ESTIMATING PREMORBID INTELLIGENCE USING THE NATIONAL ADULT READING TEST IN NEWFOUNDLAND

KAREN LYNN SHARPE, B.A.(Honours)

ESTIMATING PREMORBID INTELLIGENCE USING THE NATIONAL ADULT READING TEST IN NEWFOUNDLAND

BY

Karen Lynn Sharpe, B.A.(Honours)

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements of the degree of Master of Science

Department of Psychology

Memorial University of Newfoundland

August, 1990

St. John's Newfoundland

ISBN 0-315-65326-4

ABSTRACT

The reliability and validity of the National Adult Reading Test (NART, Nelson, 1982), a measure of predicted premorbid intelligence level, was examined. Subjects were twenty mild-moderate dementing (17 females, 3 males) and twenty nondementing (14 females, 6 males) individuals aged 59-89. The NART demonstrated high inter-rater reliability ($r = 0.96$, $p < .001$). The NART was a valid measure of intelligence in that it correlated well with, and predicted a substantial amount of variance in the Wechsler Adult Intelligence Scale-Revised Full Scale (WAIS-R FSIQ) and Verbal Intelligence Quotients (WAIS-R VIQ) in the control sample. The NART was relatively "dementia-resistant" in that NART performance did not significantly correlate with severity of dementia when the demographic variables were partialled out. Further, it was the only cognitive measure on which there was no significant difference between dementing and control subjects, when the demographic variables were partialled out. Regression equations to predict dementing subjects' premorbid WAIS-R FSIQ and VIQ from NART errors (NART FSIQ and VIQ) and from WAIS-R Vocabulary age-scaled scores (Vocabulary FSIQ and VIQ), were developed using data from the control subjects. Predicted NART FSIQ and VIQ were significantly more "dementia-resistant" than predicted Vocabulary FSIQ and VIQ. That is, post hoc Scheffe tests revealed that the NART predicted significantly higher WAIS-R FSIQs and VIQs than the Vocabulary subtest of the WAIS-R ($p < .05$). Demographic variables did not add a significant amount of predicted WAIS-R FSIQ variance when combined with either NART (4%) or Vocabulary subtest (<1%). Two WAIS-R algorithms were investigated in their ability to distinguish dementing from nondementing individuals. Coolidge's algorithm (Vocabulary age-scaled score $\geq$ 2 Block Design age-scaled score = dementia, Coolidge, Peters, Brown & Harsch, 1985) correctly classified a statistically ($p < .05$) but not a clinically significant proportion of subjects, and the V-P Split (WAIS-R VIQ - WAIS-R PIQ) did not significantly distinguish between dementing and control subjects, indicating that these algorithms may not be clinically useful in identifying dementia. The results of the present study indicate that the NART may be the procedure of choice for differential diagnosis of dementia in North America.

ACKNOWLEDGEMENTS

A great (roll your r's please) Scottish "Ta" is extended to Dr. Ronan O'Carroll for his healthy Type A, sarcastically witty, but seriously matchless supervision.

I would also like to extend my appreciation to my committee members, Drs. Mike Stones and Graham Skanes for their useful suggestions, but in particular to Dr. Skanes for his equally satirical yet immensely astute statistical knowledge. His Newfie-ism is compromised only by his inability to drink Screech.

This thesis would not have been possible without the help of the people who referred subjects to the study (Drs. Harris, Strong, Eaton, and Butler, Ms. Lee Stones and Ms. Debbie Howard), and indeed, of the subjects who participated. Thank you all for your interest in advancing scientific research.

I owe my sanity to Michael, and my gratitude to Sue, both of whom hung around despite bad moods, bad breath, and bad grammar. You're the best.

Finally, I would like to thank the Memorial University School of Graduate Studies for financial aid throughout my program, in the form of University Fellowships, Teaching Assistantships, and Research Assistantships.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Dementia is described in the revised third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-111-R) as having the features of "...impairment in short- and long-term memory, associated with impairment in abstract thinking, impaired judgement, other disturbances of higher cortical function, or personality change...Dementia may be progressive, static, or remitting" (American Psychiatric Association, 1987, pp. 103-104). Much research has been carried out investigating the etiology and medical diagnosis of dementia, however an exhaustive review will not be undertaken in this thesis. A brief overview of dementia will be given, derived from information provided in a recent review paper (Zarit & Zarit, 1983).

The cognitive changes seen in dementing individuals are presumed to be due to atrophy throughout the cerebral cortex and are not related to the normal aging process (Zarit & Zarit, 1983). The two major types of dementia are Dementia of the Alzheimer's Type (DAT) and Multi Infarct Dementia (MID), comprising an estimated 60% and 10-20% of cases of dementia in people over age 65, respectively. DAT is characterized by abnormal structures in the brain including senile plaques, neurofibrillary tangles, and granulovascular structures viewed on autopsy, and its course is typically a gradual progression of deterioration. MID occurs when an individual undergoes a series of small strokes caused by pieces of plaque on artery walls breaking off and travelling to the brain, and there they occlude cerebral blood flow, resulting in neuronal death. The course of MID is stepwise. The etiology of these types of dementia is unknown. Theories of DAT vary from genetic, to viral, to biochemical, while MID is purported to be related to the same risk factors as stroke or myocardial infarction. However, there is no clear explanation as to why the risk factors produce heart disease in one person and MID in another.

Dementia is easily recognizable in its later stages due to the vast disturbances in behavior and cognitive functioning that occur, but is often difficult to diagnose at the beginning of the process. In fact, definite medical diagnosis of dementia can only be made at post-mortem (Zarit & Zarit, 1983). The Computer Axial Tomography scan (CAT scan) was initially viewed as a promising tool for the diagnosis of dementia since it can detect cortical atrophy and enlargement of the ventricles and sulci in the brain. However, research as to the utility of this instrument in assessing dementia has revealed that some degree of cortical atrophy can occur in the normal elderly who demonstrate no cognitive impairment, and some clearly dementing individuals do not produce positive findings on the CAT scan.

Since definite medical diagnoses of dementia are impossible, and since cognitive deterioration in dementia appears to be universal and is usually the impetus for dementing patients to come to medical attention, researchers have attempted to devise psychometric measures to aid diagnosis in the early stages of this disorder. Ideally, clinicians would have a measure of cognitive function taken at a time prior to the development of a dementing process, in which case the patient's premorbid cognitive status could be compared with his/her current level of cognitive functioning. However, this luxury seldom exists in clinical practice as many dementing individuals experience their first visits to psychiatric settings only after the suspected dementing process has begun. Therefore clinicians have attempted to estimate premorbid levels of cognitive functioning, using tests which have been dubbed "dementia-insensitive" or "dementia-resistant" inventories. It is, of course, unlikely that any psychological test will be completely "dementia-insensitive" due to the immense cortical change which takes place in the latter stages of dementia. Therefore the following discussion will be concerned with relative rather than absolute "dementia-insensitivity".

Some of the earliest methods used in clinical practice for distinguishing dementing from nondementing individuals included Wechsler Adult Intelligence Scale (WAIS) or its revised version (WAIS-R) algorithms. The WAIS/WAIS-R algorithms were based on the premise that dementing individuals would perform more poorly on some tasks ("no hold" tasks) than on others ("hold" tasks), while nondementing individuals would do equally well on both. The "no hold" tasks were purportedly detrimentally affected by the dementing process while the "no hold" tasks were purportedly not detrimentally affected by the dementing process. The "no hold" tasks generally involved learning/manipulating new unpracticed information, while "hold" tasks generally involved more automatic or over-learned tasks, usually verbal. A large discrepancy between "hold" and "no hold" tasks was purported to indicate organicity. Some support for this assumption has been found in that some tasks (e.g., the Vocabulary subtest of the WAIS/WAIS-R) have been observed to "hold", better than others (e.g., the Block Design subtest of the WAIS/WAIS-R) in dementia (e.g., Coolidge, Peters, Brown & Harsch, 1985).

The research that has been carried out to determine the accuracy of many of these equations has been disappointing (Vogt & Heaton, 1977). Some of the more successful and well-researched methods of distinguishing dementing from nondementing individuals have been

those professing to estimate premorbid intellectual functioning. That is, if one can predict the level of cognitive functioning before the disease process began, this estimate can be compared to a measure of current cognitive ability. If the discrepancy is large enough, dementia is strongly suspected. Equations have been derived from large normative samples by regressing scores on purported "dementia-resistant" variables, or measures on which dementing individuals are expected to receive scores similar to nondementing individuals, (i.e., demographic information, word reading abilities, or the Vocabulary subtest of the WAIS or WAIS-R), against measures of current intelligence (typically the WAIS or WAIS-R full scale intelligence quotients [FSIQ], verbal intelligence quotients [VIQ] and performance intelligence quotients [PIQ]). Individuals' scores on the chosen "dementia-resistant" measures, are then entered into the standardized equations to determine a premorbid IQ score. If the predicted premorbid IQ minus observed IQ discrepancy is large enough, dementia is suspected.

Early detection of dementia is important to the clinician because first, it allows a diagnosis thus an explanation of behavior. Second it allows him/her to implement treatment quickly for treatable cases. A recent review paper outlined treatment directions for the subtypes of dementia which spanned from cholinesterase inhibitors (e.g., tetrahydroaminoacridine) to controlling hypertension (Whalley, 1989). Whalley (1989) concluded that at present there are no confirmed effective drugs for the treatment of dementia, but encouraged a positive outlook for the future. Early intervention will be crucial in new treatments, as if they are to work they will have to do so before neuronal death occurs. Third, early detection of dementia allows for research into possible treatments. If research is to continue, accurate diagnoses of dementia are required to ensure researchers are examining homogeneous groups. Fourth, untreatable cases may benefit from programs aimed at slowing the dementing process. Finally, detection of dementia provides a gauge of how far the disease has progressed and how fast it is progressing, thus allowing the clinician to give a prognosis. The following will review the literature on psychometric methods of distinguishing dementing from nondementing individuals, and will outline the hypotheses for a study attempting to assess the validity of the National Adult Reading Test (NART), a purported measure of premorbid intelligence, in Newfoundland.

