3 Reading subtest raw score contributed significantly to the prediction of BNT performance, accounting for 54% of the variance in naming scores. The regression equation based on these results was:

estimated BNT = 40.901 + 0.233 (WAIS-R Vocabulary raw score) + 0.375 (STW raw score) - 0.341 (WRAT-3 Reading raw score)

The multiple correlation for the above equation was 0.731. A scatterplot of the predicted BNT scores to actual/obtained BNT scores is presented in Figure 7.

The sensitivity of the new BNT predictive equation for detecting cognitive (language/naming) decline was examined in the context of discrepancy values between obtained and predicted BNT scores. The frequency distribution of discrepancy values between obtained and predicted BNT scores is shown in Figure 8.

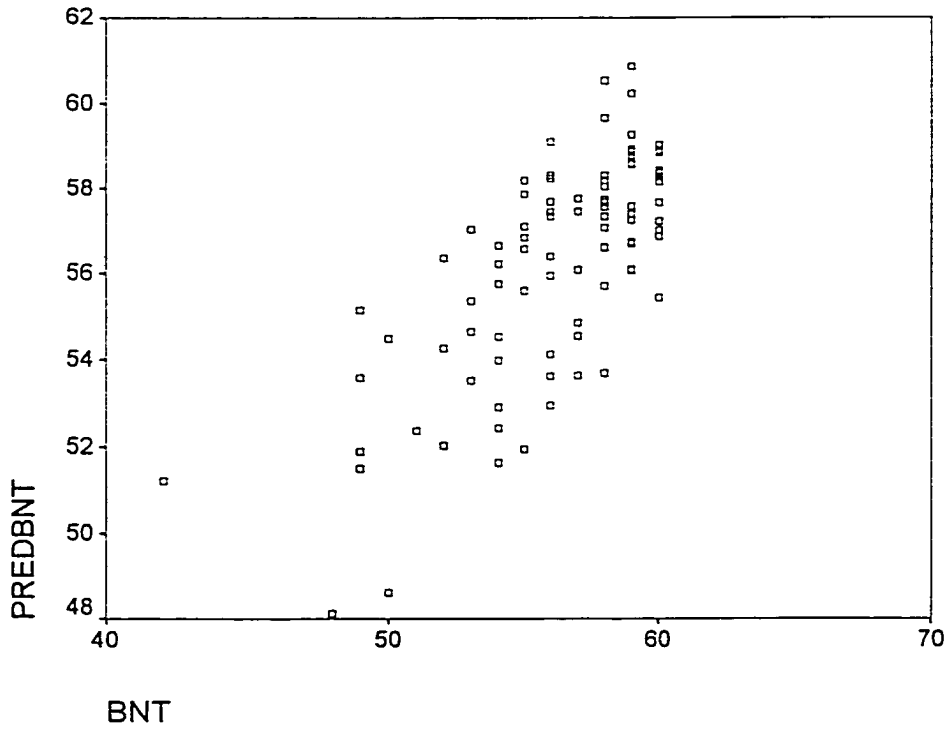

Figure 7. Scatterplot of predicted scores (WAIS-R Vocabulary, STW, WRAT-3 Reading based) to BNT scores

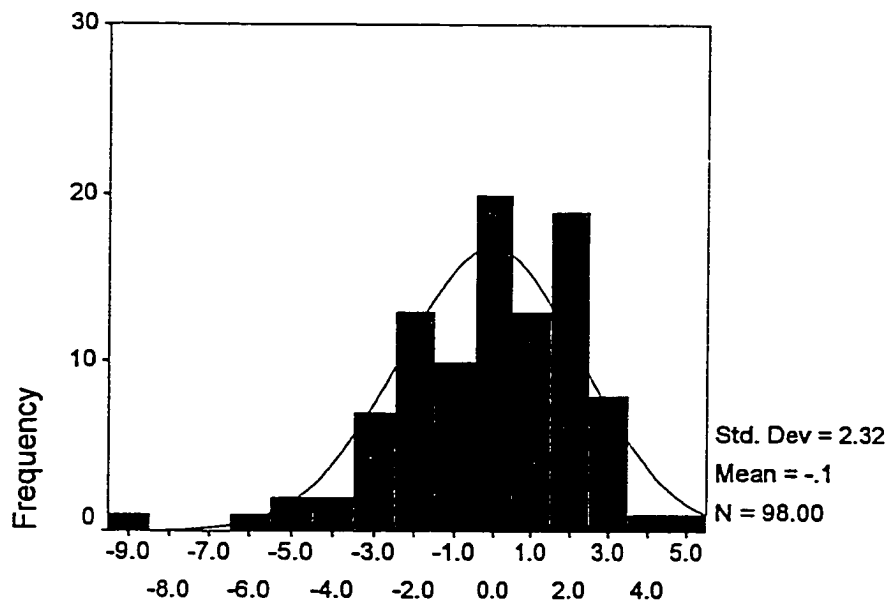

Figure 8. Frequency distribution of discrepancy values between obtained and predicted BNT scores (WAIS-R Vocabulary, STW, WRAT-3 Reading based)

The frequency distribution of BNT discrepancy scores was a nonnormal distribution (skewness = -0.844, standard error = 0.244, (ratio = -3.46); kurtosis = 1.632, standard error = 0.483 (ratio = +3.37)). However, sensitivity scores were calculated, for comparison, utilizing both the SEe method and the percentile table approach.

The percentile table for the discrepancy values between obtained and predicted BNT scores is presented in Table 26. The cognitive decline sensitivity score, based on the table, was considered to be 6 points, with a cut-off discrepancy score of 4 points and a cut-off sensitivity score of 2 points.

Cross-validation calculations were performed for the new predictive regression equation to determine the accuracy of the equation for predicting BNT performance for a participant not in the original sample. Correction of the multiple correlation (R = 0,731) with 6 predictor variables considered in the regression analysis produced an estimated cross-validation multiple correlation of 0.691 (R^2 = 0.477) based on Cattin's (1978, 1980) approach. Calculation of the cross-validation SEe (n.b., SEe = 2.3533 as reported by SPSS 7.5) utilizing Darlington's (1968) approach produced an estimated cross-validation SEe of 2.479 points. Cross-validation sensitivity scores based on the estimated cross-validation SEe were 4.08 points (cut-off discrepancy score) and 6.17 points (cognitive decline sensitivity score).

Table 26

Percentile Table of Discrepancy Values between Obtained and Predicted BNT scores based on WAIS-R Vocabulary, STW, and WRAT-3 Reading scores

Difference (BNT points)	Est.BNT Higher than Obt.BNT		Est.BNT Lower than Obt.BNT	
	(N)	(%)	(N)	(%)
1	26	26.5	35	35.7
2	11	11.2	12	12.2
3	2	2.0	4	4.1
4	4	4.1	2	2.0
5	-	-	-	-
6	1	1.0	-	-
7	-	-	-	-
8	-	-	-	-
9	1	1.0	-	-

Note. Est.= Estimated; Obt.= Obtained.

As there has been some concern in the literature regarding the ability of the WAIS-R Vocabulary subtest to resist decline in early Alzheimer's disease, a new predictive equation was developed with WAIS-R Vocabulary excluded from the regression analysis. Stepwise multiple

regression ($p < .01$ for entry) was employed to determine which of the predictor variables including NART-R errors, STW, WRAT-3 Reading subtest raw score, PPVT-R raw score, and education, were most effective at estimating BNT scores in the above normal elderly sample. Three variables including the PPVT-R raw score, STW, and WRAT-3 Reading subtest raw score contributed significantly to the prediction of BNT performance, accounting for 49% of the variance in naming scores. The regression equation based on these results was:

estimated BNT = 15.484 + 0.236 (PPVT-R raw score) +
0.343 (STW raw score) - 0.338 (WRAT-3 Reading raw score)

The multiple correlation for the above equation was 0.698. A scatterplot of predicted BNT scores (PPVT-R, STW, WRAT-3 Reading based) to actual/obtained BNT scores is presented in Figure 9.

The sensitivity of the new BNT predictive equation for detecting language decline was examined in the context of discrepancy values between obtained and predicted BNT scores. The frequency distribution of discrepancy values between obtained and predicted BNT scores is shown in Figure 10.

The frequency distribution of BNT discrepancy scores was a nonnormal distribution (skewness = -0.807, standard error = 0.244, (ratio = -3.31); kurtosis = 0.791, standard error = 0.483, (ratio = +1.64)). However, sensitivity scores

were calculated for comparison, utilizing both the SEe method and percentile table approach.

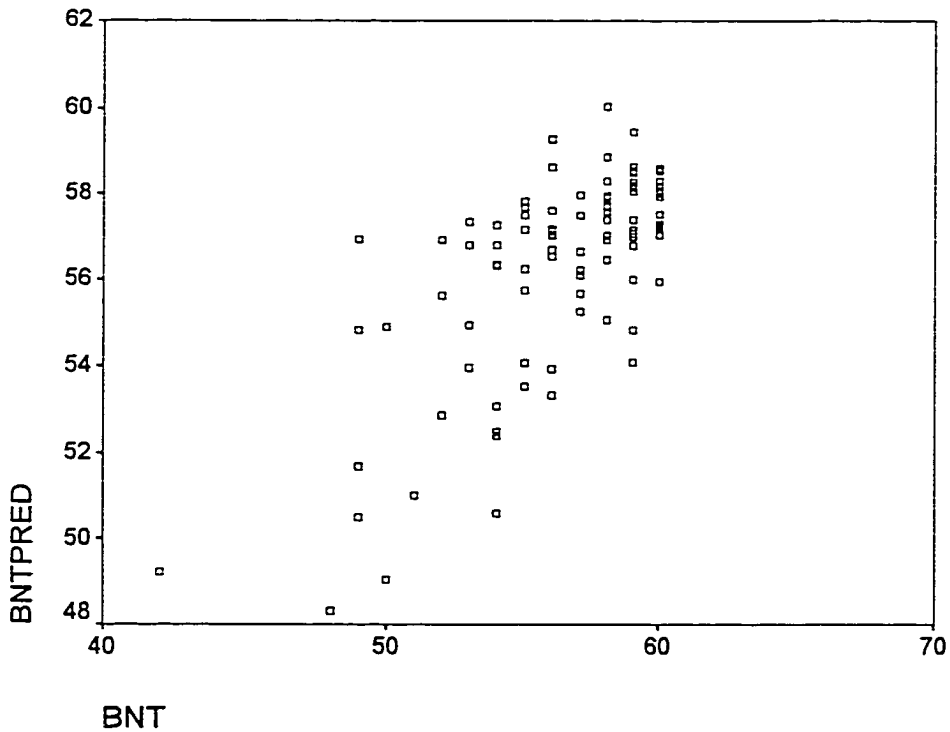

Figure 9. Scatterplot of predicted scores (PPVT-R, STW, WRAT-3 Reading based) to BNT scores

The percentile table for the discrepancy values between obtained and predicted BNT scores is presented in Table 27. The cognitive decline sensitivity score based on the percentile table was considered to be 6 points, with a cut-off discrepancy score of 4 points, and a cut-off sensitivity score of 2 points, respectively.

Cross-validation calculations were performed for the new predictive regression equation to determine the accuracy of the equation for predicting BNT performance for a

participant not in the original sample. Correction of the observed multiple correlation (R = 0.698) with 5 available predictor variables considered in the regression analysis produced an estimated cross-validation multiple correlation of 0.661 (R^2 = 0.436) based on Cattin's (1978, 1980) approach. Calculation of the cross-validation SEe (n.b., SEe = 2.4724 as reported by SPSS 7.5) utilizing Darlington's (1968) approach produced an estimated cross-validation SEe of 2.573 points. Cross-validation sensitivity scores based on the estimated cross-validation SEe were 4.23 points (cut-off discrepancy score) and 6.40 points (cognitive decline sensitivity score).

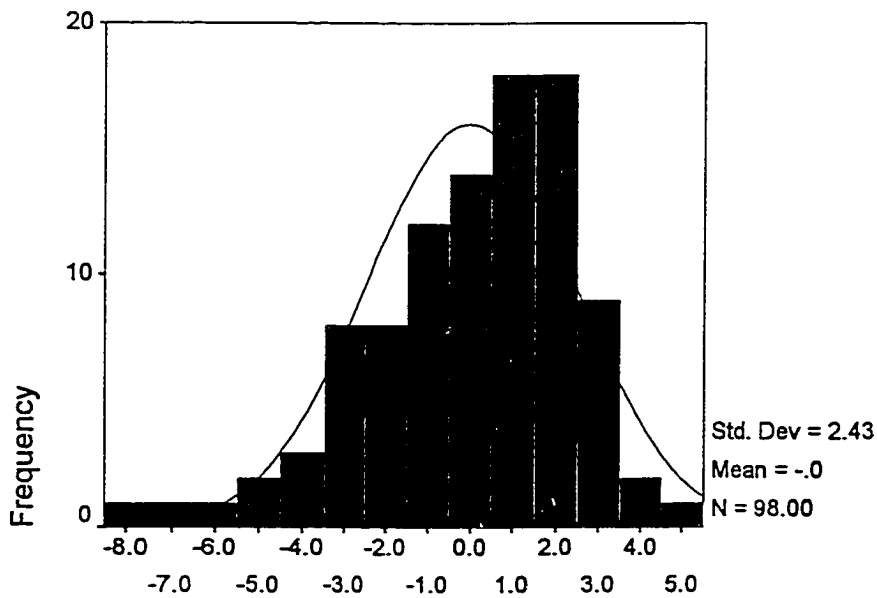

Figure 10. Frequency distribution of discrepancy values between obtained and predicted BNT scores (PPVT-R, STW, WRAT-3 Reading subtest based)

Table 27

Percentile Table of Discrepancy Values between Obtained and Predicted BNT scores based on PPVT-R, STW, WRAT-3 Reading scores

Difference (BNT Points)	Est. BNT Higher than Obt. BNT (N)	(%)	Est. BNT Lower than Obt. BNT (N)	(%)
1	21	21.4	39	39.8
2	10	10.2	12	12.2
3	5	5.1	2	2.0
4	3	3.0	3	3.0
5	1	1.0	-	-
6	-	-	-	-
7	2	2.0	-	-

Note. Est = Estimated; Obt. = Obtained.

The efficacy of significant language measures at independently estimating BNT performance and detecting cognitive decline (SEe method) is presented in Table 28. Included in Table 28 are cross-validation cut-off discrepancy and cognitive decline sensitivity scores based on test-retest reliability data of the BNT in a healthy elderly sample. BNT regression equations for independent

language predictors that demonstrated clinical utility (i.e., cross-validation cognitive decline sensitivity score less than 7 points) are included in Appendix D.

Table 28

Efficacy of Language Measures at Predicting BNT Scores

Variable	SEe	Cross-validation COD (BNTPT)	Cross-validation CDS (BNTPT)
Cross-validation Standard of Comparison			
BNT (2 yr previous) Mitrushina & Satz (1991)	2.35	3.87	5.86
Vocab. + STW + WRAT-3 Rd.	2.48	4.08	6.17
PPVT-R + STW + WRAT-3 Rd.	2.57	4.23	6.40
Vocab.	2.66	4.38	6.62
PPVT-R	2.71	4.46	6.75
STW	2.72	4.48	6.77
NART-R errors	3.04	4.99	7.55
NART errors	3.09	5.09	7.69
WRAT-3 Rd.	3.25	5.35	8.09
Education (point biserial)	3.29	5.41	8.18

Note. SEe = Standard error of estimate; COD = Cut-off

discrepancy score; CDS = Cognitive decline sensitivity score; BNTPT = BNT Points; yr = years; Vocab. = WAIS-R Vocabulary subtest raw score; WRAT-3 Rd. = WRAT-3 Reading subtest raw score.

Discussion

VIQ Predicted by VIQ Measures

The first part of the study examined the concurrent validity of several recently proposed measures of premorbid VIQ (i.e., VIQ equations and independent VIQ predictors) for estimating WAIS-R VIQ, the criterion measure, in a normal elderly Canadian sample. WAIS-R VIQs were initially calculated for persons 75 years and older by normative data provided by Ryan et al. (1990), because the WAIS-R lacks normative data for individuals in this age range. However, WAIS-R VIQs based on the Ryan et al. (1990) normative data correlated significantly with age and were not considered to provide representative estimates of WAIS-R VIQ for this sample, because they failed to adequately correct for age. Since WAIS-R VIQs based on Wechsler (1981) normative data did not correlate significantly with age (70-74 year old norms applied to VIQ scores for persons 75+ years), Wechsler normed (1981) WAIS-R VIQ was selected as the criterion measure for the above normal elderly Canadian sample.

It is suggested that the Ryan et al. (1990) normative

data failed to provide a representative estimate of WAIS-R VIQ in the above normal elderly Canadian sample due to a mismatch of demographic data between the above sample and the Ryan et al. (1990) normative group. The Ryan et al. (1990) normative sample consisted of 130 elderly Americans ranging in age from 75 to 96 years, and the majority of the sample (69.2%) had less than 12 years of education. In the above sample of normal elderly Canadians, 30 participants were 75 years or older, and of this group only 8 persons (26%) had less than 12 years of education. Therefore, it is suggested that caution be used when applying normative data to WAIS-R VIQs for persons 75 years and older that may differ demographically from the normative group. However, it should be noted that only two studies currently exist that provide WAIS-R normative data for persons 75 years and older, both with elderly American participants (Ryan et al. 1990; Ivnik et al. 1992). To date, there are no studies that have provided normative data for Canadians aged 75 years and older on the WAIS-R.

Selection of Wechsler normed (1981) WAIS-R VIQ as the criterion measure allowed for examination of the concurrent validity of several measures of premorbid VIQ (i.e., VIQ equations and independent VIQ predictors) at estimating this criterion in a normal elderly Canadian sample. This comparative examination provided an opportunity to: 1) obtain normative data as to the performance of elderly

Canadians on several premorbid VIQ measures, 2) cross-validate some of the more recently proposed measures of premorbid VIQ with a healthy elderly sample, 3) compare the utility of various measures of premorbid VIQ at estimating WAIS-R VIQ in the same normal elderly sample, and 4) identify/develop the most useful predictive measure/equation for estimating WAIS-R VIQ for elderly Canadians, and detecting cognitive decline in clinical samples.

1. *VIQ Equations*

(a) Accuracy of VIQ Equations at Predicting Actual WAIS-R VIQ Scores

The accuracy of existing VIQ equations at predicting WAIS-R VIQs in the current normal elderly Canadian sample was determined by examining the range of predicted WAIS-R VIQ scores and the mean error in prediction of actual WAIS-R VIQs (i.e., actual minus predicted VIQ scores). Review of Table 2 revealed that the Carswell et al. (1997) WAIS-R Vocabulary/NART equation (range: 89.9 to 129.3 VIQ points) and the Blair and Spreen (1989) NART-R equation (range: 87.8 to 126.9 VIQ points) produced a less restricted range of predicted WAIS-R VIQs than other estimation methods. It should be noted that in the current study, the Krull et al. (1995) WAIS-R Vocabulary/Demographics equation failed to demonstrate the breadth of predictive range previously

reported by Scott et al. (1997). Scott et al. (1997) compared the range of predicted WAIS-R FSIQs generated by various estimation methods and reported that the Krull et al. (1995) equation produced a larger range in predicted WAIS-R FSIQs (range: 73.3 to 125.3 FSIQ points) than other estimation methods without systematic over or underestimation of WAIS-R FSIQs. Although the current study examined the range of predicted WAIS-R VIQs as compared to FSIQs, the Krull et al. (1995) equation in fact produced a more restricted range of predicted WAIS-R VIQ scores (range: 90.4 to 121.2 VIQ points) than most other estimation methods. This finding may have been partly due to the fact that the Krull et al. (1995) equations were originally developed based on the WAIS-R standardization sample with an age range of 16 to 74 years, whereas the current sample was exclusively elderly, with participants ranging in age from 53 to 89 years.

Examination of the mean error in prediction of actual WAIS-R VIQ scores (i.e., actual minus predicted VIQ scores) for each VIQ prediction equation, as illustrated in Table 3, revealed that the Nelson and Willison (1991) NART equation produced the most accurate estimation of actual WAIS-R VIQs with a mean error of prediction of +0.70 VIQ points. This suggested that on average, the Nelson and Willison (1991) NART equation underestimated WAIS-R VIQs for the entire sample by less than one VIQ point. Other predictive VIQ

equations which also offered good estimates of actual VIQ scores with error scores of less than two VIQ points on average included the Blair and Spreen (1989) NART-R equation; the Carswell et al. (1997) WAIS-R Vocabulary/NART equation; and the Krull et al. (1995) WAIS-R Vocabulary/Demographics equation. VIQ equations with more limited accuracy at predicting actual WAIS-R VIQs on average included the Barona et al. (1984) Demographics equation (both with and without the Helmes (1996) age modification); the Ryan and Paolo (1992) NART equation; and the Carswell et al. (1997) NART equation.

Mean error scores were calculated by IQ range classification for each VIQ equation as illustrated in Table 4, because several studies have commented on the limitation of predictive IQ measures at estimating IQ scores outside of the median range. Examination of Table 4 revealed that all VIQ equations significantly overestimated actual WAIS-R VIQ scores of 80-89 VIQ points, and most VIQ equations underestimated actual WAIS-R VIQ scores of 130+ VIQ points. This pattern of prediction was consistent with other studies that have shown that predictive IQ measures are restricted in accurately estimating IQ scores outside of the median range. However, it should be noted that three equations including the Carswell et al. (1997) WAIS-R Vocabulary/NART equation; the Carswell et al. (1997) NART equation; and the Ryan and Paolo (1992) NART equation demonstrated greater

accuracy at predicting high extreme WAIS-R VIQs (i.e., 130+ VIQ points) than other estimation methods. These equations were developed with exclusively elderly, relatively high functioning individuals, consistent with the current normal elderly sample, which likely accounted for more accurate prediction in high extreme VIQ scores than other estimation methods.

It should also be noted that the current study was the first study to apply the Helmes (1996) age modification in the calculation of Barona et al. (1984) estimated WAIS-R VIQs for persons over 74 years of age. The results of the current study suggested that the Helmes (1996) age modification failed to significantly improve the predictive accuracy of the Barona et al. (1984) equation for estimating actual WAIS-R VIQs in the above normal elderly Canadian sample.

Summary

In summary, the current study reviewed the accuracy of existing VIQ equations at predicting actual WAIS-R VIQs in the above normal elderly sample. Given the above results, the Carswell et al. (1997) WAIS-R Vocabulary/NART equation and the Blair and Spreen (1989) NART-R equation were judged to provide the best overall estimates of WAIS-R VIQ in terms of accuracy of prediction for the entire sample and range of predicted WAIS-R VIQ scores. The Carswell et al. (1997)

WAIS-R Vocabulary/NART equation also demonstrated accuracy in predicting WAIS-R VIQ scores across most of the IQ range (i.e., 90-130+ VIQ points), and was only limited at estimating WAIS-R VIQs below 90 VIQ points in the current sample.

(b) Clinical Utility of VIQ Equations at Detecting Cognitive Decline

The approach used most frequently in the literature for determining the utility of a premorbid IQ measure has been the amount of variance accounted for by the predictive IQ measure in actual WAIS-R IQ scores. However, more recently, it was suggested that the true measure of the utility of a premorbid IQ measure is the sensitivity of the measure for detecting cognitive decline in a clinical sample, given the error of prediction of the measure at estimating the criterion in a healthy sample that has not suffered cognitive decline (Graves et al. in press; Carswell et al. 1997). Graves et al. (in press) developed a cognitive decline sensitivity score for predictive measures which was defined as the score reflecting the minimum cognitive decline that could be successfully detected 80% of the time (80% sensitivity) while using a cut-off discrepancy that would correctly classify 95% of normals (95% specificity). Graves et al. (in press) outlined two approaches for calculating cognitive decline sensitivity scores for

predictive measures including: 1) the standard error of estimate (SEe) method (dependent on a normal distribution of discrepancy values between obtained and predicted IQ scores), and 2) the percentile table method (not dependent on a normal distribution of discrepancy scores). An advantage of the SEe method was that cross-validation calculations were possible for this approximation method based on a computer program developed by Graves (1998). This computer program produced cognitive decline sensitivity scores corrected for the shrinkage in predictive accuracy expected on cross-validation, based on the estimated cross-validation SEe.

Graves et al. (in press) further indicated that the clinical utility of a potential premorbid IQ measure could be evaluated by comparing the cross-validation cognitive decline sensitivity score of the predictive measure to the cross-validation cognitive decline score calculated for the ideal situation in which a previous IQ score was actually available. Studies reporting test-retest reliability data for the WAIS-R offer such information. However, only one study has provided test-retest reliability data for healthy elderly participants on the WAIS-R. Snow et al. (1989) administered the WAIS-R to healthy elderly Canadians aged 50-84 years, one year following the initial administration of the measure. The test-retest reliability of WAIS-R VIQs for this elderly sample over a one-year period was 0.86.

Calculation of the cross-validation cognitive decline sensitivity score produced a value of 12.8 VIQ points for this sample in which previous VIQ scores were actually available. Graves et al. (in press) considered that if a predictive VIQ measure could reliably detect a decline of about 13 VIQ points, then it would be very useful in a clinical situation. Graves et al. (in press) suggested that a cross-validation cognitive decline score of 13 VIQ points, based on the best case situation in which previous VIQ scores were available, could be applied as a *standard of comparison* for predictive measures that use an estimated previous VIQ score.

Consequently, the efficacy of existing VIQ equations at predicting WAIS-R VIQs in a healthy elderly Canadian sample, and detecting cognitive decline in clinical samples was determined by comparison of their respective cross-validation cognitive decline sensitivity scores with the 13 VIQ point cross-validation test-retest standard of comparison proposed by Graves et al. (in press). Review of Table 5 revealed that the Carswell et al. (1997) WAIS-R Vocabulary/NART equation and the Krull et al. (1995) WAIS-R Vocabulary/Demographics equation compared favourably with the test-retest standard of comparison with cross-validation cognitive decline sensitivity scores of 15.2 and 15.8 VIQ points, respectively. VIQ equations which demonstrated more limited clinical utility with cognitive decline sensitivity

scores of approximately 17 to 23 VIQ points included: the Blair and Spreen (1989) NART-R equation; the Carswell et al. (1997) NART equation; the Ryan and Paolo (1992) NART equation; the Nelson and Willison (1991) NART equation; and the Barona (1984) Demographics equation.

(c) Cross-validation of VIQ Equations

An advantage of the current comparative investigation was that it provided useful cross-validation information regarding the predictive accuracy and clinical utility of several VIQ equations at estimating WAIS-R VIQs and detecting cognitive decline in a new sample (i.e., normal elderly Canadian sample). For many of these predictive VIQ equations including: the Blair and Spreen (1989) NART-R equation; the Carswell et al. (1997) WAIS-R Vocabulary/NART equation; the Carswell et al. (1997) NART equation; the Krull et al. (1995) WAIS-R Vocabulary/Demographics equation; the Nelson and Willison (1991) NART equation; and the Barona et al. (1984) Demographics equation, this was the first cross-validation with a normal elderly Canadian sample. In Table 5, the square root of the mean square error and cognitive decline sensitivity scores observed in the current sample were compared to the standard error of estimate and cognitive decline sensitivity scores (corrected for shrinkage expected on cross-validation) for the original study standardization samples.

Review of Table 5 revealed that five of the seven VIQ equations yielded either somewhat poorer results on cross-validation (i.e., Blair & Spreen (1989) NART-R equation; Ryan & Paolo (1992) NART equation; Carswell et al. (1997) NART equation), as expected (Graves et al. in press), or produced essentially the same cognitive decline sensitivity scores on cross-validation with the current normal elderly Canadian sample as compared to original sensitivity scores (corrected for shrinkage expected on cross-validation) based on their respective standardization samples (i.e., Krull et al. (1995) WAIS-R Vocabulary/Demographics equation with data referenced in Scott et al. 1997; Nelson & Willison (1991) NART equation). It should be noted that although the cognitive decline sensitivity score for the Blair and Spreen (1989) NART-R equation observed in the current sample (17.5 VIQ points) was poorer than the corrected/cross-validation sensitivity score of the original standardization sample (16.6 VIQ points), the observed cross-validation result in the current normal elderly Canadian sample was much improved as compared to previous cross-validations with American participants. Wiens et al. (1993) examined the ability of the Blair and Spreen (1989) NART-R equation at estimating WAIS-R IQs administered concurrently in a healthy young American sample with a WAIS-R VIQ cross-validation sensitivity score calculated at 19.9 VIQ points, while Berry et al. (1994) examined the accuracy of the NART-R equation

at postdicting WAIS-R IQs obtained 3.5 years earlier in a normal elderly American sample with a WAIS-R VIQ cross-validation sensitivity score calculated at 22.5 VIQ points. However, direct comparison of the cross-validation results of these studies with the original NART-R equation study (Blair & Spreen, 1989) was limited, given that the NART-R standardization study involved the concurrent administration of the NART-R and WAIS-R in a largely Canadian sample (49 Canadians, 17 Americans). Both of the above cross-validation studies consisted of exclusively American participants. Furthermore, the Berry et al. (1994) study was complicated by the fact that it examined the accuracy of the NART-R equation at postdicting WAIS-R IQs obtained 3.5 years prior to administration of the NART-R. While postdiction may provide a more accurate indication of the clinical utility of a predictive measure given that concurrent testing may maximize the correlation between predictive and criterion measures that test similar abilities (Carswell et al. 1997), use of concurrent cross-validation may provide a more conventional comparison. Thus, the current cross-validation with the above normal elderly Canadian sample was considered to be the most representative cross-validation of the NART-R equation because it involved an exclusively Canadian sample in which the NART-R and WAIS-R were administered concurrently.

Review of the cross-validation results of the remaining

predictive VIQ equations yielded better results on cross-validation than standardization studies, which was considered unusual as cross-validation typically produces worse estimates (Graves et al. in press). The Carswell et al. (1997) WAIS-R Vocabulary/NART equation likely showed improved results on cross-validation given the concurrent administration of the WAIS-R and the predictive VIQ measures used in this equation in the current study. However, the Carswell et al. (1997) equation was originally developed based on the accuracy of independent VIQ predictors (i.e., NART errors and WAIS-R Vocabulary age-scaled scores) at postdicting WAIS-R VIQs obtained 5 years earlier. As such, it was difficult to directly compare the results of the current study with the results of the original Carswell et al. (1997) study as the administration times of VIQ predictors with the criterion variable (WAIS-R VIQ) were not equivalent in these studies. The more appropriate cross-validation would involve a longitudinal cross-validation of the Carswell et al. (1997) equation. Nevertheless, the current study provided useful information as to cognitive decline sensitivity scores for the Carswell et al. (1997) equation based on concurrent testing, which can be directly applied to the clinical situation in which concurrent testing is used to detect cognitive decline.

It should be noted that although the improved result of the Carswell et al. (1997) WAIS-R Vocabulary/NART equation

on cross-validation may have been due to the concurrent administration of the WAIS-R and predictive VIQ measures in the current study, the other Carswell et al. (1997) equation, based on the NART alone, was developed from the same postdictive sample and yielded poorer results on cross-validation. However, the NART based Carswell et al. (1997) equation demonstrated poor predictive accuracy at estimating actual WAIS-R VIQs in the current study which may have contributed to the poorer sensitivity score on cross-validation, regardless of the concurrent administration of predictive and criterion VIQ measures in this study.

Cross-validation of the Barona et al. (1984) equation in the current study also produced a more clinically useful WAIS-R VIQ cognitive decline sensitivity score (23.1 VIQ points) as compared to the WAIS-R cross-validation sensitivity score based on the original standardization sample (29.4 VIQ points). Other cross-validation studies of the Barona et al. (1984) equation have produced WAIS-R VIQ cross-validation cognitive decline sensitivity scores of 26.5 to 37.9 VIQ points (Eppinger et al. 1987; Grober & Sliwinski, 1991; Paolo & Ryan 1992). Grober and Sliwinski (1991) attributed their poor cross-validation result of the Barona et al. (1984) equation (cognitive decline sensitivity score = 37.9 VIQ points) to a demographically homogeneous sample which yielded a low correlation between estimated and obtained WAIS-R VIQs. The improved results of the Barona et

al. (1984) equation on cross-validation in the current study may therefore have been attributable to a demographically diverse sample that achieved a correlation of 0.47 between estimated and obtained WAIS-R VIQs. However, it was more likely the contribution of education (point biserial) to the Barona et al. (1984) equation that improved the prediction of WAIS-R VIQs as education itself produced a correlation of 0.47 with WAIS-R VIQs in the above sample. Education has previously been shown to be a powerful predictor in demographic IQ equations (Spreen & Strauss, 1998).

The current study was also the first study to cross-validate the Barona et al. (1984) equation in a healthy elderly Canadian sample in which WAIS-R VIQs were available. Helmes (1996) previously examined the utility of the Barona et al. (1984) equation at predicting IQ performance in a normal elderly Canadian sample, but was unable to directly cross-validate the results with WAIS-R IQs as the sample had only received four subtests of the WAIS-R at the time of testing. The results of the current study were encouraging regarding the predictive/clinical utility of the Barona et al. (1984) equation in an elderly Canadian sample. Further studies with healthy elderly Canadians are required however, to confirm the above cross-validation results.

Summary

In summary, the current study reviewed the sensitivity

of existing VIQ equations at detecting cognitive decline in a new sample of normal elderly Canadians. The above results suggested that the Carswell et al. (1997) WAIS-R Vocabulary/NART equation and the Krull et al. (1995) WAIS-R Vocabulary/Demographics equation demonstrated the greatest clinical utility of the equations considered for detecting cognitive decline with cross-validation cognitive decline sensitivity scores of less than 16 VIQ points.

2. ***Development of New VIQ Prediction Equations***

In the current study, a new predictive WAIS-R VIQ regression equation utilizing WAIS-R Vocabulary age-scaled scores and Spot-the-Word (STW) subtest raw scores was developed that demonstrated significant utility for detecting cognitive decline in a clinical sample. The sensitivity of the new WAIS-R VIQ prediction equation for detecting cognitive decline was examined in the context of discrepancy values between obtained and predicted VIQ scores. Given the normal distribution of WAIS-R VIQ discrepancy scores, sensitivity scores for detecting cognitive decline were calculated based on the percentile table approach and the SEe method. The sensitivity score based on the percentile table approach was 13 VIQ points. However, the percentile table approach does not include a correction for the shrinkage expected on cross-validation. An advantage of the SEe method is that a correction for the

shrinkage expected on cross-validation can be applied to provide an estimate of the expected accuracy of prediction of the equation in a new sample (Graves, 1998). The cross-validation sensitivity score based on the estimated cross-validation SEe was calculated at 14.7 VIQ points. This cross-validated/corrected cognitive decline sensitivity score compared very favourably with the 13 VIQ point cross-validation standard of comparison proposed by Graves et al. (in press). Thus, the new WAIS-R VIQ prediction equation was judged to be of considerable clinical utility because it could reliably detect a decline of approximately 15 VIQ points between predicted and obtained WAIS-R VIQ scores. The clinical sensitivity of this new WAIS-R VIQ regression equation was in fact the most sensitive of any currently available predictive WAIS-R VIQ method for detecting cognitive decline.

Previous review of the literature has shown that existing methods of predicting premorbid VIQ have difficulties predicting VIQs outside of the median range, often resulting in serious over or underestimation of VIQ scores. The accuracy of the new WAIS-R VIQ equation at predicting extreme versus median range VIQ scores was determined by examining mean discrepancy scores by IQ range classification. The results of Table 10 showed that the new equation produced a pattern of predictive accuracy reflected by other predictive VIQ measures with a tendency to

underestimate high VIQ scores and overestimate low VIQ scores. The accuracy of prediction of median and high extreme VIQs was quite good as the mean error between obtained and predicted VIQs was no more than 5.33 VIQ points for VIQ scores between 90 to 130+ VIQ points. However, prediction of low extreme VIQ scores produced a mean error of -10.73 VIQ points for VIQ scores between 80 to 89 VIQ points. Therefore, because the accuracy of the new WAIS-R VIQ equation was quite good at predicting median and high extreme VIQ scores, but considerably more limited at predicting scores below 90 VIQ points, it is recommended that this equation only be used to predict VIQ scores at or above 90 VIQ points.

Although the Vocabulary subtest of the Wechsler scales has frequently been used for predictive purposes given its high correlation with general intellectual ability (Vanderploeg et al. 1996), it has demonstrated susceptibility to cognitive decline in clinical samples. Vocabulary scores have been reported to be significantly lower for groups of patients with Alzheimer's disease, multi-infarct dementia, alcoholic dementia, Korsakoff's dementia, Huntington's disease, and depression when compared to controls matched for age, gender and education (Nelson & McKenna, 1975; Hart et al. 1986; Crawford et al. 1988; Crawford et al. 1987; Sharpe & O'Carroll, 1991). However, of the studies that have examined Wechsler Vocabulary

performance in patients with Alzheimer's disease, all patients were at least moderately demented, with no patients in the early/mild stages of the disease process.

Given the concern regarding the resistance of the Wechsler Vocabulary subtest to decline even in the earliest stages of Alzheimer's disease, a second predictive WAIS-R VIQ equation was developed with Vocabulary excluded from the regression analysis. This second predictive WAIS-R VIQ equation utilized NART-R errors and STW raw scores. The cognitive decline sensitivity score for the new equation was 15 VIQ points based on the percentile table approach. The cross-validation cognitive decline sensitivity score based on the estimated cross-validated SEe was 16.1 VIQ points. These sensitivity scores were comparable given the normal distribution of discrepancy values between obtained and predicted VIQ scores in the current sample, although the sensitivity score based on the percentile table approach was not corrected for the shrinkage expected on cross-validation. The cross-validation cognitive decline sensitivity score (i.e., 16.1 VIQ points) compared favourably with the 13 VIQ point WAIS-R test-retest cross-validation standard of comparison (Graves et al. in press). Thus, this second predictive WAIS-R VIQ equation was judged to be of moderate clinical utility because it could reliably detect a decline of approximately 16 VIQ points between predicted and obtained WAIS-R VIQ scores.