1.1. WAIS/WAIS-R Algorithms

1.1.1. Deterioration Indices

Several researchers have proposed WAIS algorithms for discriminating cerebral dysfunction from normality (for reviews see Savage, Britton, Bolton & Hall, 1973, Vogt & Heaton, 1977). The algorithms are all based on the finding that some WAIS/WAIS-R subtests ("hold") are less "dementia-sensitive" (i.e., dementing individuals show less of a performance decline in comparison with nondementing individuals) than others, the "no hold" subtests, which dementing individuals have greater difficulty performing than normal control subjects. Therefore discrepancies between "hold" and "no hold" subtests should provide an index to distinguish dementing from nondementing individuals.

The first WAIS algorithm was Wechsler's deterioration quotient (DQ) which was determined from the Wechsler-Bellevue scale (Wechsler, 1944). It divided "hold" minus "no hold" by "hold" subtests multiplied by 100 (Wechsler, 1944). The "hold" subtests included Comprehension, Information, Object Assembly, and Picture Completion, and these were purported to be least susceptible to cognitive decay. The "no hold" subtests included Digit Span, Arithmetic, Digit Symbol, and Block Design, and were purported to be most susceptible to cognitive decay. Wechsler later revised his DQ for use with the WAIS (Wechsler, 1955) and substituted the Vocabulary subtest for Comprehension in the "hold" subtests, and Similarities for Arithmetic in the "no hold" subtests. Use of these formulae, however, produced both errors of commission and ommission since there were wide variations among individuals. For example, since the "hold" tests have a large verbal component, cognitively unimpaired individuals who were proficient at verbal tasks but poor at performance tasks, received a "dementing" profile.

An early study compared the ability of WAIS-derived formulae (e.g., Wechsler's DQ) to identify organicity in a comparison of a community aged sample ($\underline{N}$ = 29) with a group of dementing individuals ($\underline{N}$ = 42, Savage et al., 1973). The results demonstrated that the discriminatory ability of most of the formulae was poor. Of particular importance, if the hit rate was high, the number of false positives (controls categorized as dementing) was also high. For instance, using Hewson's eight ratios in which eight WAIS equations identify cut-off points

indicating Neurotic vs. Normal, vs. Organic, 58.6% of dementing subjects were correctly classified, but 38.0% of control subjects were also classified as organic. Savage and his colleagues' (1973) best classification system was the revised Wechsler DQ, but even using this index, the hit rate was fairly low (52.3%) and a sizeable proportion of control subjects were misclassified (6.9%).

A second study comparing various Wechsler deterioration indices using a large sample of 117 neurologically impaired individuals versus 116 non-impaired individuals, also reported disappointing findings (Vogt & Heaton, 1977). Vogt and Heaton (1977) found that although all indices except one exceeded the chance levels for differentiating the groups, all measures except two (Hunt, 1949 and Hewson, 1949, cited in Vogt & Heaton, 1977) misclassified large numbers of unimpaired individuals as impaired. Further, the researchers used patients with extreme impairment which they indicated probably resulted in liberal estimates of the WAIS indices' abilities to discriminate the groups, and still the indices had only limited success in their ability to discriminate. It is likely that had more mild/moderate cases been used, which is the population with which these indices would have the most clinical utility, the researchers would have found even less impressive results. Vogt and Heaton (1977) themselves concluded that even the most successful formulae as determined by their study (Hunt, 1949, Hewson, 1949, cited in Vogt & Heaton, 1977), would only have clinical utility if patients scored above the cut-off, while dismissal of a dementing process could not be done if a patient scored within the normal range.

1.1.2. Coolidge's algorithm

Coolidge and his colleagues (1985) proposed a more recent WAIS based algorithm to distinguish organic from nonorganic illness. They statistically analyzed data from an early study (Crookes, 1974) in which 148 patients with uncertain diagnoses (dementia vs. depression) were tested with the WAIS and classified after at least one year follow-up as either dementing or depressed. The Crookes study found that dementing subjects scored best on the Vocabulary subtest and worst on Block Design and Digit Symbol. Coolidge and his colleagues (1985) re-analyzed the data from the Crookes study and found that the Vocabulary subtest was the only one which did not discriminate between dementing and depressed individuals while the Block Design subtest discriminated the best. They then derived a WAIS algorithm: dementia is

indicated if the Vocabulary subtest age-scaled score is greater than or equal to twice the Block Design subtest age-scaled score ($V \geq 2BI.D$ = dementia), and applied the algorithm to the data from the Crookes study. They found that this algorithm correctly classified nearly 3/4 of demenring (74%) and depressed (74.5%) individuals. However, the authors did not adequately report what diagnostic criteria were used to assign the follow-up classifications of "dementia" or "depression", leaving the validity and reliability of diagnoses in question.

A later study investigated the accuracy of the same algorithm ($V \geq 2BI.D$) in distinguishing DAT from normal elderly controls, elderly depressed individuals, and other organic conditions (i.e., MID, Huntington's Disease, Korsakoff's Psychosis, and Alcoholic Dementia, Crawford, Parker, Besson, & Beavan, unpublished manuscript). The authors found that none of the normal elderly subjects and only 19% of the elderly depressed patients exhibited this profile. This was in contrast to 68% of DAT patients who exhibited the profile, yielding a classification accuracy of 84.1% when DAT was compared with normal elderly controls, and 74% classification accuracy when DAT was compared with elderly depressed individuals. Further, a comparison of DAT with each of the other organic conditions, indicated that the algorithm was beneficial in distinguishing DAT from Korsakoff's Psychosis (classification accuracy = 78%), Alcoholic Dementia (classification accuracy = 85%), and Huntington's Disease (classification accuracy = 75%), but less beneficial in distinguishing DAT from MID (classification accuracy = 64%). In other words, the algorithm was beneficial in differential diagnosis of DAT from normal elderly controls, elderly Depressives, Korsakoff's Psychosis, Alcoholic Dementia, and Huntington's Disease. It was less successful in differential diagnosis of DAT from MID. However, the authors' separation of these two types of dementia (DAT and MID), indicated that the algorithm ($V \geq 2BI.D.$) was probably more beneficial in distinguishing DAT from Normals and Depressives than in distinguishing MID from Normals and Depressives, however this was not statistically investigated. Therefore, although the algorithm may be relatively successful in the differential diagnosis of DAT, it may not be as successful in differential diagnosis of patients suffering from the MID form of dementia.

1.1.3. V-P Split

A commonly used index for identifying organicity in clinical practice is the WAIS/WAIS-R V-P Split. The V-P Split is determined by subtracting the total PIQ score from the total VIQ score (Field, 1960). The logic for this WAIS/WAIS-R based algorithm is the same as used in the other WAIS/WAIS-R based algorithms mentioned above. That is, PIQ is considered to contain more "no hold" tasks (i.e., tasks which require more cognitive strategy and particular types of tasks, such as visual-spatial tasks, which appear to be more sensitive to organic impairment), than VIQ. The common rule-of-thumb is that if the discrepancy is greater than or equal to 15, there is strong need for further investigation (Wechsler, 1981, p. 36).

Although this equation is commonly used in clinical practice, little research has been carried out to determine its validity. Further, much of the research that has been carried out has disregarded the direction of the discrepancy (Field, 1960; Grossman, 1983; Naglieri, 1982; Wechsler, 1981). That is, discrepancies have contained both VIQ > PIQ and PIQ > VIQ added together, even though the predicted direction for determining whether an individual is dementing clearly is VIQ > PIQ. Clinical use of frequency tables generated in the above studies to determine abnormality of observed V-P Splits, leads to potentially inaccurate judgements of the presence of cognitive impairment, due to the unproven assumption that PIQ > VIQ occurs equally frequently as VIQ > PIQ in the cognitively intact general population. Further, even with the methodological flaw of combining both VIQ > PIQ and PIQ > VIQ, Field (1960) found that a discrepancy of 15 or more points between VIQ and PIQ (VIQ > PIQ) occurred in at least 10% of his sample over age 65, suggesting that a 15 point discrepancy is not a clear indicator of cognitive impairment in the elderly. This is in contrast to the three other studies (Grossman, 1983; Naglieri, 1982; and Wechsler, 1981), all using the WAIS-R standardization sample (N = 1880), which indicated that a 15 point predicted discrepancy (i.e., VIQ greater than or equal to 15 points higher than PIQ) occurred rarely (<1%) in the non-impaired elderly population and therefore was suggestive of cognitive deficits. However, it was unclear whether Field's (1960) sample was representative of the general elderly population or whether subjects were screened for cognitive impairment. Therefore, these "equivocal" results could easily be explained if Field assessed a general elderly population, and the other researchers assessed a screened population. That is, if Field's sample included a general elderly population, it is likely that some

dementing individuals would have been included, thus potentially inflating the observed V-P Split (VIQ > PIQ), in comparison with the size of the V-P Split of the non-impaired elderly subjects comprising the WAIS-R sample employed by Grossman (1983), Naglieri (1982), and Wechsler (1981).