The accuracy of this second predictive equation at estimating extreme versus median range VIQ scores was determined by examining mean discrepancy scores by IQ range classification. The results of Table 12 showed that for median VIQ scores of 90 to 129 VIQ points, the maximum error between obtained and predicted VIQs was 5.62 VIQ points. However, prediction of high extreme scores resulted in an underestimation of actual VIQs by an average of 10.74 VIQ points for VIQ scores of 130+ VIQ points, and prediction of low extreme scores resulted in an overestimation of actual VIQs by an average of 12.19 VIQ points for VIQ scores between 80 to 89 VIQ points. It is recommended that this equation only be used to predict VIQ scores between 90 to 129 VIQ points, because this equation demonstrated significant difficulty at predicting VIQs outside of the median range.

The results of the current study were consistent with other studies that have compared the utility of several IQ predictors at estimating WAIS/WAIS-R IQ in the same sample. Such studies have typically reported that a combination of predictive measures demonstrates greater utility at estimating WAIS/WAIS-R IQ scores than any independent predictor. In the current study, two clinically useful predictive WAIS-R VIQ equations were developed which combined WAIS-R Vocabulary scores and STW scores in one equation, and NART-R errors and STW scores in the second

equation. Of note, was the contribution of STW scores to both predictive equations as the STW test has only recently been considered as a predictive IQ measure. The STW test was developed in the United Kingdom and was considered to demonstrate potential utility for estimating premorbid intellectual functioning as it correlated highly with measures of verbal intelligence including the Mill Hill Vocabulary Scale and the NART in a normal elderly sample (Baddeley et al. 1993). The current study was, in fact, the first study to determine the validity of the STW test at estimating WAIS-R VIQs in a North American sample. The results of the current study showed that the STW test not only contributed significantly to the prediction of WAIS-R VIQs when combined with other VIQ predictors (i.e., WAIS-R Vocabulary subtest, NART-R errors), but independently accounted for a significant proportion of variance in WAIS-R VIQs and produced a clinically useful cross-validation cognitive decline sensitivity score of 16.9 VIQ points.

Several other VIQ predictors also independently accounted for a significant proportion of the variance in WAIS-R VIQs and produced clinically useful cross-validation cognitive decline sensitivity scores. The clinical utility of the WAIS-R Vocabulary subtest was marginally better than the combined predictive equation of NART-R errors and STW scores, as illustrated in Table 13, as the cross-validation sensitivity score for the WAIS-R Vocabulary subtest was 15.6

VIQ points, whereas the cross-validation sensitivity score for the combined NART-R/STW equation was 16.1 VIQ points. Independent VIQ predictors which demonstrated moderate clinical utility with cross-validation cognitive decline sensitivity scores of less than 17 VIQ points included NART-R errors, and STW raw scores. Independent VIQ predictors with more limited clinical utility with cross-validation sensitivity scores of more than 17 VIQ points included NART errors, WRAT-3 Reading raw scores, and education. It should be noted that this was the first study to examine the predictive/clinical utility of the WRAT-3 Reading subtest at estimating WAIS-R VIQs in a normal elderly Canadian sample.

Comparison of the NART and NART-R as independent VIQ predictors showed that estimation of WAIS-R VIQ by NART-R errors produced a more clinically useful cross-validation cognitive decline sensitivity score (16.7 VIQ points) than estimation with NART errors (17.7 VIQ points) in the above normal elderly Canadian sample. As such, it is suggested that the above normal elderly Canadian sample improved their reading performance by completing the North American (i.e., NART-R) versus the United Kingdom (i.e., NART) edition of this test, resulting in a more clinically useful cognitive decline sensitivity score for the NART-R. Thus, the current study supported the use of the NART-R in the estimation of WAIS-R VIQ for Canadian samples.

Summary

In summary, the current study developed two new clinically useful WAIS-R VIQ predictive equations that could reliably detect a decline of 14.7 (i.e., WAIS-R Vocabulary/STW) to 16.1 VIQ points (i.e., NART-R errors/STW) between predicted and obtained WAIS-R VIQs in a clinical sample. The newly developed WAIS-R Vocabulary/STW equation demonstrated the greatest predictive/clinical utility of any currently available method for estimating premorbid VIQ.

Verbal Memory and Language Prediction

The second and third parts of the study examined the concurrent validity of VIQ measures (i.e., existing and newly developed VIQ equations and independent VIQ predictors), measures of present ability, and demographic variables at predicting verbal memory and language performance, respectively, in the same normal elderly Canadian sample. The goal of this part of the study was to identify measures that would be useful at estimating memory and language performance in normal elderly individuals, so that these measures could serve as potential predictors of premorbid functioning and detectors of decline for patients with early Alzheimer's disease.

1. ***Verbal Memory Prediction***

In the domain of verbal memory, the CVLT long delay

free recall score (CVLTLDFR) was selected as the criterion measure for prediction as the delayed recall component of verbal memory has demonstrated sensitivity to decline in the preclinical stages of Alzheimer's disease. Potential predictors of verbal memory performance were compared to a cross-validation test-retest standard of comparison, proposed by Graves et al. (in press), of CVLT long delay free recall performance over a period of time for healthy elderly participants. Paolo et al. (1997a) administered the CVLT to healthy elderly Americans aged 54 to 87 years on two occasions. The test-retest reliability of the CVLT long delay free recall score for this elderly sample over a 1.3 year period was 0.76. Using this data, a cross-validation cognitive decline sensitivity score of 4.64 points was calculated for this sample in which previous CVLT long delay free recall scores were available. Thus, a cognitive decline sensitivity score of about 5 points for CVLT long delay free recall performance was applied as a standard of comparison for predictive measures that used an estimated previous CVLTLDFR score.

(a) VIQ Measures

The efficacy of existing VIQ equations, two newly developed VIQ equations (previous section) and independent VIQ predictors at estimating long delay free recall performance on the CVLT was examined in Tables 15 and 16.

The results of Tables 15 and 16 showed that all VIQ measures considered in the current study failed to correlate significantly with verbal memory performance in the above normal elderly Canadian sample. This finding was inconsistent with a previous study in which NART errors demonstrated utility at estimating verbal memory performance in a healthy elderly sample. Schlosser and Ivison (1989) developed a regression equation utilizing NART errors, Schonell Graded Word Reading Test (SGWRT) errors and age that accounted for 55% of the variance in general memory performance on the Wechsler Memory Scale (WMS) for healthy elderly participants. However, the results of the current study were not comparable to the Schlosser and Ivison (1989) study given that different criterion measures were used in each study. The CVLT long delay free recall score used in the current study represented the number of words recalled from multi-trial learning of a 16 item word list over a 20-minute delay whereas, general memory performance on the WMS (Memory Quotient: MQ) represented a summary score of performance on seven subtests including Personal and Current Information; Orientation; Mental Control; Logical Memory; Memory Span; Visual Reproduction; and Associate Learning. Although the validity, standardization and psychometric properties of the WMS have been independently criticized (Spreen & Strauss, 1998), the factors which contributed most to the inappropriateness of the MQ of the WMS as a

comparable criterion measure to the CVLTLDFR score included: 1) a lack of a delayed recall component to the summary score; 2) a lack of differentiation among various facets of memory (verbal vs. visual) by use of a summary score; and 3) inclusion of constructs in the summary score that were not genuine measures of memory (e.g., orientation, mental control). Alternately, the CVLT long delay free recall score assessed a specific component of verbal memory (i.e., long delay free recall), and was free of contamination by other memory/nonmemory constructs.

It should be noted that although all VIQ equations and VIQ predictors failed to correlate significantly with CVLTLDFR scores in the current sample, actual WAIS-R VIQ scores demonstrated a small but significant relationship with CVLTLDFR scores ($r = 0.278$, $p < .01$). Other studies have reported variable results with regard to the relationship between verbal memory measures and WAIS-R IQs in healthy elderly samples. Ivnik et al. (1991) reported a significant correlation of 0.48 ($p < .0001$) between the delayed recall raw scores of the Logical Memory subtest of the WMS and WAIS-R FSIQs in a sample of 99 healthy elderly persons aged 65 to 97 years. Paolo et al. (1997b) however failed to find a significant relationship between any of the CVLT scores and WAIS-R FSIQs (i.e., seven subtest short form) in a sample of 137 healthy elderly persons aged 53 to 94 years. As such, it is suggested that further investigation is

necessary to determine the relationship between verbal memory performance and IQ in elderly samples. Furthermore, even if such a relationship exists, it cannot be assumed that measures that demonstrate utility at predicting performance in one cognitive domain (i.e., IQ) can be applied to predict performance in another cognitive domain (i.e., verbal memory). The results of the current study clearly showed that even though WAIS-R VIQs were significantly correlated with CVLTLDFR scores, predictive VIQ measures failed to demonstrate utility in estimating CVLTLDFR performance in a healthy elderly Canadian sample. Therefore, specific investigation is warranted to determine the utility of VIQ predictors at estimating performance in cognitive domains other than general intellectual ability.

(b) ***Present Ability Memory Measures and Demographics***

The study then examined the utility of present ability memory measures and demographic variables at predicting CVLTLDFR performance in the above normal elderly Canadian sample. Stepwise multiple regression revealed that the CVLT Recency raw score and WAIS-R Digits Forward raw score contributed significantly to the prediction of CVLTLDFR performance accounting for 32% of the variance in CVLTLDFR scores. Given the normal distribution of CVLTLDFR discrepancy scores, sensitivity scores for the CVLTLDFR equation were calculated based on both the percentile table

and SEe approximation methods. The cognitive decline sensitivity score based on the percentile table approach was 6 points, whereas the cross-validation sensitivity score corrected for the shrinkage expected on cross-validation was 6.1 points.

The clinical utility of the new CVLT long delay free recall equation was determined by comparing the sensitivity score of this predictive memory measure to the CVLTLDFR cross-validation test-retest standard of comparison. The cross-validation sensitivity score for the new predictive CVLTLDFR equation (i.e., 6.1 points) compared favourably with the cross-validation test-retest standard of comparison of approximately 5 points. However, in terms of practical application, the cross-validation standard cognitive decline sensitivity value of approximately 5 points was considered to be somewhat large even with previous CVLTLDFR scores available. This was due to the relatively low test-retest reliability of CVLTLDFR scores ($r = 0.76$), which would lower the potential validity of any predictor. Thus, the new CVLTLDFR equation was judged to be of moderate clinical utility as it could only detect a decline of 6 or more points (of a total possible 16 points) as compared to predicted performance. Although the sensitivity of the new CVLTLDFR equation was adequate for detecting decline for persons that achieved high predicted CVLTLDFR scores (i.e., 12 to 16 points), the new equation was somewhat more limited

at detecting decline in a clinically useful range for persons that achieved mid-range/average predicted CVLTLDFR scores (i.e., 8 to 11 points). For example, if the predicted CVLTLDFR score for a 71-year-old woman was 11 points (mean CVLTLDFR score for the current sample and within the average range as compared to a normative sample; Spreen & Strauss, 1998), the woman would have to achieve an actual CVLTLDFR score of 5 points before decline could be reliably detected. A low score of 5 points on CVLT long delay free recall is consistent with at least moderate impairment as compared to a normative sample (Spreen & Strauss, 1998), reducing the need for a predictive memory measure to confirm a decline in performance.

However, it should be noted that although the predictive CVLTLDFR equation may only demonstrate moderate sensitivity for detecting cognitive decline in clinical samples, it demonstrates greater utility at detecting "impaired memory" than the CVLT alone. If "impaired" performance is defined by use of a *cut-off discrepancy score* which reflects the score necessary for 95% correct classification of normals (i.e., 1.645 SD) then, using the example above, the predictive CVLTLDFR equation would produce a cut-off score of 7 points [i.e., 11 - (1.645 x 2.454)], based on the estimated cross-validation SEe of the predictive CVLTLDFR equation (i.e., 2.454), for a person with a predicted CVLTLDFR score of 11 points. As such, if

the person received an actual CVLTLDFR score of less than 7 points (low-average range as compared to a normative sample; Spreen & Strauss, 1998), memory performance would be considered to be "impaired". However, if the predictive CVLTLDFR equation was not available, a cut-off score of 6.3 points [i.e., 11 - (1.645 x 2.84)] based on the mean (i.e., approximately 11) and standard deviation (i.e., 2.84) of the CVLTLDFR score would be used for all persons in the above sample. Therefore, the predictive CVLTLDFR equation offered a higher cut-off score which demonstrated greater sensitivity at detecting impaired memory performance than data based only on the CVLT in the above example in which predicted and mean CVLTLDFR scores were identical. The predictive CVLTLDFR equation was also considered to be more useful than the CVLT alone as it could provide more appropriate cut-off scores for individuals with predicted CVLTLDFR scores that deviated from the mean. Thus, the above discussion suggested that the new predictive CVLTLDFR equation was clinically useful at detecting impaired memory performance. This was also the first study to develop an equation utilizing present ability measures to predict long delay free recall performance in the domain of verbal memory.

Various predictive memory measures were initially considered in the development of the CVLTLDFR equation including VIQ measures, present ability memory measures, and

demographic variables. The results showed that a combination of present ability measures (i.e., CVLT Recency raw score; WAIS-R Digits Forward raw score) offered the best estimate of CVLT long delay free recall performance for this normal elderly Canadian sample, and provided sensitivity scores (i.e., cognitive decline sensitivity and cut-off scores) that were useful for detecting cognitive decline and impaired memory performance. Present ability measures (i.e., CVLT Recency raw score; WAIS-R Digits Forward raw score; Sentence Repetition corrected score) were selected for consideration in the prediction of CVLT long delay free recall performance based on the premise that these measures remained relatively resistant to decline in the early stages of Alzheimer's disease, and demonstrated potential utility for predicting verbal memory performance. The results of this study confirmed the utility of the Recency effect and verbal span procedures (i.e., WAIS-R Digits Forward raw score) for predicting verbal memory performance in a normal elderly sample. The resistance of the CVLT Recency effect (raw score) to decline in early Alzheimer's disease was recently confirmed by Simon et al. (1994) who administered the CVLT to patients with mild to moderate Alzheimer's disease and normal controls matched for age and education, and found no significant difference between groups on the Recency component of the CVLT word list. Several studies have also documented the resistance of the WAIS-R Digits

Forward subtest to decline in early Alzheimer's disease (Small et al. 1997; Lafleche & Albert, 1995; Carlesimo et al. 1994; Flicker et al. 1993; Lines et al. 1991; Orsini et al. 1988; Stordant et al. 1986). Thus, the CVLT Recency raw score and WAIS-R Digits Forward subtest were considered to be representative present ability memory measures because each demonstrated utility at estimating verbal memory performance in a normal elderly sample, and remained resistant to decline in mild to moderate Alzheimer's disease.

Demographic variables have often been used in combination with present ability measures to improve the prediction of the criterion measure in the domain of general intellectual ability. Unfortunately, demographic variables failed to improve the predictive utility of the regression equation developed in the current study for estimating CVLTLDFR performance. However, age and gender independently contributed to the prediction of CVLTLDFR performance. This result was consistent with prior studies that have documented a significant relationship between verbal memory performance and age in healthy individuals, with evidence of decreased recall with increasing age (Paolo et al. 1997b). The relationship between gender and long delay free recall performance in the above elderly sample was also consistent with previous studies that have reported that woman tend to outperform men on measures of delayed recall (Kramer et al.

1988; Geffen et al. 1990; Trahan & Quintana, 1990). It should be noted that although demographic variables including age and gender correlated significantly with CVLTLDFR scores, they demonstrated limited clinical utility at independently predicting CVLTLDFR scores and detecting cognitive decline with cross-validation sensitivity scores of approximately 7 points per variable.

The only measure that independently demonstrated utility at predicting CVLT long delay free recall performance was the CVLT Recency raw score with a cross-validation cognitive decline sensitivity score of 6.1 points, based on the cross-validation SEe. This sensitivity score was, in fact, equivalent to the sensitivity score of the combined predictive regression equation (i.e., CVLT Recency raw score + WAIS-R Digits Forward raw score), as illustrated in Table 20. The equivalence of cross-validation sensitivity scores for VIQ predictors with different multiple correlations and standard error of estimates was due to the fact that cross-validation calculations based on Cattin's (1978, 1980) approach consider the number of predictor variables that are available to the original regression, not the number that ultimately are entered into the equation (Graves, 1998). Thus, for the combined predictive equation, four variables were available to the original regression, whereas only one variable was available to the regression analysis when the CVLT Recency raw score

was considered as an independent predictor. As only two variables contributed to the combined predictive equation, the remaining two variables served to reduce the cross-validation multiple correlation and sensitivity score of the combined predictive equation, resulting in equivalent cross-validation sensitivity scores for different predictors. This finding provides support for Graves' (1998) recommendation that the fewest number of relevant predictors be used when developing predictive regression equations.

Summary

In summary, existing VIQ measures failed to demonstrate utility in predicting long delay free recall performance in the domain of verbal memory in a healthy elderly Canadian sample. A new predictive regression equation was developed to estimate CVLT long delay free recall performance, utilizing the CVLT Recency raw score and WAIS-R Digits Forward subtest raw score, that was deemed clinically useful for detecting impaired memory performance. This is the first study to develop a regression equation utilizing present ability measures to predict CVLT long delay free recall performance in a healthy elderly sample, that demonstrated utility at detecting impaired memory performance.

2. ***Language Prediction***

The current study also examined the concurrent validity

of VIQ measures, present ability language measures, and demographic variables at predicting language performance in the same normal elderly sample. The Boston Naming Test (BNT) was selected as the criterion measure for prediction in the domain of language, as it was the only nonmemory measure that has added to the preclinical detection of Alzheimer's disease (Welsh et al. 1992). Potential predictors of language performance were compared to cross-validation cognitive decline sensitivity scores based on test-retest reliability data of the BNT. Mitrushina and Satz (1991) administered the BNT to healthy elderly adults aged 57 to 85 years on three occasions, each administration separated by one year. Comparison of BNT test scores at the first and third administration produced a test-retest reliability of 0.89 for this elderly sample. Using this data, a cross-validation cognitive decline sensitivity score of 5.86 points was calculated for this sample in which previous BNT scores were available. Thus, a cross-validation cognitive decline sensitivity score of about 6 points for BNT performance was applied as a standard of comparison for predictive measures that used an estimated previous BNT score.

(a) VIQ Equations

The efficacy of existing VIQ equations and newly developed VIQ equations at predicting BNT performance in a

healthy elderly Canadian sample and detecting cognitive decline in clinical samples was determined by comparison of their respective cross-validation cognitive decline sensitivity scores with the BNT cross-validation test-retest sensitivity score of 5.86 points. Only sensitivity scores of VIQ equations that accounted for more than 10% of the variance in BNT scores were considered. It should be noted that although the cross-validation test-retest standard of comparison of approximately 6 points was considered to be clinically useful, it was judged to be the upper limit of the maximum discrepancy that could be considered clinically useful given the limited error range of BNT scores in healthy samples. As the distribution of BNT scores is skewed in healthy samples, with participants generally receiving a score of 47 to 60 of a possible 60 points depending on age and education level (Tombaugh & Hubley, 1997), a sensitivity score of more than 7 points was considered to be of limited clinical utility because a person's actual BNT score would likely be in the impaired range as compared to a normative sample before cognitive decline could be assumed, rendering the prediction of premorbid performance unnecessary. Examination of Table 23 revealed that only the newly developed WAIS-R Vocabulary/STW equation and the Krull et al. (1995) WAIS-R Vocabulary/Demographics equation produced cross-validation sensitivity scores of less than 7 points. VIQ equations which demonstrated limited utility at

detecting cognitive decline in clinical samples with cross-validation sensitivity scores of more than 7 points included: the newly developed NART-R/STW equation; the Carswell et al. (1997) WAIS-R Vocabulary/NART equation; the Blair and Spreen (1989) NART-R equation; the Carswell et al. (1997) NART equation; the Ryan and Paolo (1992) NART equation; and the Nelson and Willison (1991) NART equation.

However, it should also be noted that although most VIQ equations demonstrated limited utility at detecting cognitive decline in clinical samples, they provided cut-off discrepancy scores that offered greater sensitivity at detecting "impaired" language performance than normative data based on the BNT alone. For example, if a person with 12 years of education received a predicted BNT score of 56 based on the Blair and Spreen (1989) NART-R equation, a cut-off score of 51.01 (i.e., 56 - 4.99; see Table 23 for cut-off discrepancy score of 4.99 points for the Blair and Spreen (1989) NART-R equation) would be used to detect "impaired" naming performance. However, if the cut-off score based on the VIQ equation was not available, impaired performance for the same person could not be assumed until the actual BNT score fell below 49 points (10th percentile) as compared to the BNT norms for the above sample in Table 21. Given the skewed distribution of BNT scores in normal samples, it has been recommended that performance be compared to percentile table distributions rather than mean

scores (Tombaugh & Hubley, 1997). Therefore, the VIQ equation demonstrated greater sensitivity at detecting impaired performance on the BNT, than BNT normative data alone.

(b) VIQ Predictors, Present Ability Language Measures and Demographics

The study then examined the utility of independent VIQ predictors, present ability language measures, and demographic variables at predicting BNT performance in the above normal elderly Canadian sample. Stepwise multiple regression revealed that the WAIS-R Vocabulary subtest raw score, the Spot-the-Word (STW) subtest raw score, and the WRAT-3 Reading subtest raw score accounted for 54% of the variance in BNT scores. The sensitivity of the new BNT predictive equation for detecting cognitive decline was examined in the context of discrepancy values between obtained and predicted BNT scores. As the frequency distribution of BNT discrepancy scores was a nonnormal distribution, a cognitive decline sensitivity score of 6 points was obtained from the percentile table approach for the new BNT equation. However, a cognitive decline sensitivity score was also calculated for the new BNT equation based on the SEe method, for comparison, and to obtain an estimate of the sensitivity of the new BNT equation on cross-validation. Calculations based on the

estimated cross-validated SEe produced a cross-validation cognitive decline sensitivity score of 6.17 points.

These results showed that although discrepancy scores between predicted and actual BNT scores were not normally distributed in the above sample, cognitive decline sensitivity scores based on the SEe method (i.e., 6.17 points on cross-validation) and the percentile table approach (i.e., 6 points) failed to differ significantly. This result was consistent with Tabachnick and Fidell's (1989) contention that minor deviations from normality in a large sample, although significant, often fail to deviate enough from normality to make a realistic difference in the analysis. Other studies however, have documented significant differences between sensitivity scores calculated with the above methods, when discrepancy scores have not been normally distributed. Graves et al. (in press) illustrated this clearly with data from the Nelson and McKenna (1975) study which yielded an estimated cognitive decline sensitivity score of 14 FSIQ points based on the SEe approximation method (WAIS Vocabulary subtest estimating WAIS FSIQ), while a histogram of discrepancy scores yielded an estimated cognitive decline sensitivity score of 23 FSIQ points. Thus, the above results supported the recommendation by Graves et al. (in press) that new predictive equations should provide the standard error of estimate and complete percentile table of discrepancy scores for calculation of

cognitive decline sensitivity scores.

The clinical utility of the new BNT predictive equation was determined by comparing the sensitivity score of this new predictive language measure to the BNT cross-validation test-retest sensitivity score. The sensitivity score for the new predictive BNT equation of approximately 6 points compared favourably to the cross-validation test-retest sensitivity score (i.e., 5.86 points) and was considered to be clinically useful at detecting cognitive decline with a sensitivity score of less than 7 points. For example, if a 65-year-old woman with 12 years of education obtained a predicted BNT score of 56 (i.e., mean for current sample) and an actual BNT score of 50, falling above the 10th percentile as compared to BNT norms in Table 21, a decline in performance could only be detected with the predictive language measure.

A second predictive BNT regression equation was developed given the concern regarding the resistance of the WAIS-R Vocabulary subtest to decline in early Alzheimer's disease, as discussed previously. Stepwise multiple regression revealed that the PPVT-R raw score, STW raw score and WRAT-3 Reading subtest raw score accounted for 49% of the variance in BNT scores. The cognitive decline sensitivity score for this second BNT equation based on the percentile table approach was 6 points. The cross-validation cognitive decline sensitivity score based on the estimated

cross-validation SEe was 6.40 points. It should be noted that the above sensitivity scores were essentially equivalent, even though BNT discrepancy scores were not normally distributed. Thus, the nonnormality of the frequency distribution of BNT discrepancy scores once again failed to have a significant impact on the calculation of sensitivity scores in this sample.

In terms of clinical utility, this second predictive BNT regression equation compared favourably with the original BNT equation with a cognitive decline sensitivity score of approximately 6 points. The sensitivity score of this second BNT equation also compared favourably with the BNT cross-validation test-retest standard of comparison, and was less than 7 points, which supported the utility of this equation as a sensitive measure for detecting cognitive decline in a clinical sample.

The results of the current study showed that a combination of "language" predictors offered the best estimate of BNT performance in this normal elderly Canadian sample. Independently, three measures demonstrated utility at predicting BNT performance and detecting cognitive decline in clinical samples with cross-validation cognitive decline sensitivity scores (i.e., SEe method) of less than 7 points that compared favourably with the BNT cross-validation test-retest sensitivity score. These measures included the WAIS-R Vocabulary subtest, the PPVT-R and the

STW subtest, as illustrated in Table 28. Measures of limited clinical utility with sensitivity scores of more than 7 points included the NART, NART-R, WRAT-3 reading subtest and education.

The results of the current study were consistent with a previous study that demonstrated the utility of a VIQ predictor at estimating language performance. Crawford et al. (1992) examined the validity of the NART at estimating performance on a measure of verbal fluency and developed a regression equation based on NART errors that predicted 45% of the variance in verbal fluency scores in a sample of normal individuals. The BNT equations developed in the current study both utilized two VIQ predictors (i.e., STW; WRAT-3 Reading subtest) in combination with present ability measures that contributed significantly to the prediction of BNT performance and produced sensitivity scores that demonstrated utility at detecting cognitive decline in clinical samples. As previous research has reported a significant relationship between WAIS-R VIQ and BNT performance (Thompson & Heaton, 1989), which was also found in the current study, it was not unexpected that VIQ predictors contributed significantly to the estimation of BNT performance. Thus, the current results supported the use of VIQ predictors in the estimation of BNT performance and detection of cognitive decline in clinical samples.

Measures of present ability including the WAIS-R

Vocabulary subtest, the PPVT-R, and the BDAE object-recognition task were also considered in the estimation of confrontational naming (BNT) performance in the above normal elderly sample. Results indicated that the WAIS-R Vocabulary subtest contributed significantly to the prediction of BNT performance in combination with VIQ predictors in this sample. Independently, the WAIS-R Vocabulary subtest accounted for 40% of the variance in BNT performance and demonstrated the greatest potential utility of independent predictors at detecting cognitive decline in clinical samples with a cross-validation sensitivity score of 6.62 points. This result was consistent with previous studies that have reported a significant relationship between the WAIS-R Vocabulary subtest and BNT performance in cognitively intact adults ranging in age from 25 to 88 years (Tombaugh & Hubley, 1997), and in a clinical sample primarily characterized by heterogeneous neurological conditions (Thompson & Heaton, 1989). Similarly, the PPVT-R contributed significantly to the prediction of BNT performance in combination with VIQ predictors, and independently accounted for 38% of the variance in BNT performance in the above sample. Alternately, BDAE object-recognition (number correct; latency to complete task) failed to contribute significantly to the prediction of BNT performance in this sample. This was likely due to the relative ease of the object-recognition task which produced a restricted range of

scores for both measures (i.e., number correct; latency to complete task) and limited the predictive utility of the task. Given the above results, the current study supported the use of the WAIS-R Vocabulary subtest and the PPVT-R as present ability language measures in the prediction of BNT performance and detection of cognitive decline in clinical samples.

Demographic variables failed to improve the predictive utility of the regression equations developed in the current study for estimating BNT performance. However, education (i.e., point biserial) alone contributed significantly to the prediction of BNT performance. This result was consistent with previous studies that have shown education to be positively correlated with performance on the BNT (Thompson & Heaton, 1989; King, Caine, & Cox, 1993; Tombaugh & Hubley, 1997). Other demographic variables including age and gender failed to contribute significantly to the prediction of BNT performance in the above normal elderly sample. Although most studies with healthy participants over 70 years of age have typically documented an age effect, with BNT performance declining with advancing age, gender has generally failed to demonstrate a significant relationship with BNT performance (Spreen & Strauss, 1998). Recently, Tombaugh and Hubley (1997) examined the relationship between demographic variables and BNT performance in 219 cognitively intact Canadian adults aged

25-88 years and found that while education accounted for 8.9% of the variance in BNT scores, age and gender accounted for only 1.7% and 2.0%, respectively. Thus, the results of the current study were largely compatible with the Tombaugh and Hubley (1997) study as education accounted for approximately 7% of the variance in BNT scores, whereas gender and age failed to contribute significantly to the prediction of BNT performance in the above normal elderly Canadian sample.

Summary

In summary, existing VIQ equations demonstrated variable utility in detecting cognitive decline in clinical samples, with the newly developed WAIS-R Vocabulary/STW equation demonstrating the greatest predictive/clinical utility of the equations considered. VIQ equations, however, demonstrated greater utility at detecting impaired performance on the BNT than BNT normative data alone. Two new predictive BNT regression equations were developed that could reliably detect a decline of approximately 6 points between predicted and obtained BNT scores in a clinical sample. These equations utilized VIQ predictors and present ability language measures. These are the first regression equations utilizing VIQ predictors and present ability language measures to predict BNT performance in a healthy elderly sample that demonstrated utility for detecting

cognitive decline in a clinical sample.

Limitations

Although the current study was informative regarding the use of predictive variables in estimating performance and detecting decline in cognitive domains including general intellectual ability, verbal memory, and language, limitations of the study must be acknowledged. One limitation involved the use of several predictor variables in the estimation of the criterion variable and the development of prediction equations in the above cognitive domains. Although this was consistent with the investigative/comparative nature of the study, it was inefficient from a statistical perspective because it, 1) limited the sensitivity of the regression analysis, and 2) reduced the sensitivity of significant predictors on cross-validation. Regression analysis requires that the p value be adjusted according to the number of variables considered in the analysis in an effort to avoid a Type 1 error. However, as more variables are considered in the analysis, the p value becomes more restrictive, which increases the likelihood that variables will fail to reach the criterion to enter the analysis, reducing the sensitivity of the regression analysis and increasing the possibility of a Type 2 error (Graves, 1998). In the current study, attempts were made to reduce the number of variables

considered in the regression analysis by selecting only those variables that demonstrated a significant relationship with the criterion variable, and were the most appropriate for the sample (i.e., NART-R vs. NART). However, given the breadth of variables considered in the above analyses, particulary for VIQ and language prediction, it was likely that some variables failed to reach the criterion to enter the regression analysis. Therefore, significant variables were regressed independently on the criterion variable to provide specific information regarding the predictive and clinical utility of each variable.

Cross-validation calculations were also affected by the number of predictor variables considered in the investigation. Graves (1998) reported that formulas that estimate the cross-validation multiple correlation (R) of a predictive regression equation (i.e., Cattin's 1978, 1980 approach) consider the number of predictors that were available to the original regression, not the number that ultimately were entered into the equation. Therefore, if six variables were considered in the original prediction, estimation of the cross-validation R would be based on six predictors, even if only one variable was retained for the final regression equation (Graves, 1998). Furthermore, the remaining variables not entered in the regression equation would serve to reduce the sensitivity of the equation on cross-validation (Graves, 1998). This was illustrated

earlier in the discussion as the combined predictive CVLTLDFR equation, with a higher multiple correlation and more sensitive SEe, produced the same cognitive decline sensitivity score on cross-validation as the CVLT Recency raw score alone. The combined predictive CVLTLDFR equation with two predictors in the final equation was cross-validated based on four variables considered in the original prediction, whereas, only one predictor was considered in the cross-validation of the CVLT Recency raw score. Thus, the higher number of original predictors for the combined predictive equation served to reduced the sensitivity of the equation to the value of the CVLT Recency raw score on cross-validation. Therefore, the fewest number of variables should be used when developing predictive regression equations, as recommended by Graves (1998).

A second issue that must be addressed involved the administration of predictive VIQ, memory, and language measures at the same testing period as the criterion measure in each of the above cognitive domains. Carswell et al. (1997) considered that concurrent testing could maximize the correlation between predictive and criterion measures that test similar abilities, resulting in possible overestimation of the clinical utility of a predictive measure. Carswell et al. (1997) suggested that the best estimate of the accuracy of a premorbid cognitive measure would be the "standard error of postdiction", in which performance on the criterion

measure had actually been obtained several years prior to administration of the predictive/postdictive cognitive variable. Although information regarding the postdictive utility of an estimator of cognitive performance would be useful in determining the validity of the measure at estimating current versus previous performance, such information would be difficult to apply clinically because concurrent testing is generally used to estimate premorbid performance and detect cognitive decline in patient populations. Therefore, administration of predictive and criterion measures during the same testing period in the current study was considered to be appropriate as the clinical application of predictive measures would also involve concurrent testing (Graves et al. in press).

As well, before these new predictive equations can be applied to clinical samples, cross-validation of these equations is required with both normal and clinical samples to determine their true utility as predictive estimates of cognitive performance. Psychometric cross-validation of each predictive equation revealed minimal expected shrinkage of the validity of each equation at predicting cognitive performance on cross-validation with other normal elderly samples (Graves, 1998). Nevertheless, cross-validation with other healthy elderly participants is required to confirm these results, and clinical cross-validation with patients with early Alzheimer's disease is required to determine if

these new equations are useful for detecting cognitive decline in early Alzheimer's disease.

Graves et al. (in press) indicated that another issue that must be considered in the clinical application of a predictive equation is the demographic match between the patient and the normative sample. The current sample consisted of 98 healthy participants ranging in age from 53 to 89 years, with an average level of education of 13.6 years. However, only 21% of the participants were male. Although this finding was consistent with other studies that have documented a significantly higher proportion of female than male participants with advancing age (Unverzagt et al. 1998; Small et al. 1997; Wechsler, 1997; Frataglioni et al. 1991), this factor must be acknowledged when assessing an elderly male patient.

The current study was also constrained by test development/updating in both intellectual and language domains. In the domain of VIQ prediction, the newest version of the Wechsler scales (WAIS-III; Wechsler, 1997) was released for clinical use during the course of this study. Although predictive VIQ measures considered in this study will have to be updated to determine their utility at estimating WAIS-III VIQs, it is not anticipated that there will be significant differences in the predictive utility of these measures at estimating WAIS-III VIQs given the significant relationship between WAIS-R VIQs and WAIS-III

VIQs (r = 0.94; Wechsler, 1997). However, since the current regression equation was developed to predict WAIS-R VIQs, it should not be used to predict WAIS-III VIQs until sufficient cross-validation data exists.

In the domain of language prediction, predictive BNT equations were affected by the release of the WAIS-III (i.e., Vocabulary subtest) and the newest version of the Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997). These equations will have to be updated to determine the utility of these new predictive measures for estimating BNT performance. However, until such updated information is obtained, it is recommended that the regression equations developed in the current study only be calculated with the language predictors specified for each equation.

Finally, it must be taken into account that any present ability measure used in estimating premorbid performance remains potentially susceptible to cognitive decline even in the earliest stages of Alzheimer's disease. This factor cannot be disregarded in clinical samples even though review of the literature suggested that many of the present ability measures considered in the current study were thought to remain resistant to decline at least in the early stages of Alzheimer's disease. Because there has been concern about the resistance of the WAIS-R Vocabulary subtest to decline even in the earliest stages of Alzheimer's disease, regression equations were developed with and without this

variable to try to control for this possibility. Thus, it is necessary for the practising clinician to apply caution in estimating premorbid cognition and detecting cognitive decline on the basis of present ability measures.

Clinical Implications

Detection of cognitive decline is important in clinical practise. However, interpretation of discrepancy scores based on predictive equations should only be considered in the context of a comprehensive neuropsychological examination. Graves et al. (in press) identified potential pitfalls in the clinical interpretation of discrepancy scores given that the absence of a significant discrepancy does not eliminate the possibility of cognitive decline as the predictive measure may not have demonstrated adequate sensitivity to detect decline or adequate resistance to the effects of disease, whereas, the presence of a significant discrepancy provides only one piece of evidence consistent with cognitive decline. This finding must be corroborated with other evidence of impairment as 5% of normals can achieve scores below sensitivity values, and the decline may be related to functional rather than organic factors (i.e., depression vs. dementia). Furthermore, the limitations of predictive equations, as discussed previously, also require consideration. Therefore, it is recommended that the practising clinician only use the present predictive methods

in the context of a comprehensive neuropsychological examination in which changes in various areas of functioning (e.g., employment, daily tasks) can be investigated before cognitive decline can be assumed.

Conclusions

In conclusion, the current study identified the most useful measures for predicting performance in intellectual, verbal memory, and language domains in a healthy elderly Canadian sample that demonstrated utility for detecting cognitive decline in clinical samples. Several objectives were achieved in the domain of VIQ prediction; the current study: 1) provided normative data for elderly Canadians on several existing measures of premorbid VIQ; 2) compared the utility of various measures of premorbid VIQ at estimating WAIS-R VIQ in the same normal elderly sample based on a quantitative estimate (i.e., cognitive decline sensitivity score); 3) identified the most useful existing predictive measure/equation for estimating WAIS-R VIQ for elderly Canadians and detecting cognitive decline in clinical samples (i.e., Carswell et al. 1997 equation); 4) developed a new predictive WAIS-R VIQ regression equation utilizing WAIS-R Vocabulary age-scaled scores and STW raw scores that demonstrated the greatest utility at estimating WAIS-R VIQs and detecting cognitive decline of any current VIQ predictor; and 5) cross-validated some of the more recently

proposed measures of premorbid VIQ with a healthy elderly sample. The current study also examined the utility of predictive VIQ measures, present ability measures, and demographic variables at predicting performance in verbal memory and language domains. The results of this study were considered to be of value because this was the first study to develop equations to predict long delay free recall performance (i.e., verbal memory) and confrontation naming performance (i.e., language) to aid in the detection of cognitive decline in the early stages of Alzheimer's disease. The goal of the study was to develop predictive measures of premorbid memory and language functioning that would be sensitive enough to detect decline in the early stages of Alzheimer's disease so that diagnosis and intervention could be possible much earlier in the disease process than is currently available.