One pair of researchers has presented separate VIQ > PIQ and PIQ > VIQ norms, using the non-impaired subjects comprising the WAIS-R standardisation sample (Matarazzo & Herman, 1985). They found that 8.7% of all subjects' VIQs were 15 or more points greater than their PIQs (8.5% aged 45-74), and similar figures were reported for PIQ > VIQ. This means that approximately 9/100 non-impaired individuals would be suspected to be impaired by the 15-point rule-of-thumb regarding the V-P Split. This figure does not even reach the commonly accepted statistical significance of 5%, let alone address the issue of clinical significance.

Research investigating whether or not dementing individuals differ from nondementing individuals in the size of the verbal performance discrepancy indicates that the V-P Split is not a highly promising indicator of dementia. Hart, Smith, and Swash (1986) compared elderly dementing individuals with elderly nondementing individuals on various cognitive measures. The authors found that the WAIS V-P Split was not of great enough magnitude to be a sensitive indicator of intellectual decline. In their study, some dementing individuals actually scored higher on PIQ ("no hold") than on VIQ ("hold"), and extreme discrepancies in favour of the VIQ (i.e., up to 21-25 points) were found in both dementing and control subjects.

Lezak (1983) was critical of the V-P Split on two accounts. First, she argued that VIQ and PIQ are both compilations of functions which are dissimilar to each other and have relatively low intercorrelations. Second, she noted that VIQ and PIQ overlap in functions measured. She attributed the overlap to the method which was used to assign subtests to VIQ and PIQ (dictated rather than assigned through factor analysis). Further, it has been reported that many other conditions (e.g., psychosis) may produce VIQ > PIQ, reducing this measure's specificity in identifying dementia (Guertin, Ladd, Frank, Rabin, & Hiester, 1966).

1.1.4. Subtest scatter

Another commonly used but scarcely researched clinical "indicator" of dementia, is WAIS or WAIS-R subtest scatter (Lezak, 1983). The most frequently used method of determining subtest scatter is deviations between pairs of subtests (Field, 1960). Research to determine the extent of subtest scatter in the non-impaired general population has indicated that a three point discrepancy between any two subtests occurs in less than 15% of normally functioning individuals (Wechsler, 1981, p. 36). Other research has indicated that scatter of 3.5-4.3 scaled scores occurs in 5% of the normal population and scatter of 4.6-5.5 scaled scores occurs in 1% of the normal population (Field, 1960). Further, some WAIS-R subtests (Digit Span and Object Assembly) exhibit more variability than others (Vocabulary and Information) (Lezak, 1983). In other words, there may be more variability in the non-cognitively impaired public than originally assumed and therefore clinicians should refrain from using subtest scatter as a means in and of itself for diagnosing dementia. Further, much of the research that has been carried out on subtest scatter has usually involved mentally retarded individuals or children (e.g., Coolidge, Rakoff, Schwellenbach, Bracken, & Walker, 1986; Roszkowski & Spreat, 1982; 1983), and has typically revealed that subtest scatter is relatively common in mentally retarded individuals (Coolidge et al., 1986) and lower functioning individuals in general (Roszkowski & Spreat, 1982). Therefore, subtest scatter does not appear to be specific to dementia and its use for diagnosing organicity may inaccurately identify lower functioning individuals as cognitively impaired.

1.1.5. Summary of WAIS Algorithms.

It is evident from the above discussion that much research has been carried out in search for the best configuration of WAIS or WAIS-R subtests that distinguish cognitively intact from cognitively non intact subjects. Although some configurations have been more successful than others, none has proven satisfactory in the detection of dementia (with the possible exception of Coolidge and his colleagues', 1985, formula which requires more research before clinicians can use the algorithm with confidence). As Miller (1977, p. 109) put it, "For a few psychologists the search for the right Wechsler sub-test combination still goes on rather like the medieval alchemist's search for the philosopher's stone, and with as little likelihood of ultimate success".

One explanation for the lack of success of these purported deterioration indices, is that the

WAIS/WAIS-R "hold" versus "no hold" formulae may include subtests which are not the best indicators of cognitive decay and maintenance. For instance, as mentioned previously, one group of researchers concluded from a followup study, that the Vocabulary subtest of the WAIS was the only subtest which did not differentiate between elderly dementing and elderly depressed individuals (Coolidge et al., 1985). Indeed, several researchers have noted that neurologically impaired individuals show some decrements on all WAIS or WAIS-R subtests (Lezak, 1983; Russell, 1972; Vogt & Heaton, 1977).

1.2. Measures of premorbid intelligence

Several methods of estimating premorbid cognitive functioning (i.e., the estimated level of intelligence prior to the development of a dementing disorder) have been proposed. These include the use of demographic information, the Vocabulary subtest of the WAIS or WAIS-R, and word reading abilities. The research on these indices generally indicates they have enjoyed better success than the WAIS/WAIS-R algorithms discussed above. They are reviewed below in terms of their development, success in correct classification of dementing individuals, and their success in comparison with each other.

1.2.1. Demographic variables

One commonly used method of estimating premorbid intelligence is clinical guesswork based on demographic information. The logic is that intelligence level should be related to such information as education and occupation (other demographic information has also been used as discussed below). However, this can lead to large miscalculations with the elderly due to often scanty education records and limited early opportunities leading to occupations below their capabilities. Further, many elderly females have never been gainfully employed and estimates have been based on the crude measure of husband's occupation.

Nevertheless, some recent work has been carried out in the United States, the United Kingdom, and Canada in which demographic information has been entered into regression equations to yield estimated premorbid Performance, Verbal and Full Scale WAIS/WAIS-R IQ equivalents (Barona, Reynolds, & Chastain, 1984; Blair & Spreen, 1989; Crawford, Stewart, Cochrane, Foulds, Besson, & Parker, 1989; Eppinger, Craig, Adams & Parsons, 1987; Karzmark,

Heaton, Grant & Matthews, 1985; Wilson, Rosenbaum, Brown, Rourke, Whitman, & Grisell, 1978). The pioneering work in this area was carried out by Wilson and his colleagues (1978), who used most of the WAIS standardization sample (N = 1700) to develop demographic equations to predict WAIS IQ. They built regression equations which contained information about age, sex, race, education, and occupation which predicted 54% of the variance in WAIS FSIQ, 53% in WAIS VIQ, and 42% in WAIS PIQ. A further study using 140 neurological patients versus 140 non-neurological patients, compared the efficacy of these demographic equations with the efficacy of the revised Wechsler DQ in classifying cases within the two groups (Wilson, Rosenbaum & Brown, 1979). They found that in both the initial run and the double cross validation run, the demographic equations were superior to the revised Wechsler DQ (71.8% vs. 63.2% and 72.8% vs 61.8% correct classification of cases, respectively). The Wilson equations were cross validated by one group using a large sample (N = 491) of subjects without neurological problems (Karzmark et al., 1985). Karzmark and his colleagues (1985) compared WAIS obtained scores with the predicted scores using the Wilson demographic formulae. They found that demographic variables predicted less of the variance in WAIS FSIQ in their sample than in the original Wilson study. That is, only 46% of the FSIQ variance was accounted for by the Wilson formulae for their employed subjects and the amount dropped to 42% when considering the total sample, in comparison to 54% in the Wilson et al. (1978) study. This might be explained, however, given that the Wilson formulae were developed using a sample from 1955. It would be expected that the relationship between demographic variables and IQ would change over the 30 years between the Karzmark et al. (1985) study and collection of the WAIS standardization sample (Wechsler, 1955).

Barona created similar formulae for estimation of WAIS-R IQ scores (Barona et al., 1984). Using the WAIS-R standardization sample (N = 1880), demographic estimated premorbid IQ equations were determined by regressing the Wilson demographic variables (i.e., age, sex, race, education, and occupation) plus two new demographic variables (urban vs rural residence and handedness), against WAIS-R FSIQ, VIQ, and PIQ. The Barona demographic equations had a lower predictive accuracy than the Wilson equations, predicting only 36% of the variance in WAIS-R FSIQ, 38% in WAIS-R VIQ and 24% in WAIS-R PIQ. The equations were cross-validated on a group of 80 neurologically normal subjects, and accuracy of the demographic

equations in distinguishing between a group of 83 brain impaired subjects and the neurologically normal controls was determined (Eppinger et al., 1987). Cross-validation produced a substantial increase in predictive accuracy over the Barona study, with the demographic equations predicting 58% of the variance in WAIS-R FSIQ, 61% in WAIS-R VIQ, and 36% in WAIS-R PIQ (Eppinger et al., 1987). Further, there were no significant differences between neurologically normal and brain impaired subjects on demographically estimated IQ, but there were significant differences between groups on obtained IQ. However, although there were large significant differences between estimated and obtained IQ for brain impaired groups, there were also significant differences between estimated and obtained IQ scores for control subjects, indicating that the Barona demographic index (Barona et al., 1984) consistently overestimated IQ.

Barona and Chastain (1986) attempted to develop more accurate demographic equations to predict WAIS-R IQ. They re-analyzed the data from the WAIS-R standardization sample, excluding the first two age categories (16-17 years and 18-19 years), since these subjects' occupational codings were primarily based on the head of household's occupation, and excluding races other than Black or White, since there were very small numbers in the "other" category. The re-analysis on the smaller sample (N = 1433), produced demographic equations which predicted more variance in WAIS-R FSIQ, VIQ and PIQ than the initial Barona et al. (1984) study (43% vs. 36%, 47% vs. 38%, and 28% vs. 24%, respectively). These results indicate that the revised Barona and Chastain (1986) equations should be favoured over the original equations (Barona et al., 1984), when predicting premorbid IQ for Black or White patients over the age of nineteen.

Similar equations have been developed in the U.K. to estimate premorbid WAIS IQ from demographic information (Crawford, Stewart, Cochrane, Foulds, Besson, & Parker, 1989). Crawford and his colleagues (1989b) found that a demographic equation consisting of social class, age, education, and sex, predicted 50% of the variance in WAIS FSIQ and VIQ, and 30% of the variance in WAIS PIQ. The researchers concluded that demographic equations may have some utility in predicting premorbid intelligence.

These data suggest some promise in the use of demographic variables to estimate premorbid cognitive functioning, however, some problems have been described in the literature

(Eppinger et al., 1987; Silverstein, 1987). These include the restricted range of predicted IQ using Barona and his colleagues' (1984) equations (69-120), the difficulty of fitting many occupations into the Barona classification system, the lack of finely-tuned discriminations between higher education, and the lack of control for individuals who are high or low achievers (Eppinger et al., 1987). These problems might intensify while using the equations with an elderly population. With this specialized group there may not be enough variation in type of occupation or amount of education to be discriminatory, due to reduced life opportunities. Further, one researcher found that neither the Wilson nor the Barona demographic formulae were accurate in classifying patients in the seven Wechsler IQ categories (Very Superior, Superior, High Average, Average, Low Average, Borderline, and Mentally Retarded, Silverstein, 1987). Even the best estimators in Silverstein's (1987) study (i.e., Wilson equations) misclassified more than half of the subjects (i.e., placed them in a category above or below what they should have been in). In addition, Wilson and his colleagues (1978) themselves pointed out that their predicted premorbid demographic equations based on subjects from 1955 would overestimate IQ due to the increased educational attainment of individuals since then. The overestimation of IQ by the Wilson formulae was substantiated in a study comparing the WAIS, demographic estimates using Wilson's formulae, and another test of intelligence (the Quick Test) in a sample of 50 patients with mixed psychiatric diagnoses (Law, Price, & Herbert, 1981). They found that the Wilson equations for estimating premorbid IQ significantly overestimated WAIS IQ by an average of 7.6 IQ points.

Further evidence for the Wilson et al. (1978) formulae overestimating IQ was found in a study comparing neurologically intact psychiatric inpatients with psychiatric outpatients (Klesges, Sanchez, & Stanton, 1981), and has also been found in a series of studies (Bolter, Gouvier, Veneklasen & Long, 1982; Gouvier, Bolter, Veneklasen & Long, 1983; Klesges, Fisher, Vasey & Pheley, 1985) investigating the utility of his equations in differentiating brain injured from "pseudo neurological" patients (i.e., patients who were referred for neurological testing resulting in normal test results). In general, these studies found that the Wilson et al. (1978) demographic IQ estimates (estimated WAIS FSIQ, VIQ and PIQ) overestimated premorbid IQ, therefore risking a greater possibility of misclassifying non brain damaged patients as brain damaged as a result of the discrepancy analysis (i.e., a high premorbid measure is taken as evidence of organicity). In addition, the groups of researchers investigated the utility of incorporating Wilson et al.'s (1978)

suggested educational adjustment (educational weights x 0.82), to reflect the increase in educational attainment since 1955, when Wilson et al.'s (1978) sample was collected. All of the above studies (with the exception of the cognitively unimpaired psychiatric outpatients in Klesges et al., 1981), found these adjustments resulted in little to no improvement in predictive accuracy of the demographic equations, and in some cases actually decreased the predictive accuracy. Although Klesges et al. (1981) cautiously approved the clinical use of educationally-adjusted Wilson et al. (1978) demographic estimates of premorbid IQ, the other groups of researchers mentioned above (Bolter et al., 1982; Gouvier et al., 1983; Klesges et al., 1985), discouraged their use due to questionable validity.

The above studies (Bolter et al., 1982; Gouvier et al., 1983; Klesges et al., 1981; 1985) have been criticized by Crawford (1989) on several accounts. First, all of the studies used clinical subjects (i.e., psychiatric patients or "pseudo neurological patients" with no evidence of organic deterioration on a CAT scan). However, intollectual impairment in these clinical samples is likely. For example, negative CAT scan results do not rule out cognitive impairment. Further, schizophrenia and/or its associated medications likely produces intellectual impairment. Crawford (1989) therefore argued that these subjects were not adequate controls, a better control sample being cognitively unimpaired subjects. Indeed, Crawford (1989) pointed out that the control subjects used in Klesges et al. (1981, psychiatric outpatients), and Klesges et al. (1985, "pseudo neurological" patients) had lower mean IQ scores than would be expected in a cognitively intact US sample (1.5 and 0.5 S.D. lower, respectively).

Intellectual impairment in the comparative sample would distort research findings in two ways. First, there would be weaker correlations between demographically predicted and observed IQ, and second the demographic variables would appear to overestimate premorbid IQ in comparison to obtained IQ. These were precisely the findings of the above studies (Bolter et al., 1982; Gouvier et al., 1983; Klesges et al., 1981; 1985).

Crawford (1989) outlined two further methodological flaws in the above studies (Bolter et al., 1982; Gouvier et al., 1983; Klesges et al., 1981; 1985), which leaves their conclusions in question. First, the authors neglected to ensure that the range of demographic variables occurring in their samples reflected the range of these variables in the general population.

Therefore it is possible that spuriously low estimates of the population correlation between predicted and observed IQ were obtained. Second, two of the studies (Bolter et al., 1982 and Gouvier et al., 1983) used small sample sizes (23 "pseudo neurological" patients, 11 recovered and 11 non-recovered brain injured patients). Use of multiple predictor variables, such as demographic variables, with small samples can produce misleading high or low correlations by chance (Crawford, 1989). In light of the above methodological problems, Crawford (1989) argued that the conclusions reported by Klesges et al. (1981; 1985), Bolter et al. (1982) and Gouvier et al. (1983), that is, that demographic variables overestimated IQ and predicted a small amount of variance in WAIS IQ, were unwarranted.

A further concern regarding the use of demographic equations to predict premorbid intelligence is that they can only be used for predicting intelligence of individuals with similar characteristics as the standardization samples. For instance, the relationship between demographic variables and IQ may not be equivalent between countries (Crawford, Stewart, Parker, Besson, & Cochrane, 1989). That is, people may require greater intelligence in one country to attain a similar education level in another. In other words, applying the currently available equations to a sample from a different country might not yield an accurate prediction of premorbid intelligence, and therefore unique equations should be developed for each unique population studied.

In defence of the demographic method of predicting IQ, however, Crawford and his colleagues (1989b) pointed out that prediction of *premorbid* functioning by demographic equations has the great advantage of being completely independent of current cognitive functioning. Further, in comparison with the use of word reading abilities (e.g., the NART) to estimate premorbid IQ, demographic equations can be used with certain patients (e.g., dyslexics and illiterates) with which the NART cannot be used with (Crawford, 1989).

1.2.2. WAIS/WAIS-R Vocabulary subtest

Some researchers have noted that the Vocabulary subtest of the WAIS/WAIS-R is the least "dementia-sensitive" (i.e., least likely to suffer performance decline in dementing individuals) of all the WAIS/WAIS-R subtests (Coolidge et al., 1985; Lezak, 1983). Further, Coolidge and his colleagues (1985) found it to be the only WAIS subtest that did not discriminate groups (dementing vs. control subjects). That is, although neither the Vocabulary subtest nor any other cognitive test is likely to be completely "dementia-insensitive", it has been observed that dementing subjects perform better on it than on any other WAIS/WAIS-R subtest. In addition, the Vocabulary subtest has been reported to correlate the highest with both the WAIS and the WAIS-R FSIQ in comparison to the other Wechsler subtests, the correlations ranging between 0.85-0.87 (Wechsler, 1955; 1981). Therefore, the Vocabulary subtest appears to have satisfied some of the criteria necessary to be used as a premorbid IQ indicator namely, that it correlates highly with WAIS/WAIS-R FSIQ and that it appears to discriminate the least between dementing and nondementing individuals, in comparison with the other WAIS/WAIS-R subtests. However, in a review article, Crawford (1989) reported that several studies (Russell, 1972; Swiercinsky & Warnock, 1977, cited in Crawford, 1989; Vogt & Heaton, 1977), have found that non cognitively impaired individuals performed significantly higher on the Vocabulary subtest than cognitively impaired individuals. Crawford (1989) cautioned against ready acceptance of these results, however, since the researchers did not take either educational differences (Russell, 1972; Vogt & Heaton, 1977), or premorbid intelligence differences (Swiercinsky & Warnock, 1977, cited in Crawford, 1989) between the groups into consideration. Therefore, it is impossible to discern whether performance variability was due to poorer (e.g., educational) backgrounds of the cognitively impaired groups rather than cognitively impaired subjects' reduced ability to perform well on the Vocabulary subtest.

Nelson and McKenna (1975) proposed a method for estimating a premorbid IQ score from the Vocabulary subtest alone. By regressing control subjects' scores on the Vocabulary subtest against their scores on the WAIS FSIQ, VIQ and PIQ, they created regression equations into which Vocabulary subtest age-scaled scores of individuals could be entered to produce predicted premorbid WAIS IQ scores. These Vocabulary IQs could then be compared with the current IQ measures of WAIS FSIQ, VIQ and PIQ in neurologically impaired individuals, to determine the

extent of the discrepancy between predicted and obtained IQ, and clinically this method has been used to estimate the degree of cognitive deterioration. Published studies on the predicted Vocabulary IQ have primarily appeared in the context of comparing it with estimated premorbid IQs determined by word reading tests, and will therefore be discussed below under the heading of 2.4.3. NART in comparison with other methods of distinguishing dementing from nondementing individuals.