References

Almkvist, O., & Backman, L. (1993). Detection and staging of early clinical dementia. Acta Neurologica Scandinavica, 88, 10-15.

Austin, M.P., Ross, M., Murray, C., O'Carroll, R.E., Ebmeier, K.P., & Goodwin, G.M. (1992). Cognitive function in major depression. Journal of Affective Disorders, 25, 21-30.

Baddeley, A.D. (1986). Working memory. Oxford, England: Oxford University Press.

Baddeley, A. (1992). Working memory. Science, 255, 556-559.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1992). The Speed and Capacity of Language-Processing Test. Bury St. Edmunds: Thames Valley Test Company.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word test: A robust estimate of verbal intelligence based on lexical decision. British Journal of Clinical Psychology, 32, 55-65.

Baddeley, A.D., & Warrington, E.K. (1970). Amnesia and the distinction between long- and short-term memory. Journal of Verbal Learning and Verbal Behaviour, 9, 176-189.

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

Bayles, K.A., & Tomoeda, C.K. (1983). Confrontation naming impairment in dementia. Brain and Language, 19, 98-114.

Beardsall, L., & Brayne, C. (1990). Estimation of verbal intelligence in an elderly community: A prediction analysis using a shortened NART. British Journal of Clinical Psychology, 29, 83-90.

Beardsall, L., & Huppert, F.A. (1994). Improvement in NART word reading in demented and normal older persons using the Cambridge Contextual Reading Test. Journal of Clinical and Experimental Neuropsychology, 16, 232-242.

Beardsall, L., & Huppert, F. (1997). Short NART, CCRT and Spot-the-Word: Comparisons in older and demented persons. British Journal of Clinical Psychology, 36, 619-622.

Becker, J.T., Boller, F., Saxton, J., & McGonigle-Gibson, K.L. (1987). Normal rates of forgetting of verbal and non-verbal material in Alzheimer's disease. Cortex, 23, 59-72.

Berry, D.T.R., Carpenter, G.S., Campbell, D.A., Schmitt, F.A., Helton, K., & Lipke-Molby, T. (1994). The New Adult Reading Test-Revised: Accuracy in estimating

WAIS-R IQ scores obtained 3.5 years earlier from normal older persons. Archives of Clinical Neuropsychology, 9, 239-250.

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

Boekamp, J.R., Strauss, M.E., & Adams, N. (1995). Estimating premorbid intelligence in African-American and White elderly veterans using the American version of the National Adult Reading Test. Journal of Clinical and Experimental Neuropsychology, 17, 645-653.

Boller, F., Becker, J.T., Holland, A.L., Forbes, M.M., Hood, P.C., & McGonigle-Gibson, K.L. (1991). Predictors of decline in Alzheimer's disease. Cortex, 27, 9-17.

Bondi, M.W., Monsch, A.U., Galasko, D., Butters, N., Salmon, D.P., & Delis, D.C. (1994). Preclinical cognitive markers of dementia of the Alzheimer type. Neuropsychology, 8, 374-384.

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.

Brink, T.L., Yesavage, T.L., Lum, O., Heersma, P.H., Adey, M., & Rose, T.L. (1982). Screening tests for geriatric depression. Clinical Gerontologist, 1, 437-444.

Brooks, D.N. (1975). Long- and short-term memory in head-injured patients. Cortex, 11, 329-340.

Bucks, R.S., Scott, M.I., Pearsall, T., & Ashworth, D.L. (1996). The Short NART: Utility in a memory disorders clinic. British Journal of Clinical Psychology, 35, 133-141.

Buschke, H. (1973). Selective reminding for analysis of memory and learning. Journal of Verbal Learning and Verbal Behaviour, 12, 543-550

Canadian Study of Health and Aging Working Group (1994). Canadian study of health and aging: Study methods and prevalence of dementia. Canadian Medical Association Journal, 150, 899-913.

Capitani, E., Della Sala, S., Logie, R.H., & Spinnler (1992). Recency, primacy, and memory: Reappraising and standardising the serial position curve. Cortex, 28, 315-342.

Carlesimo, G.A., Fadda, L., Lorusso, S., & Caltagirone, C. (1994). Verbal and spatial memory spans in Alzheimer's and multi-infarct dementia. Acta Neurologica Scandinavica, 89, 132-138.

Carswell, L.M., Graves, R.E., Snow, W.G., & Tierney, M.C.

(1997). Postdicting verbal IQ of elderly individuals. Journal of Clinical and Experimental Neuropsychology, 19, 914-921.

Carvajal, H., Shaffer, C., & Weaver, K.A. (1989). Correlations of scores of maximum security inmates on Wechsler Adult Intelligence Scale-Revised and Peabody Picture Vocabulary Test-Revised. Psychological Reports, 65, 268-270.

Cattin, P. (1978). A predictive-validity-based procedure for choosing between regression and equal weights. Organizational Behaviour and Human Performance, 22, 93-102.

Cattin, P. (1980). Note on the estimation of the squared cross-validated multiple correlation of a regression model. Psychological Bulletin, 87, 63-65.

Cherry, B.J., Buckwalter, J.G., & Henderson, V.W. (1996). Memory span procedures in Alzheimer's disease. Neuropsychology, 10, 286-293.

Conway, S.C., & O'Carroll, R.E. (1997). An evaluation of the Cambridge Contextual Reading Test (CCRT) in Alzheimer's disease. British Journal of Clinical Psychology, 36, 623-625.

Cooper, D., & Fraboni, M. (1988). Relationship between the Wechsler Adult Intelligence Scale-Revised and the Wide Range Achievement Test-Revised in a sample of normal adults. Educational and Psychological Measurement, 48,

799-803.

Corrigan, S.K., & Berry, D.T.R. (1992). Prediction of IQ in normal older persons: A comparison of two methods. Archives of Clinical Neuropsychology, 7, 323-324.

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.R., & Allan, K.M. (1997). Estimating premorbid WAIS-R IQ with demographic variables: Regression equations derived from a UK sample. The Clinical Neuropsychologist, 11, 192-197.

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., Moore, J.W., & Cameron, I.M. (1992). Verbal fluency: A NART-based equation for the estimation of premorbid performance. British Journal of Clinical Psychology, 31, 327-329.

Crawford, J.R., Parker, D.M., Allan, K.M., Jack, A.M., &

Morrison, F.M. (1991). The Short NART: Cross-validation, relationship to IQ and some practical considerations. British Journal of Clinical Psychology, 30, 223-229.

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., & DeLacey, 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., Foulds, J., Besson, J.A.O., & Parker, D.M. (1989). Estimating premorbid IQ from demographic variables: A regression equation derived from a UK sample. British Journal of Clinical Psychology, 28, 275-278.

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.

Cummings, J.L., Benson, D.F., Hill, M., & Read, S. (1985).

Aphasia in dementia of the Alzheimer type. Neurology, 35, 394-397.

Cummings, J.L., Houlihan, J.P., & Hill M. (1986). The pattern of reading deterioration in dementia of the Alzheimer type: Observations and implications. Brain and Language, 29, 315-323.

Darlington, R.B. (1968). Multiple regression in psychological theory and practise. Psychological Bulletin, 69, 161-182.

Della Sala, S., Pasetti, C., & Sempio, P. (1987). Deficit of the "Primacy Effect" in Parkinsonians interpreted by means of the working memory model. Swiss Archives of Neurology, Neurosurgery and Psychiatry, 138, 5-14.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). California Verbal Learning Test: Adult Version. San Antonio, TX: The Psychological Corporation.

Dunn, L.M., & Dunn, E.S. (1981). Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service.

Dunn, L.M., & Dunn, E.S. (1997). Peabody Picture Vocabulary Test-III. Circle Pines, MN: American Guidance Service.

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.

Flicker, C., Ferris, S.H., Crook, T., & Bartus, R.T. (1987). Implications of memory and language dysfunction in the naming deficit of senile dementia. Brain and Language, 31, 187-200.

Flicker, C., Ferris, S.H., & Reisberg, B. (1993). A two year longitudinal study of cognitive function in normal aging and Alzheimer's disease. Journal of Geriatric Psychiatry and Neurology, 6, 84-96.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-Mental State". A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12, 189-198.

Fox, N.C., Warrington, E.K., Seiffer, A.L., Agnew, S.K., & Rossor, M.N. (1998). Presymptomatic cognitive deficits in individuals at risk of familial Alzheimer's disease: A longitudinal prospective study. Brain, 121, 1631-1639.

Fratiglioni, L., Grut, M., Forsell, Y., Viitanen, M., Grafstrom, M., Holmen, K., Ericsson, K., Backman, L., Ahlbom, A., & Winblad, B. (1991). Prevalence of Alzheimer's disease and other dementias in an elderly urban population: Relationship with age, sex and education. Neurology, 41, 1886-1892.

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

Reading Test. Cortex, 27, 367-376.

Fuld, P.A., Masur, D.M., Blau, A.D., Crystal, H., & Aronson, M.K. (1990). Object memory evaluation for prospective detection of dementia in normal functioning elderly: Predictive and normative data. Journal of Clinical and Experimental Neuropsychology, 12, 520-528.

Geffen, G., Moar, K.J., O'Hanlon, A.P., Clark, C.R., & Geffen, L.B. (1990). Performance measures of 16-86-year-old males and females on the Auditory Verbal Learning Test. The Clinical Neuropsychologist, 4, 45-63.

Glanzer, M., & Cunitz, A.R. (1966). Two storage mechanisms in free recall. Journal of Verbal Learning and Verbal Behaviour, 5, 351-360.

Glosser, G., Libon, D.J., & Friedman, R.B. (1993). Comparison of two reading tests to estimate premorbid intelligence. Archives of Clinical Neuropsychology, 8, 228.

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.

Goodglass, H., & Kaplan, E. (1983). Boston Diagnostic Aphasia Examination (BDAE). Philadelphia: Lea and

Febiger.

Graves, R.E. (1998). Validity of multiple regression prediction models: Issues for neuropsychologists. (Unpublished manuscript).

Graves, R.E., Carswell, L.M., & Snow, W.G. (in press). An evaluation of the sensitivity of premorbid IQ estimators for detecting cognitive decline. Psychological Assessment.

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

Halperin, J.M., Healey, J.M., Zeitchik, E., Ludman, W.L., & Weinstein, L. (1989). Developmental aspects of linguistic and mnestic abilities in normal children. Journal of Clinical and Experimental Neuropsychology, 11, 518-528.

Harrison, P.L., Kaufman, A.S., Hickman, J.A., & Kaufman, N.L. (1988). A survey of tests used for adult assessment. Journal of Psychoeducational Assessment, 6, 188-198.

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

Hathaway, S.R., & McKinley, J.C. (1943). Booklet for the Minnesota Multiphasic Personality Inventory. New York:

The Psychological Corporation.

Helmes, E. (1996). Use of the Barona method to predict premorbid intelligence in the elderly. The Clinical Neuropsychologist, 10, 255-261.

Henderson, V.W., & Buckwalter, J.G. (1994). Cognitive deficits of men and women with Alzheimer's disease. Neurology, 44, 90-96.

Houlihan, J.P., Abrahams, J.P., LaRue, A.A., & Jarvik, L.F. (1985). Qualitative differences in vocabulary performance of Alzheimer versus depressed patients. Developmental Neuropsychology, 1, 139-144.

Howieson, D.B., Dame, A., Camicioli, R., Sexton, G., Payami, H., & Kaye, J.A. (1997). Cognitive markers preceding Alzheimer's dementia in the healthy oldest old. Journal of the American Geriatric Society, 45, 584-589.

Incalzi, R.A., Capparella, O., Gemma, A., Marra, C., & Carbonin, P.U. (1995). Effects of aging and of Alzheimer's disease on verbal memory. Journal of Clinical and Experimental Neuropsychology, 17, 580-589.

Ivnik, R.J., Smith, G.E., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1991). Wechsler Memory Scale: IQ-Dependent norms for persons aged 65 to 97 years. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3, 156-161.

Jacobs, D.M., Sano, M., Dooneief, G., Marder, K., Bell, K.L., & Stern, Y. (1995). Neuropsychological detection

and characterization of preclinical Alzheimer's disease. Neurology, 45, 957-962.

Jastak, S., & Wilkinson, G.S. (1984). WRAT-R: Wide Range Achievement Test. Los Angeles: Western Psychological Services.

Johnstone, B., Callahan, C.D., Kapila, C.J., & Bouman, D.E. (1996). The comparability of the WRAT-R reading test and NAART as estimates of premorbid intelligence in neurologically impaired patients. Archives of Clinical Neuropsychology, 11, 513-519.

Johnstone, B., Hexum, C.L., & Ashkanazi, G. (1995). Extent of cognitive decline in traumatic brain injury based on estimates of premorbid intelligence. Brain Injury, 9, 377-384.

Johnstone, B., & Wilhelm, K.L. (1996). The longitudinal stability of the WRAT-R Reading subtest: Is it an appropriate estimate of premorbid intelligence? Journal of the International Neuropsychological Society, 2, 282-285.

Kaplan, E.F., Goodglass, H., & Weintraub, S. (1983). The Boston Naming Test (2nd. ed.). Philadelphia: Lea & Febiger.

Kareken, D.A., Gur, R.C., & Saykin, A.J. (1995). Reading on the Wide Range Achievement Test-Revised and parental education as predictors of IQ: Comparison with the Barona formula. Archives of Clinical Neuropsychology,

10, 147-157.

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.

Kaszniak, A.W., Garron, D.C., & Fox, J. (1979). Differential effects of age and cerebral atrophy upon span of immediate recall and paired-associate learning in older patients suspected of dementia. Cortex, 15, 285-295.

King, D.A., Caine, E.D., & Cox, C. (1993). Influences of depression and age on selected cognitive functions. The Clinical Neuropsychologist, 7, 443-453.

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

Kopelman, M.D. (1985). Rates of forgetting in Alzheimer's type dementia and Korsakoff's syndrome. Neuropsychologia, 23, 623-638.

Kramer, J.H., Delis, D.C., & Daniel, M. (1988). Sex differences in verbal learning. Journal of Clinical Psychology, 44, 907-915.

Krull, K.R., Scott, J.G., & Sherer, M. (1995). Estimation of premorbid intelligence from combined performance and demographic variables. The Clinical Neuropsychologist, 9, 83-88.

Kupke, T.E. (1996). NAART prediction of 10 year old ASVAB

scores. Poster presented at the 104th Annual Meeting of the American Psychological Association, Toronto, Canada.

LaBarge, E., Edwards, D., & Knesevich J. Wm. (1986). Performance of normal elderly on the Boston Naming Test. Brain and Language, 27, 380-384.

Lafleche, G., & Albert, M.S. (1995). Executive function deficits in mild alzheimer's disease. Neuropsychology, 9, 313-320.

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

Lezak, M.D. (1995). Neuropsychological assessment. (3rd ed). New York: Oxford University Press.

Lines, C.R., Dawson, C. Preston, G.C., Reich, S., Foster, C., & Traub, M. (1991). Memory and attention in patients with senile dementia of the Alzheimer type and in normal elderly subjects. Journal of Clinical and Experimental Neuropsychology, 13, 691-702.

Linn, R.T., Wolf, P.A., Bachman, D.L., Knoefel, J.E., Cobb, J.L., Belanger, A.J., Kaplan, E.F., & D'Agostino, R. (1995). The "preclinical phase" of probable Alzheimer's disease: A 13-year prospective study of the Framingham cohort. Archives of Neurology, 52, 485-490.

Maddrey, A.M., Cullum, C.M., Weiner, M.F., & Filley, C.M. (1996). Premorbid intelligence estimation and level of dementia in Alzheimer's disease. Journal of the

International Neuropsychological Society, 2, 551-555.

Malec, J.F., Ivnik, R.J., & Smith, G.E. (1993). Neuropsychology and normal aging: The clinician's perspective. In: R.W. Parks, R.F. Zec & R.S. Wilson (Eds.). Neuropsychology of Alzheimer's disease and other dementias. New York: Oxford University Press, 81-111.

Mangiaracina, J., & Simon, M.J. (1986). Comparison of the PPVT-R and WAIS-R in state hospital psychiatric patients. Journal of Clinical Psychology, 42, 817-820.

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

Masur, D.M., Sliwinski, M., Lipton, R.B., Blau, A.D., & Crystal, H.A. (1994). Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. Neurology, 44, 1427-1432.

Matarazzo, 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.

Maxwell, J.K., & Wise, F. (1984). PPVT validity in adults: A measure of vocabulary, not of intelligence. Journal of Clinical Psychology, 40, 1048-1053.

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

H.J. Eysenck (Ed.). Handbook of Abnormal Psychology. New York: Basic Books.

Miller, E. (1971). On the nature of the memory disorder in presenile dementia. Neuropsychologia, 9, 75-81.

Miller, E. (1973). Short- and long-term memory in presenile dementia (Alzheimer's disease). Psychological Medicine, 3, 221-224.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Mitrushina, M., Drebing, C., Uchiyama, C., Satz, P., Van Gorp, W., & Chervinsky, A. (1994). The pattern of deficit in different memory components in normal aging and dementia of the Alzheimer's type. Journal of Clinical Psychology, 50, 591-595.

Mitrushina, M., & Satz, P. (1991). Effect of repeated administration of a neuropsychological battery in the elderly. Journal of Clinical Psychology, 47, 790-800.

Mitrushina, M., Uchiyama, C., & Satz, P. (1995). Heterogeneity of cognitive profiles of normal aging: Implications for early manifestations of Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 17, 374-382.

Mittenberg, W., Thompson, G.B., Schwartz, J.A., Ryan, J.J., & Levitt, R. (1991). Intellectual loss in Alzheimer's dementia and WAIS-R intrasubtest scatter. Journal of

Clinical Psychology, 47, 544-547.

Morris, R.G. (1984). Dementia and the functioning of the articulatory loop system. Cognitive Neuropsychology, 1, 143-157.

Morris, R.G. (1986). Short-term forgetting in senile dementia of the Alzheimer's type. Cognitive Neuropsychology, 3, 77-97.

Morris, R.G. (1987). Articulatory rehearsal in Alzheimer's type dementia. Brain and Language, 30, 351-362.

Morris, R.G. (1994). Working memory in Alzheimer-type dementia. Neuropsychology, 8, 544-554.

Mortensen, E.L., Gade, A., & Reinisch, J.M. (1991). A critical note on Lezak's "Best Performance Method" in clinical neuropsychology. Journal of Clinical and Experimental Neuropsychology, 13, 361-371.

Moscovitch, M. (1982). A neuropsychological approach to perception and memory in normal and pathological aging. In: F.I.M. Craik and S. Trehub (Eds.). Aging and cognitive processes. New York: Plenum Press, 55-78.

Murdoch, B.E., Chenery, H.J., Wilks, V., & Boyle, R.S. (1987). Language disorders in dementia of the Alzheimer type. Brain and Language, 31, 122-137.

Naugle, R.I., Cullum, C.M., & Bigler, E.D. (1990). Evaluation of intellectual and memory function among dementia patients who were intellectually superior. The Clinical Neuropsychologist, 4, 355-374.

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.

Nelson, H.E., & O'Connell, A. (1978). Dementia: The estimation of pre-morbid intelligence levels using the New Adult Reading Test. Cortex, 14, 234-244.

Nelson, H.E., & Willison, J.R. (1991). The revised National Adult Reading Test-Test Manual. In: H. Nelson (Ed.). Windsor: NFER.

Norusis, M.J. (1996). SPSS for Windows Base System User's Guide (release 7.5). Upper Saddle River, NJ: Prentice Hall.

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. (1995). The assessment of premorbid ability: A critical review. Neurocase, 1, 83-89.

O'Carroll, R.E., Baikie, E.M., & Whittick, J.E. (1987). Does the National Adult Reading Test hold in dementia? British Journal of Clinical Psychology, 26, 315-316.

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 premorbid intellectual ability in the Alcoholic Korsakoff syndrome. Psychological Medicine, 22, 903-909.

O'Carroll, R.E., Prentice, N., Murray, C., Van Beck, M., Ebmeier, K.P., & Goodwin, G.M. (1995). Further evidence that reading ability is not preserved in Alzheimer's disease. British Journal of Psychiatry, 167, 659-662.

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) as a measure of premorbid ability. Schizophrenia Research, 8, 137-141.

Orsini, A., Trojano, L., Chiacchio, L., & Grossi, D. (1988). Immediate memory spans in dementia. Perceptual and Motor Skills, 67, 267-272.

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.

Paolo, A.M., Ryan, J.J., & Tröster, A.I. (1997). Estimating premorbid WAIS-R intelligence in the elderly: An extension and cross validation of new regression

equations. Journal of Clinical Psychology, 53, 647-656.

Paolo, A.M., Ryan, J.J., Tröster, A.I., & Hilmer, C.D. (1996a). Utility of the Barona demographic equations to estimate premorbid intelligence: Information from the WAIS-R standardization sample. Journal of Clinical Psychology, 52, 335-343.

Paolo, A.M., Ryan, J.J., Tröster, A.I., & Hilmer, C.D. (1996b). Demographically based regression equations to estimate WAIS-R subtest scales scores. The Clinical Neuropsychologist, 10, 130-140.

Paolo, A.M., Tröster, A.I., & Ryan, J.J. (1997a). Test-retest stability of the California Verbal Learning Test in older persons. Neuropsychology, 11, 613-616.

Paolo, A.M., Tröster, A.I., & Ryan, J.J. (1997b). California Verbal Learning Test normative data for the elderly. Journal of Clinical and Experimental Neuropsychology, 19, 220-234.

Paolo, A.M., Tröster, A.I., Ryan, J.J., & Koller, W.C. (1997). Comparison of NART and Barona demographic equation premorbid IQ estimates in Alzheimer's disease. Journal of Clinical Psychology, 53, 713-722.

Paque, L., & Warrington, E.K. (1995). A longitudinal study of reading ability in patients suffering from dementia. Journal of the International Neuropsychological Society, 1, 517-524.

Parkinson, S.R., Lindholm, J.M., & Inman, V.W. (1982). An

analysis of age differences in immediate recall. Journal of Gerontology, 37, 425-431.

Parks R.W., Zec, R.F., & Wilson, R.S. (1993). Neuropsychology of Alzheimer's disease and other dementias. New York: Oxford University Press.

Patterson, K., Graham, N., & Hodges, J.R. (1994). Reading in dementia of the Alzheimer's type: A preserved ability? Neuropsychology, 8, 395-407.

Pepin, E.P., & Eslinger, P.J. (1989). Verbal memory decline in Alzheimer's disease: A multiple-processes deficit. Neurology, 39, 1477-1482.

Petersen, R.C., Smith, G.E., Ivnik, R.J., Kokmen, E., & Tangalos, E.G. (1994). Memory function in very early Alzheimer's disease. Neurology, 44, 867-872.

Prout, H.T., & Schwartz, J.F. (1984). Validity of the Peabody Picture Vocabulary Test-Revised with mentally retarded adults. Journal of Clinical Psychology, 40, 584-587.

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

Raguet, M.L., Campbell, D.A., Berry, D.T.R., Schmitt, F.A., & Smith, G.T. (1996). Stability of intelligence and intellectual predictors in older persons. Psychological Assessment, 8, 154-160.

Rey, A. (1964). L'Examen Clinique en Psychologie. Paris:

Press Universitaire de France.

Ripich, D.N., Petrill, S.A., Whitehouse, P.J., & Ziol, E.W. (1995). Gender differences in language of AD patients: A longitudinal study. Neurology, 45, 299-302.

Robinson-Whelen, S., & Storandt, M. (1992). Immediate and delayed prose recall among normal and demented adults. Archives of Neurology, 49, 32-34.

Roper, B.L., Bieliauskas, L.A., & Peterson, M.R. (1996). Validity of the Mini-Mental State Examination and the Neurobehavioural Cognitive Status Examination in cognitive screening. Neuropsychiatry, Neuropsychology, and Behavioural Neurology, 9, 54-57.

Ross, T.P., Lichtenberg, P.A., & Christensen, B.K. (1995). Normative data on the Boston Naming Test for elderly adults in a demographically diverse medical sample. The Clinical Neuropsychologist, 9, 321-325.

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., Paolo, A.M., & Brungardt, T.M. (1990). Standardization of the Wechsler Adult Intelligence Scale-Revised for persons 75 years and older. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2, 404-411.

Schlosser, D., & Ivison, D. (1989). Assessing memory deterioration with the Wechsler Memory Scale, the

National Adult Reading Test, and the Schonell Graded Word Reading Test. Journal of Clinical and Experimental Neuropsychology, 11, 785-792.

Schlottmann, R., & Johnsen, D. (1991). The Intellectual Correlates Scale and the prediction of premorbid intelligence in brain-damaged adults. Archives of Clinical Neuropsychology, 6, 363-374.

Scott, J.G., Krull, K.R., Williamson, D.J.G., Adams, R.L., & Iverson, G.L. (1997). Oklahoma Premorbid Intelligence Estimation (OPIE): Utilization in clinical samples. The Clinical Neuropsychologist, 11, 146-154.

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.

Simon, E., Leach, L., Winocur, G., & Moscovitch, M. (1994). Intact primary memory in mild to moderate Alzheimer disease: Indices from the California Verbal Learning Test. Journal of Clinical and Experimental Neuropsychology, 16, 414-422.

Small, B.J., Herlitz, A., Fratiglioni, L., Almkvist, O., & Bäckman, L. (1997). Cognitive predictors of incident Alzheimer's disease: A prospective longitudinal study. Neuropsychology, 11, 413-420.

Smith, G.E., Bohac, D.L., Ivnik, R.J., & Malec, J.F. (1997). Using word recognition tests to estimate premorbid IQ

in early dementia: Longitudinal data. Journal of the International Neuropsychological Society, 3, 528-533.

Snow, W.G., Tierney, M.C., Zorzitto, M.L., Fisher, R.H., & Reid, D.W. (1989). WAIS-R test-retest reliability in a normal elderly sample. Journal of Clinical and Experimental Neuropsychology, 11, 423-428.

Spinnler, H., Della Sala, S., Bandera, R., & Baddeley, A.D. (1988). Dementia, aging and the structure of human memory. Cognitive Neuropsychology, 5, 193-211.

Spreen, O., & Benton, A.L. (1969). Neurosensory Center Comprehensive Examination for Aphasia. Victoria, BC: University of Victoria Neuropsychology Laboratory.

Spreen, O., & Strauss, E. (1998). A Compendium of Neuropsychological Tests (2nd Edition): Administration, Norms and Commentary. New York: Oxford University Press.

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.

Stevenson, J.D. (1986). Alternate form reliability and concurrent validity of the PPVT-R for referred

rehabilitation agency adults. Journal of Clinical Psychology, 42, 650-653.

Storandt, M., Botwinick, J., & Danzinger, W.L. (1986). Longitudinal changes: Patients with mild SDAT and matched healthy controls. In: L.W. Poon (Ed.). Clinical memory assessment of older adults. Washington, DC: American Psychological Association.

Storandt, M., Botwinick, J., Danzinger, W.L., Berg, L., & Hughes, C.P. (1984). Psychometric differentiation of mild senile dementia of the Alzheimer type. Archives of Neurology, 41, 497-499.

Storandt, M., & Hill, R.D. (1989). Very mild senile dementia of the Alzheimer type. Archives of Neurology, 46, 383-386.

Storandt, M., Stone, K., & LaBarge, E. (1995). Deficits in reading performance in very mild dementia of the Alzheimer's type. Neuropsychology, 9, 174-176.

Sullivan, E.V., Sagar, H.J., Gabrieli, J.D.E., et al., (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.

Sweet, J.J., Moberg, P.J., & Tovian, S.M. (1990). Evaluation of Wechsler Adult Intelligence Scale-Revised premorbid IQ formulas in clinical populations. Psychological Assessment, 2, 41-44.

Tabachnick, B.G., & Fidell, L.S. (1989). Using Multivariate Statistics (Second Edition). New York, NY: HarperCollins Publishers Inc.

Taylor, K.I., Salmon, D.P., Rice, V.A., Bondi, M.W., Hill, L.R., Ernesto, C.R., & Butters, N. (1996). Longitudinal examination of American National Adult REading Test (AMNART) performance in dementia of the Alzheimer type (DAT): Validation and correction based on degree of cognitive decline. Journal of Clinical and Experimental Neuropsychology, 18, 883-891.

Thompson, L.L., & Heaton, R.K. (1989). Comparison of different versions of the Boston Naming Test. Clinical Neuropsychologist, 3, 184-192.

Tombaugh, T.N., & Hubley, A.M. (1997). The 60-item Boston Naming Test: Norms for cognitively intact adults aged 25 to 88 years. Journal of Clinical and Experimental Neuropsychology, 19, 922-932.

Trahan, E.E., & Quintana, J.W. (1990). Analysis of gender effects upon verbal and visual memory performance in adults. Archives of Clinical Neuropsychology, 5, 325-334.

Tulsky, D.S., Zhu, J., & Prifitera, A. (1996). An introduction to the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III). Paper presented at the 104th Annual Convention of the American Psychological Association, Toronto, Canada.

Tulving, E., & Colotla, V.A. (1970). Free recall of trilingual lists. Cognitive Psychology, 1, 86-98.

Tulving, E., & Patterson, R.D. (1968). Functional units and retrieval processes in free recall. Journal of Experimental Psychology, 77, 239-248.

Tuokko, H., Vernon-Wilkinson, R., Weir, J., & Beattie, B.L. (1991). Cued recall and early identification of dementia. Journal of Clinical and Experimental Neuropsychology, 13, 871-879.

Unverzagt, F.W., Hui, S.L., Farlow, M.R., Hall, K.S., & Hendrie, H.C. (1998). Cognitive decline and education in mild dementia. Neurology, 50, 181-185.

Vanderploeg, R.D., & Schinka, J.A. (1995). Predicting WAIS-R IQ premorbid ability: Combining subtest performance and demographic variable predictors. Archives of Clinical Neuropsychology, 10, 225-239.

Vanderploeg, R.D., Schinka, J.A., & Axelrod, B.N. (1996). Estimation of WAIS-R premorbid intelligence: Current ability and demographic data used in a best-performance fashion. Psychological Assessment, 8 404-411.

Van Gorp, W.G., Satz, P., Kiersch, M.E., & Henry, R. (1986). Normative data on the Boston Naming Test for a group of normal older adults. Journal of Clinical and Experimental Neuropsychology, 8, 702-705.

Waugh, N.C., & Norman, D.A. (1965). Primary memory. Psychological Review, 72, 89-104.

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. (1945). A standardized memory scale for clinical use. Journal of Psychology, 19, 87-95.

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

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

Wechsler, D. (1987). Wechsler Memory Scale-Revised. New York: Psychological Corporation.

Wechsler, D. (1997). Wechsler Adult Intelligence Scale-Third Edition (WAIS-3) Wechsler Memory Scale-Third Edition (WMS-3) Technical Manual. San Antonio: The Psychological Corporation.

Welch, L.W., Doineau, D., Johnson, S., & King, D. (1996). Educational and gender normative data for the Boston Naming Test in a group of older adults. Brain and Language, 53, 260-266.

Welsh, K., Butters, N., Hughes, J., Mohs, R., & Heyman, A. (1991). Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. Archives of Neurology, 48,

278-282.

Welsh, K.A., Butters, N., Hughes, J.P., Mohs, R.C., & Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease. Archives of Neurology, 49, 448-452.

Welsh, K.A., Butters, N., Mohs, R.C., Beekly, D., Edland, S., Fillenbaum, G., & Heyman, A. (1994). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): Part V. A normative study of the neuropsychological battery. Neurology, 44, 609-614.

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

Wiederholt, W.C., Cahn, D., Butters, N., Salmon, D.P., Kritz-Silverstein, D., & Barrett-Connor, E. (1993). Effects of age, gender and education on selected neuropsychological tests in an elderly community cohort. Journal of the American Geriatric Society, 41, 639-647.

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.

Wilkinson, G.S. (1993). WRAT-3. Wilmington, DE: Wide Range Inc.

Williams, J.M. (1992). The Memory Assessment Scales. Lutz,

FL: Psychological Assessment Resources.

Williams, J.M. (1997). The prediction of premorbid memory ability. Archives of Clinical Neuropsychology, 12, 745-756.

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., Bacon, L.D., Fox, J.H., & Kaszniak, A.W. (1983). Primary memory and secondary memory in dementia of the Alzheimer's type. Journal of Clinical Neuropsychology, 5, 337-344.

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.

Wrobel, N.H., & Wrobel, T.A. (1996). The problem of assessing brain damage in psychiatric samples: Use of personality variables in prediction of WAIS-R scores. Archives of Clinical Neuropsychology, 11, 625-635.

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 A

WAIS-R VIQ Regression Equations

The following WAIS-R VIQ regression equations were considered in the estimation of intellectual, verbal memory, and language performance in the normative sample:

1. ***Carswell et al. (1997) equation[1]:***
   estimated WAIS-R VIQ = 97.066 + 2.101 (age-scaled WAIS-R Vocabulary score) - 0.661 (NART errors)
2. ***Carswell et al. (1997) equation[2]:***
   estimated WAIS-R VIQ = 130.895 - 0.972 (NART errors)
3. ***Krull et al. (1995) equation:***
   estimated WAIS-R VIQ = 65.87 + 0.87 (education) - 1.53 (race) - 0.50 (occupation) + 0.79 (WAIS-R Vocabulary raw score)
4. ***Blair and Spreen (1989) equation:***
   estimated WAIS-R VIQ = 128.7 - 0.89 (NART-R) errors
5. ***Ryan and Paolo (1992) equation:***
   estimated WAIS-R VIQ = 132.3893 - 1.164 (NART errors)
6. ***Nelson and Willison (1991) equation:***
   estimated WAIS-R VIQ = 127.4 - 1.14 (NART errors)
7. ***Barona et al. (1984) equation:***
   estimated WAIS-R VIQ = 54.23 + 0.49 (age) + 1.92 (sex) + 4.24 (race) + 5.25 (education) + 1.89 (occupation) + 1.24 (urban - rural residence)

A modification for age, as described by Helmes (1996) was applied to the Barona et al. (1984) formula to allow for the inclusion of very old elderly individuals (74 years+) in the study.

Appendix B

WAIS-R VIQ Regression Equations for Independent VIQ Predictors

1) estimated WAIS-R VIQ = 65.739 + 3.765 (**VAS**)
2) estimated WAIS-R VIQ = 123.687 - 0.658 (**NART-R errors**)
3) estimated WAIS-R VIQ = 124.234 - 0.856 (**NART errors**)
4) estimated WAIS-R VIQ = 44.246 + 1.277 (**STW**)
5) estimated WAIS-R VIQ = 41.502 + 1.391 (**WRAT-3 Reading raw score**)
6) estimated WAIS-R VIQ = 112.210 + 4.346 (**education, point biserial; ed $\geq$ 13 = +1, ed < 13 = -1**)

Note. VAS = WAIS-R Vocabulary age-scaled score; NART-R errors = New Adult Reading Test-Revised errors; NART errors = National Adult Reading Test errors; STW = Spot-the-Word subtest raw score; WRAT-3 Reading raw score = Wide Range Achievement Test-3rd Edition raw score; ed = education.

Appendix C

CVLTLDFR Regression Equations for Independent Predictors

1) estimated CVLTLDFR = 4.570 + 0.476 (**CVLT Recency raw score**)
2) estimated CVLTLDFR = 19.721 - 0.120 (**Age**)
3) estimated CVLTLDFR = 8.304 + 0.311 (**WAIS-R Digits Forward raw score**)
4) estimated CVLTLDFR = 12.881 - 1.466 (**Sex**)

Appendix D

BNT Regression Equations for Independent Predictors

1) estimated BNT = 40.272 + 0.291 (**WAIS-R Vocabulary raw score**)
2) estimated BNT = 6.261 + 0.297 (**PPVT-R raw score**)
3) estimated BNT = 33.668 + 0. 423 (**STW raw score**)

Appendix E

Raw Data Set

Par.#	Age	Sex	Ed	MMSE	GDI	NRTER	NRTIQ
1	80	1	14	30	2	23	101
2	68	1	16	30	3	4	123
3	53	1	22	30	2	4	123
4	69	1	12	30	3	7	119
5	68	1	15	30	1	11	115
6	80	1	13	29	2	10	116
7	82	1	11	26	4	10	116
8	74	2	7	28	6	10	116
9	79	1	12	29	1	17	108
10	83	1	15	29	5	9	117
11	73	1	13	28	4	11	115
12	74	1	13	30	4	6	121
13	63	1	11	30	0	12	114
14	80	1	18	30	1	2	125
15	83	2	8	26	10	11	115

16	74	1	12	29	3	33	90
17	88	2	12	30	4	10	116
18	64	2	19	30	3	9	117
19	68	1	15	29	2	6	121
20	74	1	14	29	1	7	119
21	70	1	17	30	5	11	115
22	74	1	14	29	5	8	118
23	65	1	17	30	1	7	119
24	72	1	17	30	9	5	122
25	77	1	10	29	1	18	107
26	67	1	12	30	1	10	116
27	67	1	12	30	4	13	113
28	69	1	11	29	1	27	97
29	75	2	10	29	3	32	91
30	74	1	9	30	0	18	107
31	78	1	14	27	1	21	103
32	75	1	6	27	7	23	101
33	74	2	12	28	7	21	103
34	79	2	15	30	3	8	118

35	76	1	17	29	3	15	110
36	74	1	16	30	2	8	118
37	78	1	8	29	3	19	106
38	83	1	18	30	1	13	113
39	58	1	13	26	1	10	116
40	82	1	15	27	1	11	115
41	75	2	10	29	4	23	101
42	66	1	17	29	8	10	116
43	70	1	10	29	7	11	115
44	63	1	14	30	6	5	122
45	72	1	10	27	2	30	93
46	76	1	12	28	9	19	106
47	57	1	16	29	5	12	114
48	72	1	12	27	6	18	107
49	77	1	18	27	2	16	109
50	57	1	10	28	6	16	109
51	68	1	10	30	2	11	115
52	66	2	11	29	12	5	122
53	68	1	14	30	11	14	111

54	74	2	18	30	0	11	115
55	71	1	11	30	5	9	117
56	71	1	17	29	2	6	121
57	63	1	12	30	8	10	116
58	73	2	12	29	12	15	110
59	67	1	9	28	1	23	101
60	70	1	15	29	9	5	122
61	66	1	12	28	3	28	95
62	74	1	14	29	2	6	121
63	87	1	11	27	10	10	116
64	71	1	12	26	1	12	114
65	67	1	19	29	8	19	106
66	67	1	18	29	7	8	118
67	63	1	16	30	6	9	117
68	87	1	11	28	8	21	103
69	66	1	13	27	2	15	110
70	66	1	15	29	2	3	124
71	68	2	19	26	2	9	117
72	72	1	12	29	3	18	107

73	73	1	10	30	1	22	102
74	89	1	13	28	4	13	113
75	66	1	18	30	4	20	105
76	78	2	15	29	1	16	109
77	69	1	12	29	9	20	105
78	75	1	12	28	4	15	110
79	72	2	17	30	7	10	116
80	67	1	11	29	6	26	98
81	64	1	16	29	6	18	107
82	73	2	18	30	1	14	111
83	73	2	17	30	1	7	119
84	71	1	13	30	1	7	119
85	82	2	18	26	1	9	117
86	70	2	14	30	3	9	117
87	69	1	15	29	6	21	103
88	89	1	14	28	2	18	107
89	72	2	12	29	11	17	108
90	76	1	11	26	4	14	111
91	72	1	12	29	3	13	113

92	58	1	10	28	3	33	90
93	79	2	12	29	3	15	110
94	65	1	12	28	2	14	111
95	71	1	17	30	3	7	119
96	69	1	19	27	1	7	119
97	61	1	14	30	1	16	109
98	77	2	17	29	1	8	118

Note. Definitions: Par.# = Participant number; Ed = Education; MMSE = Mini-Mental State Examination; GDI = Geriatric Depression Index; NRTER = National Adult Reading Test (NART) errors; NRTIQ = NART Verbal IQ. Sex was dummy variable coded with females = 1 and males = 2.