1.2.3. Word reading ability: The Schonell Graded Word Reading Test

One pair of researchers noted through clinical observation that reading ability (accuracy of oral pronunciation) was relatively well-preserved in dementing individuals (Nelson & McKenna, 1975). Further, they reasoned that since reading of complex sentences requires use of syntax and semantics and therefore is more cognitively demanding than reading singular words, the latter would be deemed more useful in estimating pre-existing intellectual functioning (Nelson & McKenna, 1975). Research into the usefulness of word reading as a valid estimator of premorbid intellectual functioning began with the Schonell Graded Word Reading Test (SGWRT), constructed for use with children from most elementary levels (Nelson & McKenna, 1975). The authors administered the SGWRT and the WAIS to 98 neurologically intact subjects (one group of hospitalized and one group of nonhospitalized controls) and 45 hospitalized dementing subjects. Findings from this study supported the hypothesis that word reading ability and general intelligence were positively correlated in normal adults ($r = 0.75$, $p < .001$, $r = 0.78$, $p < .001$, $r = 0.56$, $p < .001$ between reading score and WAIS FSIQ, WAIS VIQ, and WAIS PIQ, respectively). Further, there were no significant differences between dementing and control groups on reading ability (SGWRT scores), but there were significant differences between the two groups on all measures of the WAIS. This led the authors to conclude that although reading ability is likely affected by severity of dementia, it can be maintained at a high level despite deterioration in other skills. Ruddle and Bradshaw (1982) replicated Nelson and McKenna's (1975) study using 78 normal controls, 75 patients suspected of cognitive impairment not due to a dementing disorder (divided into confirmed cortical atrophy and equivocal cases), and 22 dementing subjects. They regressed the SGWRT against WAIS FSIQ, VIQ and PIQ in their control subjects and found no significant difference between Nelson and McKenna's equations for predicting WAIS FSIQ from the SGWRT and their own. In addition, they found that all of the patient groups had significantly

higher mean discrepancy scores than the control subjects (calculated by SGWRT predicted minus observed WAIS IQ), but the dementing group displayed the highest mean discrepancy scores. The authors concluded that (a) SGWRT provided a reliable estimate of general intelligence, (b) the regression equation for predicting WAIS FSIQ by SGWRT proposed by Nelson and McKenna was accurate and useful to clinical practice, and (c) that use of discrepancy scores between SGWRT predicted WAIS IQ and current WAIS IQ would produce fewer false negatives in the diagnosis of dementia than in the diagnosis of less well defined neurological impairment.

Although the above mentioned studies indicated some promise for the use of the SGWRT in diagnosing dementia, criticisms against using this measure included its low IQ ceiling of 115, and that the test included a mixture of regular and irregular words, many of which were long (Nelson & O'Connell, 1978). Nelson and O'Connell (1978) found that dementing individuals made significantly more errors than normal controls in reading long regular words. The authors concluded that this was due to the ability of control subjects to correctly guess the pronunciation of the words by applying intelligent guesswork, a task which was difficult for the dementing subjects. That is, even though a neurologically intact person has never seen a word before, if regular grammatical rules are applied, the word can be pronounced correctly, while dementing individuals are less likely to guess the correct pronunciation.

1.2.4. A less dementia sensitive word reading test?: The National Adult Reading Test

The National Adult Reading Test (NART) was developed in an attempt to produce a measure which was relatively unaffected by the dementing process, and to increase the IQ ceiling that the SGWRT offered (Nelson, 1982). Nelson (1982) reasoned that a better estimate of premorbid IQ would involve a measure which minimized cognitive strategy at the time of testing. The NART attempts to do just this by simply having the individual read a list of 50 short irregular words (e.g., gauche), and accuracy of pronunciation is scored. Short words are easier than long words for dementing individuals to process and the irregularity demands previous familiarity with both pronunciation and spelling in order to produce a correct response. In other words, subjects are unable to use standard grammatical rules to successfully pronounce the word.

1.2.4.1. Psychometric properties: Reliability

The NART has received considerable support from studies attempting to establish its reliability. The test has good split-half reliability, as demonstrated by the standardization sample (Chronbach's Alpha, r=0.93, Nelson, 1982), and another study investigating NART performance of 201 neurologically normal subjects (Spearman-Brown formula, r=0.90, Crawford, Stewart, Garthwaite, Parker, & Besson, 1988). The NART's test-retest reliability over a 10-day period was extremely high (r = 0.98) in a group of 61 neurologically normal subjects (Crawford, Parker, Stewart, Besson, & DeLacey, 1989). Adequate inter-rater reliability was established in a study comparing 10 experienced Clinical Psychologist NART users' independent scoring of the NART for 12 psychology outpatients' responses, (Kendall's coefficient of concordance, W=0.88, O'Carroll, 1987). This effect was replicated by Crawford and his colleagues (1989a), in which 5 experienced and 5 inexperienced NART users' scoring was compared on 40 NART recordings of nonpatients. The correlations between all pairs of experienced raters and between all raters together ranged from 0.96-0.98. Although the raters differed significantly in the strictness with which they scored the NART, 82% of the words had a 90% or greater agreement rate and 64% of the words had a 95% or greater agreement rate. Thus the NART has been shown to be internally consistent, to have good test-retest reliability, and to have high inter-rater reliability, regardless of whether the scorer is experienced or inexperienced with the test.

1.2.4.2. Psychometric properties: Validity

Of paramount importance, psychometric instruments must have construct validity. That is, they must measure the construct that they purport to measure. In the case of the NART, it must be demonstrated that the NART measures intelligence level. The NART was standardized on a group of 120 inpatients with extra-cerebral disorders, between the ages of 20-70 years (Nelson, 1982). Subjects completed a prorated WAIS (seven subtests) and the NART, and from this data regression equations were developed to predict WAIS IQ from NART error scores. Nelson (1982) found that the NART predicted 55% of the variance in WAIS FSIQ, 60% in WAIS VIQ, and 32% in WAIS PIQ. The NART was cross-validated on a larger sample than the standardization sample (i.e., N = 151 non-neurologically impaired subjects) with a wider age range (i.e., ages 16-88, Crawford et al., 1989a). Subjects were administered the NART and the entire WAIS. The cross-validation study revealed that the NART increased the predicted amount of variance in full WAIS

IQ measures as compared to the standardization study in which a prorated WAIS was used (66% vs. 55% in WAIS FSIQ, 72% vs. 60% in WAIS VIQ, and 33% vs. 32% in WAIS PIQ). Further, addition of quadratic and cubic functions of the NART error score to the regression models did not significantly increase predicted IQ variance, indicating that there were no floor or ceiling effects in the relationship between NART and WAIS IQ. A combination of the standardization sample with the cross-validation sample allowed new regression equations to be built, which predicted 57% of the variance in WAIS FSIQ, 63% in WAIS VIQ and 31% in WAIS PIQ. These studies suggest that NART performance is a reasonably good predictor of WAIS FSIQ and WAIS VIQ, although it appears to be less good at predicting WAIS PIQ.

The NART's ability to measure the construct of intelligence was further established in a study revealing that the NART loaded highly (-0.85) on general intelligence as measured by a principle components analysis in a group of neurologically normal subjects who were administered the full WAIS and the NART (Crawford, Stewart, Cochrane, Parker, and Besson, 1989). Further, two studies have investigated the NART's correlation with the WAIS and the WAIS-R (Crawford, Allan, Besson, Cochrane, & Stewart, 1990, Crawford, Morrison, Jack, Cochrane, Allan, & Besson, 1990). In the first study, WAIS and WAIS-R performance was compared in a U.K. matched samples design, in which 100 pairs of neurologically intact subjects completed the NART and either the WAIS or the WAIS-R (Crawford et al., 1990a). The NART correlated well with both of the current FSIQ measures (r = -0.78 with WAIS FSIQ, r = -0.72 with WAIS-R FSIQ). The second study replicated the design of the first (Crawford et al., 1990c). Fifty-four matched pairs of neurologically intact subjects were administered the NART and either the WAIS or the WAIS-R. The results were consistent with the first study, in that NART correlated highly with both the WAIS FSIQ (r = -0.76) and the WAIS-R FSIQ (r = -0.79). Therefore the NART appears to be a valid measure of both WAIS and WAIS-R FSIQ in the U.K.

Work in Canada has indicated that the NART may also be a valid and reliable measure of intelligence in North America (Blair and Spreen, 1989). The researchers created a Canadian revised NART to overcome difficulties with regard to variations in pronunciation between the U.K. and North America. They administered the original NART plus 54 new words along with the full WAIS-R to a sample of 66 U.S. and Canadian neurologically intact subjects. They performed a series of item analyses to determine the words which correlated best with WAIS-R FSIQ (words

with $r \geq 0.2$). They found a total of 61 words which met this criterion. The revised word list contained 38 of the original NART words plus 23 new words. The researchers found the revised NART to have good reliability in terms of a high internal consistency (alpha = 0.935), and 'virtually perfect' inter-rater reliability ($r = 0.99$). Further the revised NART demonstrated good validity in that it accounted for 56% of the variance in WAIS-R FSIQ.

A second type of construct validity must be demonstrated for the NART, namely, is the NART relatively "dementia-resistant"? Several groups of researchers have attempted to answer this question by comparing WAIS and NART performances of neurologically impaired individuals with neurologically normal individuals. If the NART truly was resistant to the effects of dementia, then the neurologically impaired individuals would perform equally well as control subjects on this test, while they would perform more poorly than control subjects on a measure of current intelligence (e.g., WAIS/WAIS-R).

Nelson and O'Connell (1978) found that patients with evidence of bilateral cortical atrophy received lower scores on all WAIS IQs than the NART standardization sample, yet there was no significant difference between the groups on NART performance. Similarly, a group in the U.S. found that 20 outpatients with mild-moderate DAT performed more poorly than 20 demographically-matched normal elderly volunteers on measures of episodic memory and conscious search of semantic memory, yet there were no significant differences between groups on NART performance (Nebes, Martin, & Horn, 1984). An Australian group also found that their sample of Alzheimer's patients did not significantly differ from cognitively intact control subjects on NART and SGWRT performance, whereas their dementing sample performed significantly more poorly than controls on the Wechsler Memory Scale (Schlosser & Ivison, 1989). Finally, in a study to determine in which conditions of cortical atrophy the NART "held", cognitively intact control subjects' NART scores were compared with NART scores of subjects suffering from differing types of intellectual decline (Crawford, Parker, & Besson, 1988). Subjects with Dementia of the Alzheimer's type (DAT), Multi Infarct Dementia (MID), Alcoholic Dementia, and closed head injury, received NART scores which did not significantly differ from demographically matched neurologically intact control subjects, but those with Huntington's Disease and Korsakoff's Psychosis scored significantly lower than controls, indicating that the NART is a relatively "dementia-resistant" psychometric test for many, but not all disorders involving organicity.

One group of researchers employed discriminant function analysis to determine whether neurologically impaired individuals would perform equally well on the NART as control subjects (Crawford, Hart, & Nelson, 1990). More specifically, they examined the hypotheses that 1. NART by itself would not directly discriminate group (impaired vs. not impaired) and that 2. NART in combination with the WAIS would improve classification accuracy over the WAIS alone. The first series of discriminant function analyses compared a group of 32 dementing patients with 151 "healthy" controls, predicted NART IQs determined by Crawford, Stewart, Parker, Besson, and Cochrane's (1989) regression equations including NART plus demographic variables. The second series of discriminant function analyses compared 40 subjects with CAT scan evidence of cortical atrophy (predicted NART IQs determined by Crawford, Parker, Stewart, Besson, & DeLacey's, 1989, regression equations for prediction of a short form of the WAIS) with the same 151 "healthy" controls (predicted NART IQs determined by Nelson & O'Connell's, 1978, regression equations for prediction of a short form of the WAIS). The researchers found that although the WAIS scales correctly classified a substantial percentage of subjects in both sets of discriminant function analyses, in 5 out of 6 cases the NART significantly improved classification. Further, as expected, the biserial correlation coefficients between NART estimated IQ on its own and group membership were nonsignificant, while the biserial correlation coefficients between group membership and WAIS IQ measures increased in magnitude when NART IQ estimates were partialled out. The researchers concluded that NART on its own did not discriminate the groups, whereas NART in conjunction with the WAIS improved discrimination between impaired and non-impaired individuals over WAIS by itself.

Another method of examining a test's "dementia-insensitivity" is to determine to which degree it correlates with dementia severity. That is, if the test is relatively "dementia-resistant", one would expect that mildly and severely dementing individuals would perform similarly on the index. O'Carroll and Gilleard (1986) compared NART and a half-length version of the Mill Hill Vocabulary Synonym's Scale scores (MHVS, a vocabulary test with fewer demand characteristics than the Vocabulary subtest of the WAIS/WAIS-R), in subjects with varying degrees of dementia. The researchers found no significant correlation between measures of dementia severity and the NART or the MHVS. Further, Crawford and his colleagues (1988 conference abstract cited in Crawford, 1989) found that the Mini Mental State Examination (MMSE), a measure of dementia

severity, was significantly correlated with NART in a sample of Parkinson's patients. However, when the researchers divided their sample into those scoring above and those scoring below the dementia cut-off on the MMSE, they found the MMSE did not significantly correlate with the NART in the dementia subgroup, but did significantly correlate with the NART in the non-dementia subgroup. The authors interpreted this to mean that the MMSE is sensitive to variation in premorbid IQ (i.e., not distribution free), rather than that the NART is significantly correlated with degree of dementia severity.

A further method for establishing validity of a test such as the NART, which purports to be a stable, non-deteriorating measure of intelligence, is the longitudinal study. If the NART is indeed a relatively "dementia-resistant" test, then test scores of individuals with progressive brain deterioration should not decay over time. O'Carroll, Baikie, and Whittick (1987) administered a dementia scale, the NART, and a half-length version of the MHVS to a sample of demenring individuals and re-administered the inventories a year later. They found that the NART was the only test which did not significantly decline at follow up.

Despite all this supportive evidence for the NART, there have been some studies which have not found it to "hold" as well as originally suggested (Brayne & Beardsall, 1990; Hart et al., 1986; Stebbins, Wilson, Gilley, Bernard, & Fox, 1987; Wood, Copeland, Forshaw, Muthu, Abed, Sharma, & Dewey, 1984). One group administered the NART, the MHVS, and several indices designed to detect dementia, to a large randomly selected community sample over the age of 65 (Wood et al., 1984). They divided their sample into normal controls, early dementing individuals and definite dementing individuals on the basis of dementia score. The researchers found that the dementia index correlated negatively with all the psychometric measures, including the NART. However, in a followup study using the same sample (Searle, 1984), the NART was the only measure which did not show significant performance decline for dementing individuals, but this was only the case when probable cases of dementia were considered separately. That is, when possible and probable cases were combined, there were significant differences in NART performance from time one to time two. The authors however, suggested that subjects labelled "possible cases of dementia" may not have been dementing, but may have been cases of long-standing low intelligence.

Brayne and Beardsall (1990) conducted a large-scale community study of 365 elderly women. The researchers compared NART performance and performance on a mini-neuropsychological test battery (the CAMCOG), in women diagnosed as dementing on the Cambridge Mental Disorders of the Elderly Examination (CAMDEX) with those scoring within the normal range on this diagnostic interview. They found that those diagnosed as dementing scored significantly lower on both the CAMCOG and the NART than those diagnosed as normal. However, the 75-79 year olds diagnosed as mild/moderately dementing scored slightly better on the NART than those diagnosed as mildly dementing, although the numbers of subjects were small.

There were several methodological problems with this study. First, the researchers did not adequately control for demographic variability between the groups (dementing vs. control). Thus, if dementing subjects had, for instance, lower education than control subjects, the observed group difference in NART performance might best be explained by poor education of dementing subjects rather than lack of validity of the NART. Secondly, the researchers themselves pointed out that their diagnostic interview (the CAMDEX), included the mini-neuropsychological test battery, the CAMCOG. Although the diagnosis was made prior to calculation of CAMCOG scores, the impression of performance on these tests may have biased the diagnoses towards dementia when in fact, a proportion of those subjects may merely have had long-standing lower intelligence. In other words, some of the "dementing" group may well have had long-standing lower intelligence rather than dementia, and therefore would be expected to score lower on the NART than those classified as normal, due to lower premorbid intelligence.

One study found that the NART estimated premorbid WAIS FSIQ (using the Nelson, 1982, regression equation) significantly distinguished between dementing and normal elderly control subjects, with the patient sample scoring lower than controls on the NART (Hart et al., 1986). However, the researchers compared two other methods of predicting WAIS FSIQ (the Vocabulary subtest of the WAIS and the SGWRT, using Nelson & McKenna's, 1975, regression equation) with the NART, and they found the NART provided the highest estimate of FSIQ. The authors therefore concluded that the NART was not totally resistant to dementia, but was the best premorbid indicator investigated in the study.

A further study which questioned the "holding" abilities of the NART, compared the NART with Wilson and his colleagues' (1979) demographic formula's ability to estimate premorbid intelligence (Stebbins et al., 1987). They found that the NART estimates for their dementing sample (N = 122) differed for moderately and severely dementing individuals than for mildly dementing individuals and controls. However, the authors neglected to mention which procedure for measuring premorbid intelligence (NART vs. Wilson's demographic equation) was superior. They concluded that although the NART may be a promising clinical tool, its applicability to moderately or severely dementing individuals may be limited. Indeed, it is improbable that any psychometric measure will prove to be completely "dementia-insensitive", but the research to date indicates that in comparison with other current methods of estimating premorbid intelligence, the NART seems to be the best. Further, although the results of this study suggest that it is likely that reading abilities may be affected in severely dementing individuals, its primary clinical use is intended to be with patients in the early stages of dementia, where diagnosis is typically a problem. Further, this study was summarized in a brief abstract with limited information, the authors only mentioned that the groups were equated for education. It is therefore impossible to determine if severe, moderate, and mild groups were well matched in terms of other demographic variables known to be related to IQ (e.g., occupation). If, for instance, severely dementing individuals were employed in less prestigious occupations than mildly dementing individuals, the difference in NART scores might be best explained in terms of long-standing lower intelligence in the severely dementing group, rather than a reduced ability of the NART to 'hold' with increased dementia severity.

In answer to the potential problem of the NART not "holding" with some individuals, Crawford and his colleagues have developed demographic equations to predict NART error score (Crawford, Allan, Cochrane, & Parker, unpublished manuscript). That is, it may be useful to estimate NART error score to determine if observed NART performance is worse than expected given an individual's demographic background, particularly if the individual is in the severe stages of dementia. Since demographic variables are correlated with IQ and are completely free from current cognitive ability, they could be used to predict NART error scores. Further, if the obtained NART error score is sufficiently larger than the predicted NART error score, the obtained score should be suspect in terms of underestimating premorbid IQ, assuming both are valid. The

authors administered the NART and collected demographic data (education, age, social class, and sex), from a large sample (N = 659) of cognitively intact subjects. They found a significant multiple correlation between demographic variables and the NART (R = 0.70, p <.0001). Further, they regressed demographic variables against NART error scores to develop regression equations to predict NART IQ. The authors noted that if the discrepancy between obtained and predicted NART error score was greater than 11.4 points, this would be indicative that an individual's NART performance was significantly worse than would be expected from his/her demographic background.