Raw Data Continued:

Par.#	BDAEC	BDAEL	BNT	WRAT3	SENTR	STW	PPVTR
1	10	30.76	54	43	15	45	153
2	9	41.07	60	55	22	57	173
3	10	22.91	57	57	20	55	172
4	10	27.26	55	53	20	54	174
5	10	26.77	59	51	20	50	167

6	9	50.59	60	50	20	54	173
7	10	39.03	55	55	22	56	168
8	10	42.86	58	55	22	58	170
9	10	44.93	57	48	16	49	166
10	10	28.18	60	53	19	59	167
11	10	29.58	56	51	23	53	170
12	10	39.16	59	51	19	58	171
13	10	26.22	57	53	18	53	171
14	10	29.06	59	56	21	59	172
15	10	32.88	56	48	18	58	170
16	10	42.46	54	38	16	38	148
17	9	27.07	53	50	17	55	169
18	10	33.75	58	54	21	57	172
19	10	28.73	54	52	18	54	171
20	10	28.50	58	54	19	58	173
21	8	85.51	55	50	17	48	163
22	10	36.67	59	54	18	59	172
23	10	30.22	56	54	18	49	169
24	10	27.02	59	55	20	59	174

25	9	29.22	59	49	20	55	171
26	9	31.52	60	53	18	57	170
27	10	23.06	54	52	18	54	169
28	10	24.72	53	40	20	47	152
29	10	41.37	60	41	23	48	165
30	10	34.79	51	47	18	46	151
31	10	28.47	42	52	19	45	152
32	10	30.81	55	47	18	48	161
33	9	31.88	54	50	16	53	151
34	9	47.39	57	53	18	57	171
35	10	24.76	59	52	20	58	173
36	9	27.88	59	54	22	58	168
37	9	42.38	58	49	22	53	172
38	10	25.47	52	53	19	54	173
39	10	25.41	59	54	20	53	172
40	10	30.74	53	54	18	56	171
41	9	38.33	58	44	20	57	169
42	9	31.62	56	53	17	56	173
43	10	26.44	58	51	24	57	174

44	9	39.13	60	55	19	60	172
45	9	27.35	50	45	18	41	147
46	10	27.93	52	51	17	45	166
47	10	33.53	56	52	19	58	173
48	10	25.42	54	51	16	53	173
49	9	26.70	60	50	23	58	170
50	10	20.96	58	48	20	54	170
51	9	36.36	55	51	19	56	171
52	10	23.89	59	54	25	57	172
53	9	28.08	54	52	17	52	172
54	10	32.09	56	52	20	52	175
55	10	30.46	56	55	18	56	172
56	9	29.65	59	54	18	59	174
57	10	31.20	59	53	17	57	174
58	9	26.51	58	52	20	57	173
59	10	38.89	54	47	18	41	167
60	10	29.36	60	54	23	59	174
61	10	47.25	49	45	16	46	146
62	10	22.21	59	53	18	60	175

63	8	38.59	49	53	21	52	167
64	9	30.04	53	51	20	49	169
65	10	31.17	55	53	19	52	173
66	10	26.78	56	55	19	56	174
67	10	24.51	59	55	18	55	175
68	9	43.45	58	51	19	56	170
69	10	19.03	60	55	19	56	174
70	10	19.41	60	55	23	57	174
71	10	34.79	58	54	18	55	171
72	10	22.86	56	48	21	56	173
73	10	28.91	60	47	19	48	169
74	10	43.93	49	54	18	55	173
75	10	22.90	58	55	19	56	171
76	9	31.19	55	50	17	56	169
77	10	23.69	56	50	17	53	171
78	10	24.03	56	51	17	53	172
79	10	25.34	60	52	19	56	174
80	10	29.67	49	42	19	41	154
81	10	23.97	56	51	17	45	168

82	10	38.81	60	49	19	56	169
83	9	36.99	58	52	21	57	173
84	9	25.38	57	55	20	59	173
85	10	26.75	58	54	20	58	172
86	9	28.08	55	54	25	58	171
87	10	19.80	59	48	17	53	173
88	9	27.35	58	50	19	52	172
89	10	25.02	58	52	17	51	168
90	10	35.87	57	52	21	53	170
91	10	22.34	52	51	20	51	169
92	10	36.31	48	44	17	40	144
93	10	36.45	50	49	21	51	163
94	9	27.11	59	50	24	51	161
95	10	19.99	59	51	18	55	169
96	10	24.05	59	55	21	58	171
97	10	19.43	57	49	18	52	169
98	9	21.71	60	54	21	59	173

Note. Definitions: BDAEC = number correct on the Boston Diagnostic Aphasia Examination (BDAE) Word-Picture Matching

subtest; BDAEL = Latency in seconds to complete BDAE Word-Picture Matching subtest; BNT = Boston Naming Test total correct score; WRAT3 = Wide Range Achievement Test-3 Reading subtest score; SENTR = Sentence Repetition corrected score; STW = Spot-the-Word subtest raw score; PPVTR = Peabody Picture Vocabulary Test-Revised raw score.

Raw Data Continued:

Par.#	NR2ER	NR2IQ	BREQ	CRSEQ	KREQ	WRVIQ	VRAW
1	26	105.6	114.3	100.7	99.8	110	40
2	7	122.5	118.1	123.8	118.8	112	63
3	5	124.2	117.1	128.0	121.2	130	66
4	6	123.4	105.7	119.7	112.6	110	58
5	17	113.6	112.9	119.2	117.2	124	62
6	9	120.7	114.3	119.8	116.4	113	61
7	11	118.9	101.9	119.8	113.4	114	60
8	15	115.3	90.5	119.8	111.7	120	62
9	20	110.9	102.9	108.9	105.3	102	50
10	16	114.5	114.3	118.4	114.0	115	58
11	18	112.7	111.5	119.2	115.1	125	60
12	8	121.6	113.4	128.8	121.1	130	67

13	17	113.6	99.9	110.1	103.1	108	47
14	2	126.9	119.6	129.3	121.2	132	66
15	23	108.2	96.7	112.9	101.5	103	48
16	44	89.5	106.2	98.3	103.2	96	46
17	17	113.6	109.6	115.6	109.5	111	54
18	13	117.1	119.5	120.5	117.2	116	61
19	9	120.7	112.9	118.3	110.1	106	53
20	12	118.0	113.3	119.7	114.0	117	58
21	16	114.5	118.6	117.1	115.7	108	59
22	8	121.6	113.4	121.2	115.6	117	60
23	11	118.9	112.4	117.6	112.6	110	57
24	5	124.3	118.6	123.2	118.8	124	63
25	26	105.6	101.4	114.6	113.3	119	60
26	11	118.9	103.8	115.6	110.1	107	56
27	19	111.8	105.7	113.7	110.3	111	55
28	39	93.9	100.5	100.2	99.1	102	42
29	39	93.9	105.3	101.1	107.5	117	52
30	25	106.5	100.9	106.2	100.7	94	44
31	27	104.7	113.8	106.3	106.1	104	48

32	34	98.4	89.1	100.7	94.5	105	39
33	37	95.8	110.0	99.9	94.2	103	34
34	14	116.2	115.8	119.1	113.2	116	57
35	19	111.8	119.1	116.5	116.5	113	60
36	11	118.9	118.6	116.9	113.3	116	56
37	24	107.3	96.2	107.6	104.6	104	50
38	13	117.1	119.6	115.8	114.1	118	57
39	10	119.8	112.4	117.7	114.8	104	59
40	12	118.0	112.4	115.0	107.9	116	51
41	33	99.3	101.5	109.2	109.5	109	57
42	7	122.5	118.1	119.8	118.0	119	62
43	9	120.7	100.9	115.0	109.4	117	55
44	6	123.4	112.4	118.9	112.4	120	56
45	36	96.7	99.1	94.0	91.0	95	33
46	21	110.1	108.6	107.6	106.8	104	50
47	11	118.9	117.6	118.5	117.3	119	61
48	20	110.9	108.1	112.5	112.4	106	57
49	17	113.6	119.1	113.8	114.1	111	57
50	23	108.2	96.2	111.7	107.6	109	54

51	18	112.7	98.6	115.0	106.8	105	53
52	6	123.4	104.3	121.0	113.1	113	59
53	16	114.5	112.8	110.9	108.5	108	51
54	12	118.0	120.5	117.1	115.7	125	59
55	10	119.8	97.2	116.3	108.4	110	55
56	5	124.3	118.6	122.5	117.3	117	61
57	12	118.0	105.2	113.5	107.1	122	51
58	16	114.5	110.0	120.7	118.7	119	65
59	29	102.9	98.6	109.2	111.5	102	59
60	5	124.3	113.4	123.2	117.2	122	62
61	39	93.9	101.9	97.5	96.6	84	39
62	3	126.0	111.5	124.6	118.3	126	64
63	13	117.1	102.4	115.6	109.4	117	55
64	13	117.1	106.2	112.2	106.3	113	50
65	19	111.8	118.1	109.7	114.1	106	57
66	7	122.5	118.1	121.2	117.3	117	61
67	9	120.7	117.6	120.5	116.5	114	60
68	24	107.3	102.4	108.4	108.6	104	54
69	17	113.6	112.9	114.5	114.8	116	59

70	6	123.4	112.9	126.6	118.7	121	64
71	12	118.0	120.0	118.4	114.9	106	58
72	20	110.9	106.2	110.4	110.3	120	55
73	27	104.7	97.2	107.7	107.6	106	54
74	18	112.7	112.9	109.5	103.2	108	45
75	15	115.4	118.1	113.2	118.0	121	62
76	18	112.7	115.8	111.7	112.4	112	56
77	19	111.8	105.7	113.2	114.2	112	60
78	20	110.9	106.7	108.2	101.6	106	44
79	11	118.9	120.5	119.8	117.3	117	61
80	36	96.7	100.5	100.9	98.3	99	41
81	24	107.3	117.6	110.4	111.7	97	54
82	16	114.5	120.5	115.1	115.7	128	59
83	6	123.4	120.5	117.6	112.5	114	55
84	5	124.3	111.5	119.7	113.5	128	58
85	9	120.7	121.5	120.5	116.5	121	60
86	8	121.6	115.3	120.5	115.6	126	60
87	28	103.8	110.9	112.6	115.9	115	61
88	16	114.5	109.2	112.5	111.7	117	57

89	18	112.7	106.2	108.9	105.0	117	49
90	15	115.4	101.5	113.0	110.2	111	56
91	15	115.4	106.2	111.6	106.3	106	50
92	46	87.8	99.9	89.9	90.4	82	31
93	23	108.2	108.6	110.3	104.7	118	48
94	13	117.1	107.6	117.2	116.3	114	62
95	12	118.0	118.6	119.7	114.1	120	57
96	7	122.5	118.1	119.7	114.9	117	58
97	22	109.1	112.4	107.5	106.1	110	48
98	8	121.6	121.0	121.2	118.0	120	62

Note. Definitions: NR2ER = New Adult Reading Test-Revised (NART-R) errors; NR2IQ = NART-R VIQ; BREQ = Barona et al., (1984) Demographics equation with Helmes (1996) age modification; CRSEQ = Carswell et al., (1997) WAIS-R Vocabulary/NART equation; KREQ = Krull et al., (1995) WAIS-R Vocabulary/Demographics equation; WRVIQ = WAIS-R VIQ; VRAW = WAIS-R Vocabulary subtest raw score.

Raw Data Continued:

Par.#	VAS	DSF	DSAS	INFAS	ARTAS	CPAS	SIMAS
1	9	9	15	10	12	10	14
2	14	6	8	12	9	12	16
3	16	10	11	14	15	14	16
4	13	10	14	10	12	10	13
5	14	10	14	14	16	13	13
6	14	7	9	13	12	11	14
7	14	8	11	13	10	13	14
8	14	11	13	13	14	12	15
9	11	5	8	11	8	11	12
10	13	12	16	14	8	13	13
11	14	13	19	13	11	14	15
12	17	9	13	15	14	15	16
13	10	8	9	11	13	13	13
14	16	10	14	15	13	14	18
15	11	5	6	14	9	9	14
16	11	5	7	9	10	9	12
17	12	9	11	10	15	13	10

18	14	12	13	15	10	12	11
19	12	8	10	12	10	11	12
20	13	7	9	14	15	12	14
21	13	8	10	13	8	12	12
22	14	8	11	14	11	13	14
23	12	8	12	12	13	9	12
24	14	11	15	13	13	13	15
25	14	12	14	14	13	13	13
26	12	7	10	9	14	12	11
27	12	9	14	9	13	12	12
28	10	8	12	9	12	10	12
29	12	11	15	12	13	13	14
30	10	6	8	10	6	10	12
31	11	10	16	9	8	9	11
32	9	8	11	10	7	14	13
33	8	7	9	12	11	11	12
34	13	7	10	13	13	11	15
35	14	11	15	11	10	11	14
36	12	7	13	11	13	13	14

37	11	6	9	12	9	10	14
38	13	8	11	13	14	13	14
39	13	11	11	14	7	8	11
40	12	11	14	13	13	13	13
41	13	5	9	14	13	10	11
42	14	7	10	16	14	12	13
43	12	8	11	14	12	13	14
44	12	12	14	14	14	11	14
45	8	9	12	8	10	9	11
46	11	9	11	10	9	11	12
47	14	12	13	15	13	10	13
48	13	9	11	10	9	10	12
49	13	10	14	13	8	11	12
50	12	10	11	12	12	11	12
51	12	10	11	14	9	9	11
52	13	11	12	10	12	13	13
53	11	9	11	11	13	11	12
54	13	13	19	14	15	12	14
55	12	4	7	14	11	12	14

56	14	6	9	13	14	12	15
57	11	10	14	16	15	10	16
58	16	9	13	14	13	12	13
59	13	8	11	10	6	12	11
60	14	12	14	14	14	15	13
61	9	6	8	5	6	7	9
62	15	11	14	16	14	13	15
63	12	10	11	12	15	14	13
64	11	12	13	10	13	12	14
65	12	9	11	12	10	11	10
66	14	10	13	14	13	12	11
67	14	10	11	13	13	11	11
68	12	10	12	10	8	10	12
69	13	10	14	13	14	11	12
70	15	11	14	14	13	13	12
71	13	6	8	13	10	12	12
72	12	10	15	15	13	13	13
73	12	6	9	13	11	9	12
74	10	9	10	11	12	11	13

75	14	10	14	14	15	12	13
76	12	6	9	13	11	13	14
77	14	9	11	12	12	11	12
78	10	10	13	12	9	10	14
79	14	9	11	14	15	12	12
80	10	7	8	11	14	6	12
81	12	5	6	11	9	9	12
82	13	10	14	15	16	14	16
83	12	9	11	14	13	11	13
84	13	12	17	16	17	11	15
85	14	11	15	14	15	12	13
86	14	13	19	11	17	10	14
87	14	8	10	12	15	11	13
88	13	12	15	14	11	10	14
89	11	9	14	16	11	10	14
90	12	11	14	14	8	11	12
91	11	9	11	10	8	10	14
92	7	5	8	6	8	6	7
93	11	8	11	15	15	14	13

94	14	10	12	13	15	11	11
95	13	11	16	14	14	11	14
96	13	10	13	14	14	11	12
97	10	11	12	13	16	9	11
98	14	10	14	16	11	12	14

Note. Definitions: VAS = WAIS-R Vocabulary subtest age-scaled acore; DSF = WAIS-R Digit Span Forward raw score; DSAS = WAIS-R Digit Span age-scaled score; INFAS = WAIS-R Information subtest age-scaled score; ARTAS = WAIS-R Arithmetic subtest age-scaled score; CPAS = WAIS-R Comprehension subtest age-scaled score; SIMAS = WAIS-R Similarities subtest age-scaled score.