1.2.4.3. NART in Comparison With Other Methods of Distinguishing Dementing from Nondementing Individuals

Some clinicians have used the WAIS Vocabulary subtest age-scaled score to estimate premorbid intellectual functioning. Studies have shown, however, that this index is inferior to measures of word reading ability (i.e., SGWRT and NART) for this purpose, for both dementing and depressed individuals (Crawford et al., 1988a; Crawford, Besson, Parker, Sutherland & Keen, 1987; Hart et al., 1986; Nelson & McKenna, 1975). Nelson and McKenna (1975) compared word reading ability as measured by the SGWRT, with performance on the WAIS Vocabulary subtest in a sample of 98 hospitalized control subjects with extra-cerebral disorders and 45 hospitalized dementing subjects. They found that the mean Vocabulary age-scaled score of the dementing subjects was significantly lower than the mean Vocabulary age-scaled score of the control subjects. In comparison, the mean SGWRT score of dementing subjects was not significantly different from the mean SGWRT score of controls. Further, the researchers regressed both SGWRT and Vocabulary subtest age-scaled scores of control subjects against control subjects' scores on WAIS FSIQ, to create regression equations to predict WAIS FSIQ from either NART error scores or Vocabulary subtest age-scaled scores. Individual subjects' scores on SGWRT and their age-scaled scores on the Vocabulary subtest were then entered into their respective regression equations to predict WAIS FSIQ, and discrepancy scores were calculated (SGWRT predicted FSIQ - WAIS FSIQ and Vocabulary predicted FSIQ - WAIS FSIQ). The results showed that there was less overlap between dementing and control subjects' discrepancy scores when SGWRT was used to predict premorbid WAIS FSIQ than when the Vocabulary subtest was used to predict premorbid WAIS FSIQ. The researchers concluded that

discrepancy between SGWRT predicted FSIQ and WAIS current FSIQ was a better indicator of dementia than the discrepancy between Vocabulary predicted FSIQ and WAIS FSIQ. Unfortunately, Nelson and McKenna (1975) did not control for demographic variability between the groups. This presents difficulties for using the control subjects' data to predict dementing subjects' premorbid IQs, since this procedure assumes sample equivalence. Therefore, Nelson and McKenna's regression equations could be inaccurate, however, as mentioned earlier, Ruddle and Bradshaw (1982) found no significant differences between their regression equation to predict WAIS FSIQ from SGWRT, and that of Nelson and McKenna (1975).

In a study attempting to determine the most efficacious method of determining premorbid intelligence, a sample of DAT outpatients was compared with elderly controls who were functioning independently in the community (Hart et al., 1986). Premorbid IQ estimations were determined by the NART (using Nelson's, 1982, regression equations), the Vocabulary subtest age-scaled score and the SGWRT (both determined by Nelson and McKenna's, 1975, regression equations). Results showed that the NART was the procedure of choice, as it predicted a significantly higher mean WAIS FSIQ for the DAT group than either the Vocabulary subtest of the WAIS or the SGWRT. A later study compared the efficacy of NART estimated WAIS FSIQ (regression equation used not identified) with Vocabulary estimated WAIS FSIQ (using Nelson & McKenna's, 1975, regression equation) in a sample of 39 depressed inpatients and 39 demographically-matched normal control subjects (Crawford et al., 1987). The results indicated that the NART estimated a significantly higher WAIS FSIQ than the Vocabulary subtest in 79% of depressed subjects and the difference remained significant regardless of age (less than or greater than 60 years old). There was no significant difference between NART and Vocabulary predicted WAIS FSIQ in the control group. Further, there was no significant difference between depressed and control subjects on NART performance, whereas the depressed subjects performed significantly more poorly than controls on the Vocabulary subtest. The researchers concluded that NART performance was more "resistant" to the cognitive effects of depression than performance on the Vocabulary subtest of the WAIS. Finally, one group of researchers compared NART estimated WAIS FSIQ (determined by Nelson's, 1982, regression equation) with Vocabulary estimated WAIS FSIQ (determined by Nelson & McKenna's, 1975, regression equation), in a sample divided into six groups of organic conditions (Crawford et al., 1988a). The

results were that the NART estimated a significantly higher premorbid WAIS FSIQ than the Vocabulary subtest in the organic sample as a whole. Further, NART estimated WAIS FSIQ scores produced no significant differences between control subjects and four out of six organic conditions, while Vocabulary estimated WAIS FSIQ scores produced no significant differences between control subjects and only one out of six organic conditions.

In summary the research to date appears to support the view that the NART is a superior estimator of premorbid IQ than the Vocabulary subtest of the WAIS. This difference may be due to the increased cognitive effort required to succeed on the Vocabulary task (defining words) in comparison with the more "automatic" procedure required in the NART (oral pronunciation of short irregular words, Crawford et al., 1987).

Two other commonly used measures of premorbid intelligence are the Schonell Graded Word Reading Test (SGWRT, a word reading test employing long regular words) and the Mill Hill Vocabulary Scale (MHVS, a vocabulary test with fewer demand characteristics than the WAIS/WAIS-R Vocabulary subtest). Both of these tests have been extensively compared with the NART, and in every study the NART has been shown to be superior. Nelson and O'Connell (1978) found that there were no significant differences between dementing patients with EMI scans showing cortical atrophy and normal control subjects on either the NART or the SGWRT, but that the trend in the data was suggestive of impairment in some aspect of reading ability as measured by the SGWRT. Further, the authors concluded that the NART was superior to the SGWRT due to its higher ceiling level and its absence of long regular words, which dementing individuals read significantly more poorly (on the SGWRT) than control subjects.

The MHVS has been compared with the NART in several different laboratories. One group found that age was a significant predictor of MHVS score but not NART score in a "healthy old" community sample (Binks & Davies, 1984). In a set of studies, both the MHVS and the NART initially did not differentiate between subjects with differing degrees of dementia (O'Carroll & Gilleard, 1986), but on one year followup, the MHVS scores declined for the dementing group while the NART scores did not (O'Carroll et al., 1987). Finally, one group of researchers compared several measures purporting to estimate premorbid intelligence: the NART, the SGWRT, and the WAIS Vocabulary subtest (Hart et al., 1986). They found that the NART was

the procedure of choice, as it yielded significantly higher WAIS FSIQ estimates than the other two predictive measures.

Recent work has been carried out in both the U.S. and the U.K. to develop demographic equations to predict premorbid intelligence, as mentioned in a previous section (see 2.1. Demographic variables, above). U.S. (Wilson et al., 1978; Karzmark et al., 1985; Barona et al., 1984; 1986), and U.K. (Crawford et al., 1989b) demographic equations have been shown to predict a fair amount of variance in WAIS and WAIS-R FSIQ (36%-58%) and VIQ (38%-61%), but the NART has been shown to predict a larger amount of variance in WAIS FSIQ (55%-66%) and VIQ (60-72%, Crawford et al., 1989a; Nelson, 1982). Neither demographic equations nor the NART have been shown to be particularly effective at predicting variance in WAIS/WAIS-R PIQ (24%-42% and 31%-33%, respectively, Barona et al., 1984, Crawford et al.; 1989a; b, Karzmark et al., 1985; Wilson et al., 1978; Nelson, 1982).

1.2.5. NART in Combination With Demographic Equations

The current direction of research in the U.K. is combining psychometric and demographic approaches, in attempts to increase the predictive accuracy of premorbid IQ measures (Crawford et al, 1989d). The logic of this approach is as follows: There is considerable covariance between the NART and demographic variables, (for instance, education is highly correlated with IQ). Therefore, combining these variables in a regression equation will not have an additive effect on the IQ variance predicted. However, it is still possible that such a combination will have a cumulative effect. That is, some of the variance in either set of variables will not be shared but may still predict IQ.

Crawford and his colleagues (1989d) determined the predictive accuracy of the psychometric and demographic methods combined, using a sample of 151 cognitively normal subjects. They found that the NART was the single best predictor of IQ, however, the addition of demographic variables significantly increased the predictive accuracy, the combined approach accounting for 7% more variance in WAIS FSIQ (73%) and 6% more variance in WAIS VIQ (78%) and PIQ (39%). A study to determine the construct validity of the combined premorbid IQ equation (NART + demographics) indicated that the combination loaded very highly on general

intelligence (g = 0.90) as measured by a principal component analysis (Crawford, Cochrane, Besson, Parker, & Stewart, 1990). The combined equation had a g-value higher than the NART alone (g = 0.85). In fact, its g-loading was higher than any of the individual WAIS subtests. Crawford and his colleagues (1989d) revealed that a discrepancy of 15 IQ points between observed and predicted FSIQ (determined by the combined equation) was found in only 1% of the normal population, thus a discrepancy of this size could be considered as being highly suggestive of cognitive deterioration.

Not all research has found that NART combined with demographic information results in improved premorbid IQ estimation. A group in Canada determined the predictive accuracy of their revised North American NART in combination with demographic variables (Blair and Spreen, 1989). The researchers regressed NART error scores with demographic information (age, education, race, sex, occupation, handedness, and region of residence, following Barona et al., 1984) against WAIS-R FSIQ to determine whether demographic variables in combination with the NART improved the predictive accuracy over the NART on its own. The Canadian group found that when they added the demographic variables with the NART-R into the regression equation predicting WAIS-R FSIQ, the demographic variables added only 3% of predicted WAIS-R FSIQ variance. The researchers concluded that demographic variables did not significantly improve predictive accuracy in their sample. It is difficult to compare these Canadian results with the U.K. results, however, since the U.K. NART and the North American NART-R contained different stimuli. In addition, the U.K. study used the WAIS as the comparative measure while the North American study used the WAIS-R. Further, different demographic data was used (e.g., handedness, region of residence and race were not used in the U.K. demographic equations developed by Crawford et al., 1989b). Clearly more work is required in this area to determine the benefits of combining demographic and psychometric approaches in predicting premorbid IQ.