Raw Data Continued:

Par.#	CV15	CVREC	CVLDF	CVLDC	BREQ2	RYEQ	CREQ2
1	57	11	9	11	113.3	105.6	108.5
2	62	15	12	14	118.1	127.7	127.0
3	47	16	9	11	117.1	127.7	127.0
4	60	15	13	13	105.7	124.2	124.1
5	59	15	16	15	112.9	119.6	120.2

6	35	12	8	10	113.3	120.7	121.2
7	55	16	10	12	100.9	120.7	121.2
8	31	8	7	6	90.5	120.7	121.2
9	41	11	6	8	102.4	112.6	114.4
10	48	13	10	11	113.3	121.9	122.2
11	52	14	9	9	111.5	119.6	120.2
12	53	17	13	13	113.4	125.4	125.1
13	54	15	12	13	99.9	118.4	119.2
14	51	17	15	13	118.6	130.1	128.9
15	32	9	9	8	95.7	119.6	120.2
16	60	17	13	13	106.2	93.9	98.8
17	28	5	7	8	108.1	120.8	121.2
18	59	19	11	11	119.5	121.9	122.2
19	45	12	11	12	112.9	125.4	125.1
20	58	14	11	13	113.3	124.2	124.1
21	41	12	9	9	118.6	119.6	120.2
22	62	16	15	15	113.4	123.1	123.1
23	57	15	10	10	112.4	124.2	124.1
24	59	13	14	14	118.6	126.6	126.0

25	48	12	13	14	100.9	111.4	113.4
26	54	17	11	12	103.8	120.7	121.2
27	68	17	15	16	105.7	117.3	118.3
28	50	14	13	11	100.5	100.9	104.7
29	70	17	13	13	104.8	95.1	99.8
30	56	16	9	8	100.9	111.4	113.4
31	47	14	11	11	113.3	107.9	110.5
32	37	12	8	8	88.6	105.6	108.5
33	37	14	10	10	110.0	107.9	110.5
34	42	15	8	9	115.3	123.1	123.1
35	55	15	8	9	118.6	114.9	116.3
36	62	18	14	14	118.6	123.1	123.1
37	33	10	8	7	95.7	110.3	112.4
38	57	13	12	13	118.6	117.3	118.3
39	40	7	6	7	112.4	120.7	121.2
40	46	13	9	11	111.4	119.6	120.2
41	40	8	7	10	101.0	105.6	108.5
42	42	17	7	7	118.1	120.7	121.2
43	46	13	11	13	100.9	119.6	120.2

44	70	18	16	16	112.4	126.6	126.0
45	48	13	11	11	99.1	97.5	101.7
46	36	10	8	8	108.1	110.3	112.4
47	54	13	10	10	117.6	118.4	119.2
48	44	5	8	9	108.1	111.4	113.4
49	41	10	9	8	118.6	113.8	115.3
50	66	17	14	14	96.2	113.8	115.3
51	54	10	13	13	98.6	119.6	120.2
52	40	14	9	8	104.3	126.6	126.0
53	54	10	12	12	112.8	116.1	117.3
54	59	15	12	12	120.5	119.6	120.2
55	45	12	9	9	97.2	121.9	122.2
56	61	15	14	15	118.6	125.4	125.1
57	63	13	16	16	105.2	120.7	121.2
58	54	14	12	13	110.0	114.9	116.3
59	37	15	7	7	98.6	105.6	108.5
60	56	12	10	12	113.4	126.6	126.0
61	60	15	13	13	101.9	99.8	103.7
62	63	15	12	11	111.5	125.4	125.1

63	41	15	9	9	100.9	120.7	121.2
64	72	20	16	16	106.2	118.4	119.2
65	48	16	10	11	118.1	110.3	112.4
66	52	15	12	11	118.1	123.1	123.1
67	57	14	13	14	117.6	121.9	122.2
68	39	11	9	10	100.9	107.9	110.5
69	62	17	16	15	112.9	114.9	116.3
70	68	18	15	15	112.9	128.9	127.9
71	44	18	9	8	120.0	121.9	122.2
72	65	16	14	14	106.2	111.4	113.4
73	63	16	15	15	97.2	106.8	109.5
74	54	16	11	11	111.4	117.3	118.3
75	67	18	16	16	118.1	109.1	111.5
76	49	15	10	11	115.3	113.8	115.3
77	53	15	11	11	105.7	109.1	111.5
78	56	17	13	13	106.2	114.9	116.3
79	43	10	9	11	120.5	120.7	121.2
80	61	13	16	16	100.5	102.1	105.6
81	55	13	9	8	117.6	111.4	113.4

82	58	12	14	14	120.5	116.1	117.3
83	50	13	10	10	120.5	124.2	124.1
84	58	13	13	14	111.5	124.2	124.1
85	23	11	0	1	120.5	121.9	122.2
86	55	10	15	15	115.3	121.9	122.2
87	57	15	13	14	110.9	107.9	110.5
88	53	17	9	11	107.7	111.4	113.4
89	39	11	12	12	106.2	112.6	114.4
90	34	6	6	6	101.0	116.1	117.3
91	49	15	6	7	106.2	117.3	118.3
92	54	15	9	8	99.9	93.9	98.8
93	37	10	6	7	108.1	114.9	116.3
94	59	14	13	11	107.6	116.1	117.3
95	64	15	13	13	118.6	124.2	124.1
96	60	13	14	14	118.1	124.2	124.1
97	64	19	16	16	112.4	113.8	115.3
98	43	7	9	11	120.5	123.1	123.1

Note. Definitions: CV15 = California Verbal Learning Test (CV) Trials 1-5 raw score; CVREC = CV Recency raw score;

CVLDF = CV long delay free recall raw score; CVLDC = CV long delay cued recall raw score; BREQ2 = Barona et al., (1984) Demographics equation; RYEQ = Ryan et al., (1992) NART equation; CREQ2 = Carswell et al., (1997) NART equation.

Raw Data Continued:

Par.#	CVA1	CVA5	CVB	CVSDF	CVSDC	CVSEM	CVSER
1	5	13	6	9	8	1.9	2.6
2	9	14	9	14	13	2.8	2.8
3	6	11	5	10	10	1.6	5.0
4	7	16	9	13	12	2.3	1.2
5	6	16	8	16	16	1.8	1.7
6	5	9	4	7	10	1.6	5.8
7	8	14	4	8	11	2.2	2.2
8	4	10	3	6	10	1.9	3.6
9	3	11	6	9	6	1.9	1.1
10	5	12	8	10	10	0.8	5.9
11	8	13	6	9	10	0.8	1.4
12	5	13	3	13	13	2.6	1.8
13	6	13	7	14	14	3.3	1.4

14	6	16	5	12	13	2.8	2.4
15	4	8	3	7	8	0.5	2.1
16	6	14	10	14	10	2.7	0.4
17	5	7	4	6	8	0.7	7.1
18	9	12	7	12	10	1.6	2.5
19	5	11	6	11	12	1.8	2.1
20	7	12	8	13	13	1.3	2.1
21	5	10	7	7	9	1.4	0.0
22	6	14	4	12	14	3.5	0.8
23	8	15	4	11	13	2.1	1.7
24	6	15	7	12	14	2.5	3.0
25	6	13	6	11	15	1.6	4.0
26	8	12	6	11	10	1.5	1.3
27	8	16	5	15	16	2.0	1.9
28	6	12	9	11	13	1.0	2.4
29	9	16	10	12	13	2.6	1.8
30	7	14	7	10	8	2.0	2.6
31	4	11	6	8	11	2.0	1.5
32	5	11	2	8	8	1.3	3.7

33	5	9	3	8	8	0.5	3.7
34	6	10	4	7	8	2.1	1.7
35	6	14	4	8	8	1.5	4.4
36	6	16	8	14	14	2.3	5.7
37	3	9	5	6	7	1.0	4.8
38	9	14	4	12	13	2.9	1.3
39	7	9	7	4	5	1.3	1.7
40	5	13	9	10	12	1.3	1.5
41	6	11	6	8	10	1.6	0.6
42	4	10	8	6	7	1.1	0.6
43	5	11	7	9	12	1.2	0.0
44	9	16	11	15	16	3.8	0.4
45	6	13	6	10	11	1.9	1.5
46	3	10	5	6	8	1.2	2.6
47	6	15	6	12	13	2.1	1.8
48	4	12	8	8	11	1.8	1.6
49	7	10	7	10	9	1.7	0.6
50	8	16	8	14	14	2.6	1.1
51	7	14	8	12	13	3.3	0.5

52	4	10	5	10	9	1.0	8.1
53	5	14	6	12	12	3.4	1.4
54	7	16	8	14	13	2.7	1.3
55	6	10	7	10	9	2.0	0.0
56	6	16	9	15	12	3.3	0.4
57	8	16	12	14	15	2.7	2.4
58	8	14	8	12	13	2.3	2.2
59	3	10	6	8	8	0.8	1.3
60	8	12	8	11	10	2.3	4.3
61	8	13	7	13	16	2.4	1.2
62	7	15	8	11	14	2.3	1.2
63	4	10	6	5	8	1.0	2.3
64	10	16	7	16	16	3.7	0.7
65	4	12	8	9	12	1.3	2.5
66	6	12	6	11	13	1.4	0.9
67	7	14	7	13	14	0.9	4.7
68	5	11	3	9	10	2.3	1.2
69	7	16	9	15	16	2.3	1.6
70	9	16	9	15	15	3.2	2.6

71	4	12	6	7	8	1.1	1.1
72	8	15	7	13	15	3.0	1.9
73	9	15	8	13	15	2.9	2.0
74	7	14	5	10	11	2.4	0.0
75	7	16	7	16	16	3.3	0.4
76	7	12	7	10	11	1.5	0.5
77	7	15	4	10	12	2.2	0.0
78	9	13	5	10	12	3.4	0.0
79	3	12	8	9	10	1.9	1.7
80	7	16	8	16	16	2.1	2.5
81	7	12	7	9	8	2.6	0.9
82	7	14	7	13	12	3.1	2.1
83	6	14	9	7	11	1.6	2.4
84	10	13	6	12	13	1.7	2.1
85	4	4	4	0	4	1.5	2.9
86	6	16	6	14	15	0.4	10.3
87	8	14	6	12	13	3.2	0.0
88	7	9	7	11	9	1.9	1.4
89	3	11	5	12	12	0.8	5.4

90	3	10	8	7	9	1.2	4.1
91	6	12	7	9	10	1.3	3.9
92	6	14	8	10	7	1.5	0.9
93	6	7	5	6	7	1.8	2.5
94	7	13	10	10	10	2.1	2.1
95	8	13	7	12	13	3.4	1.6
96	5	16	6	14	14	3.8	0.8
97	9	16	10	14	16	2.2	2.7
98	6	12	8	10	12	1.0	6.5

Note. Definitions: CVA1 = California Verbal Learning Test (CV) Trial A1 raw score; CVA5 = CV Trial A5 raw score; CVB = CV Trial B raw score; CVSDF = CV short delay free recall raw score; CVSDC = CV short delay cued recall raw score; CVSEM = CV Semantic Clustering score; CVSER = CV Serial Clustering score.

Raw Data Continued:

Par.#	CVPP	CVPM	CVPRE	CVSLP	CVRC	CVPER	CVFRI
1	28	53	19	2.1	86	1	2
2	29	47	24	0.9	79	1	0
3	23	43	34	1.0	83	0	1
4	33	42	25	1.7	91	2	2
5	32	42	25	2.1	93	4	2
6	31	34	34	1.0	81	2	1
7	20	51	29	1.8	88	0	2
8	32	42	26	1.3	67	1	1
9	27	46	27	2.0	80	3	8
10	33	40	27	1.4	83	0	0
11	31	42	27	1.0	87	4	0
12	19	49	32	2.0	92	1	1
13	30	43	28	1.7	85	1	1
14	24	43	33	2.6	89	1	5
15	31	41	28	0.9	67	2	2
16	28	43	28	2.0	93	1	1
17	61	21	18	0.6	86	0	3

18	24	44	32	0.8	91	2	1
19	19	44	27	1.4	85	0	1
20	33	43	24	1.0	85	3	0
21	27	44	29	1.4	81	3	0
22	29	45	26	1.7	94	0	0
23	32	42	26	1.4	83	1	1
24	25	53	22	2.2	89	0	0
25	35	40	25	1.7	74	2	1
26	22	46	31	1.0	74	2	2
27	26	49	25	1.9	96	1	0
28	28	44	28	1.3	82	1	1
29	24	51	24	1.7	98	2	5
30	29	43	29	1.5	88	1	0
31	21	49	30	1.5	75	4	2
32	30	38	32	1.8	69	1	0
33	30	32	38	1.1	68	2	2
34	21	43	36	1.1	72	0	2
35	29	44	27	1.8	85	0	1
36	29	42	29	2.5	93	4	0

37	42	27	30	1.4	83	2	1
38	28	49	23	1.1	86	2	2
39	38	45	18	0.7	71	0	2
40	35	37	28	1.8	82	1	1
41	30	50	20	1.4	66	1	4
42	19	40	40	1.3	81	0	1
43	35	37	28	1.3	89	1	0
44	29	46	26	1.7	98	0	0
45	29	44	27	1.7	77	1	5
46	31	42	28	1.5	65	0	1
47	31	44	24	2.1	82	0	0
48	39	50	11	1.8	75	0	9
49	20	56	24	0.7	68	0	5
50	27	47	26	1.9	94	2	0
51	33	48	19	1.9	92	0	0
52	38	28	35	1.1	80	0	1
53	28	54	19	2.1	95	0	0
54	24	51	25	2.0	93	0	1
55	29	44	27	1.1	77	1	1

56	28	48	25	2.8	93	0	0
57	29	51	21	2.0	91	3	0
58	33	41	26	1.4	90	0	1
59	27	32	41	1.7	85	0	4
60	34	45	21	0.8	80	0	6
61	23	52	25	0.8	83	1	0
62	25	51	24	2.0	92	0	7
63	29	34	37	1.2	77	0	0
64	26	46	28	1.4	100	0	0
65	29	38	33	1.5	75	2	3
66	23	48	29	0.9	78	1	2
67	28	47	25	2.0	91	1	6
68	15	56	28	1.4	75	0	4
69	27	45	27	2.2	91	0	0
70	26	47	26	1.6	96	2	0
71	14	45	41	1.8	84	2	4
72	31	45	25	1.4	94	0	5
73	25	49	25	1.6	88	0	1
74	28	43	30	1.6	80	0	1

75	27	46	27	2.0	98	1	0
76	35	35	31	1.0	81	2	4
77	32	40	28	2.0	95	0	0
78	32	38	30	0.9	86	0	4
79	33	44	23	2.0	81	0	0
80	33	46	21	2.3	98	1	2
81	31	45	24	1.2	86	0	13
82	24	55	21	2.0	93	0	0
83	22	52	26	2.1	72	2	0
84	28	50	22	1.2	78	0	0
85	35	17	48	0.0	63	0	1
86	29	53	18	2.3	85	0	0
87	33	40	26	1.4	93	0	4
88	25	43	32	0.8	77	0	0
89	38	33	28	1.8	89	1	1
90	50	32	18	1.8	83	0	3
91	39	31	31	1.3	86	1	1
92	28	44	28	1.8	88	1	0
93	38	35	27	0.5	77	1	3

94	29	47	24	1.3	89	0	10
95	28	48	23	1.3	88	1	5
96	30	48	22	2.6	98	0	0
97	27	44	30	1.8	90	3	0
98	42	42	16	1.3	68	3	3

Note. Definitions: CVPP = California Verbal Learning Test (CV) Percent Primacy region score; CVPM = CV Percent Middle region score; CVPRE = CV Percent Recency region score; CVSLP = CV Slope score; CVRC = CV Recall Consistency score; CVPER = CV Perseverations score; CVFRI = CV Free Recall Intrusions score.

Raw Data Continued:

Par.#	CVCRI	CVTI	CVRH	CVDSM	CVFP	CVRB
1	0	2	13	93	0	-0.5
2	0	0	16	95	2	0.3
3	2	3	14	93	1	-0.3
4	2	4	15	98	0	0.0
5	0	2	16	100	0	0.0
6	1	2	12	89	1	-0.6

7	2	4	16	100	0	0.0
8	2	3	14	95	0	-0.3
9	5	13	15	93	2	0.3
10	0	0	15	93	2	0.3
11	0	0	14	95	0	-0.3
12	1	2	14	93	1	-0.3
13	0	1	15	98	0	0.0
14	2	7	16	100	0	0.0
15	1	3	14	93	1	-0.3
16	2	3	16	100	0	0.0
17	6	9	12	84	3	-0.1
18	2	3	16	98	1	0.0
19	1	2	14	93	1	-0.3
20	0	0	15	95	1	0.0
21	0	0	12	86	2	-0.3
22	0	0	16	98	1	0.0
23	1	2	15	98	0	0.0
24	0	0	15	98	0	0.0
25	0	1	15	98	0	0.0

26	1	3	15	98	0	0.0
27	0	0	16	100	0	0.0
28	0	1	14	95	0	-0.3
29	0	5	16	98	1	0.0
30	1	1	16	98	1	0.0
31	4	6	11	89	0	-0.7
32	0	0	16	93	3	0.5
33	3	5	14	86	4	0.3
34	2	4	13	89	2	-0.2
35	1	2	16	98	1	0.0
36	1	1	13	93	0	-0.5
37	2	3	16	86	6	0.7
38	3	5	15	93	2	0.3
39	4	6	15	95	1	0.0
40	0	1	15	98	0	0.0
41	4	8	11	84	2	-0.4
42	0	1	15	98	0	0.0
43	0	0	13	93	0	-0.5
44	0	0	16	100	0	0.0

45	1	6	15	93	2	0.3
46	2	3	15	91	3	0.5
47	1	1	16	98	1	0.0
48	5	14	15	95	1	0.0
49	2	7	16	89	5	0.6
50	0	0	16	100	0	0.0
51	0	0	13	93	0	-0.5
52	0	1	14	95	0	-0.3
53	0	0	14	95	0	-0.3
54	1	2	16	98	1	0.0
55	0	1	15	93	2	0.3
56	0	0	16	98	1	0.0
57	0	0	16	100	0	0.0
58	1	2	16	98	1	0.0
59	2	6	14	93	1	-0.3
60	4	10	15	98	0	0.0
61	0	0	16	100	0	0.0
62	3	10	15	91	3	0.5
63	0	0	13	91	1	-0.5

64	0	0	16	100	0	0.0
65	2	5	15	98	0	0.0
66	2	4	16	98	1	0.0
67	2	8	15	98	0	0.0
68	2	6	15	98	0	0.0
69	0	0	15	98	0	0.0
70	0	0	15	98	0	0.0
71	2	6	16	80	9	0.8
72	2	7	16	100	0	0.0
73	0	1	15	98	0	0.0
74	0	1	13	93	0	-0.5
75	0	0	16	100	0	0.0
76	1	5	11	89	0	-0.7
77	2	2	16	98	1	0.0
78	1	5	15	95	1	0.0
79	0	0	14	95	0	-0.3
80	0	2	16	100	0	0.0
81	6	19	12	89	1	-0.6
82	0	0	16	100	0	0.0

83	0	0	13	89	2	-0.2
84	0	0	15	98	0	0.0
85	5	6	11	70	8	0.2
86	0	0	16	100	0	0.0
87	1	5	16	100	0	0.0
88	0	0	14	93	1	-0.3
89	2	3	16	100	0	0.0
90	2	5	16	82	8	0.7
91	1	2	9	84	0	-0.7
92	3	3	14	95	0	-0.3
93	2	5	11	86	1	-0.7
94	4	14	16	95	2	0.3
95	4	9	14	95	0	-0.3
96	0	0	15	98	0	0.0
97	0	0	16	100	0	0.0
98	4	7	16	98	1	0.0

Note. Definitions: CVCRI = California Verbal Learning Test (CV) Cued Recall Intrusions score; CVTI = CV Total Intrusions score; CVRH = CV Recognition Hits score; CVDSM =

CV Discriminability score; CVFP = CV False Positives score; CVRB = CV Response Bias score. The participant that was excluded from verbal memory analysis due to questionable/impaired performance was participant #85.