1.3. Summary

In summary, clinicians would benefit greatly from reliable and valid psychometric instruments which can help differentiate dementing from nondementing individuals. WAIS/WAIS-R algorithms were some of the first psychometric attempts to aid in diagnosis of dementia, but most have demonstrated little success. The best methods to date appear to be those which

attempt to estimate premorbid intelligence (i.e., word reading abilities, demographic information, and the Vocabulary subtest of the WAIS/WAIS-R). Comparative research favours the NART as the procedure of choice (e.g., Crawford et al., 1989d; Hart et al., 1986). The NART appears to have a respectable history of studies supporting its reliability and validity as an estimator of premorbid intelligence. However, most of the studies to date have used U.K. samples. Before the NART can be used in Canada with confidence it is necessary that it be validated on a Canadian sample. This is particularly important considering that scoring depends on pronunciation of words, which will likely vary with regional accents. Further, some British spellings and pronunciations may be different than those used in North America. At least one item on the NART has ceased to be a word in North America: The word 'gaoled' is spelt 'jailed' in North America. In addition the NART has been almost solely validated against the WAIS, which was standardized in 1955. Therefore it is necessary to compare the NART with the more recent WAIS-R. This study will examine the validity of the the NART as an estimator of premorbid intelligence, using the WAIS-R as the measure of current IQ, and using a Newfoundland sample of dementing and normal elderly adults.

1.4. Experimental Hypotheses

The following six hypotheses are derived from the experimental literature reviewed above:

(1) NART errors will be significantly correlated with, and will therefore predict a significant amount of variance in WAIS-R IQ (FSIQ, VIQ and PIQ) in the control sample, the correlation between NART errors and PIQ being the lowest.

(2) There will be a stronger correlation between WAIS-R scores and dementia score than between NART error score and dementia score.

(3) All WAIS-R measures will more clearly differentiate group (dementing vs. control) than NART error scores.

(4) The NART will estimate a higher WAIS-R FSIQ and VIQ in dementing subjects than the Vocabulary subtest of the WAIS-R, using regression equations determined from the control subjects in the present sample.

(5) NART error scores and Vocabulary age-scaled scores in combination with demographic information, will estimate higher WAIS-R IQs than NART or Vocabulary IQ estimates alone.

(6) Coolidge and his colleagues' (1985) equation (Vocab $\geq$ 2 BI.D = dementia) will correctly classify a statistically and clinically significant proportion of dementing and control subjects. On the other hand, the V-P Split will not significantly distinguish between groups (dementing vs. control).

METHODS

2.1. Subjects

In total, 20 dementing individuals with a diagnosis of either Dementia of the Alzheimer's Type (DAT) or Multi Infarct Dementia (MID) were compared with 20 cognitively intact elderly controls (see Table 1 in Results section for descriptor variable summary statistics by group, dementing subjects vs control subjects). Information from patient files was used by the attending physician to complete the revised eight-item Hachinski Index, an objective test to distinguish DAT from MID (Rosen, Terry, Fuld, Katzman, & Peck, 1980, see Research Instruments below). The revised Hachinski Index revealed that 16/20 of the dementing subjects were probably suffering from DAT (definite diagnosis is only possible at autopsy), and 4/20 were probably suffering from MID. Subjects had a mean age of 75.93 years, dementing individuals (D.I.) ranging in age from 59-89 years (M = 78.95, S.D. = 7.33) and control subjects (C) ranging in age from 69-82 years (M = 72.90, S.D. = 4.18). All subjects were caucasian, and all spoke English as their mother tongue. The D.I.'s were 17 females and 3 males and the C's were 14 females and 6 males. Subjects were better educated than would be expected for this age group in Newfoundland. That is, number of years of full-time education ranged between 0-25 years, D.I.s ranging in education from 3-25 years (M = 10.55, S.D. = 5.12) and Cs ranging in education from 0-20.25 years (M = 11.90, S.D. = 5.49). However, these are only approximate figures since subjects commonly preferred to report last grade finished, rendering it necessary for the author to estimate number of years of education. The criteria applied were as follows: each school grade was equivalent to one year, "highschool" was equivalent to 12 years, post-secondary education included 12 years for highschool plus one year for each year of full-time college or university training, and each completed post-secondary course was equivalent to 0.25 years, as recommended by Crawford and his colleagues (1989b). Occupational status was determined by the Office of Population, Censuses and Surveys scale (OPCS, 1980). The scale includes five broad categories, and each

person receives a number from 1 (professional) to 5 (unskilled). In the case of women, their occupations prior to marriage were used if they did not maintain a career during marriage, and if they had never been employed in the work force, their husbands' occupations were recorded. At least one subject in each group (D.I. vs. C) scored within each of the five categories, and in general, the subjects held higher occupational codes than would be expected in this age-cohort in Newfoundland. Seventy-five percent (15/20) of subjects in both conditions were taking some form of medication, but none were on a regime likely to interfere with cognitive performance as judged by the referring physician.

2.1.1. Subject selection: Place of recruitment

Dementing subjects were recruited attenders at the psycho-geriatric day hospital at the Leonard A. Miller Center, St. John's, Newfoundland (N=6 daypatients), and outpatients from the caseloads of a St. John's based geriatric psychiatrist (N=8 outpatients) and a geriatric general practitioner (N=5 outpatients). In addition, due to difficulties in recruiting subjects (see 2.1.3. Subject selection: Unusable subjects, below), one dementing individual was tested on her second, third, and fourth day of admission to a geriatric inpatient ward at the Miller Center. Control subjects were recruited from the medical Geriatric Day Hospital at the Miller Center (N=3 day patients), from the day care at St. Luke's (old folks) Homes, St. John's, Newfoundland (N=2 daypatients), and from two general practitioners (N=7 outpatients). Further, due to the difficulty in recruiting appropriate control subjects, some C's were recruited from the subject pool of the Gerontology Clinic at Memorial University of Newfoundland (N=8 nonpatients). In summary the final dementing sample consisted of 6 dementing daypatients, 13 dementing outpatients and 1 dementing inpatient. The control sample consisted of 5 medical daypatients, 7 medical outpatients and 8 nonpatients.

2.1.2. Subject selection: Recruitment criteria

Initially, an age criterion of $\leq$ 84 years was imposed, in keeping with Binks and Davies' (1984) finding that the NART was not age-sensitive up to age 84 in their study. However, due to the difficulty in finding subjects within this age range, and since a cross-validation study of the NART by Crawford and his colleagues (1989a) revealed that the test was not age-sensitive to the late 80's, and a further study revealed that there was no curvilinear relationship between age and

NART performance (Crawford et al., 1988b), this criterion was adjusted and subjects up to age 89 were accepted. No subject had a previous psychiatric history, as defined by prior contact with a psychiatrist or a psychologist. This was an attempt to ensure elimination of subjects presenting with 'pseudo dementia' due to depressive illness, or subjects with other psychological disorders which might interfere with cognitive functioning. Further, to exclude subjects with possible Korsakoff's Psychosis or Alcoholic Dementia, no subjects were recruited if they had an alcohol dependence history. Subjects were also excluded if they had a history of head injury that resulted in post-traumatic amnesia or coma. All subjects had adequate visual acuity and were literate as measured by a five-item practice reading test (see Appendix-B). In addition, only subjects who had adequate orientation, test motivation, and task comprehension were included. Criteria specific to group (dementing vs. control) included entry into the dementia category only if subjects met DSM-III-R criteria for dementia, and inclusion into the control category only if subjects were free from a history of strokes and scored greater than 7 on the abbreviated Dementia Scale (Qureshi & Hodkinson, 1974, see 2.2. Research Instruments below). An attempt was made to compare only day- and outpatient dementing subjects with day- and outpatient medical control subjects. This was first to collect a control sample which best matched the dementing sample in terms of the stress associated with having to seek medical attention. Second, day- and outpatients were preferred to inpatients since this increased the likelihood that dementing subjects would be in the mild to moderate range of dementia, which in turn would increase the likelihood that they would be able to comprehend and follow task instructions. However it was not possible to strictly adhere to this criterion and 8 nonpatient control subjects and one inpatient dementing subject were included as mentioned above (see 2.1.1. Subject selection: Place of recruitment).

2.1.3. Subject selection: Unusable subjects.

Two geriatric general practitioners, one family general practitioner, one geriatric psychiatrist, one social worker, and one research coordinator agreed to refer appropriate subjects to the study. Despite efforts to adhere to the exclusion criteria adopted by the author (see 2.1.2. Subject selection: Recruitment criteria above), 11 dementing individuals (D.I.) and 12 control subjects (C) proved to be either inappropriate referrals or dropped out of the study prior to completion of the assessment. Specifically, seven D.I.s did not meet the exclusion criteria

(outlined above): Two dementing subjects had suffered a head injury in the past, three had a psychiatric history and two could not understand test directions. In addition, four D.I.'s who met the appropriate criteria were unable or unwilling to complete the protocol. Three of these individuals were unable to complete the protocol within one sitting, and subsequently two stated that they were too busy to continue with the study, and the other was unable to be re-contacted. One dementing individual's consent form was signed by his wife, but he subsequently refused to be tested. In addition, 11 C's did not meet the exclusion criteria: Three C's had a psychiatric history, five had experienced a head injury in the past, one had a history of stroke, and two had a combination of two of the above. Finally, one C who met the appropriate exclusion criteria was unable to complete the protocol on the first sitting, and was too busy to continue with the study when re-contacted.

2.2. Research Instruments

2.2.1. Practice Reading Test

Before administration of the NART and in order to test visual acuity, ability to follow task instructions, and literacy, subjects were given a list of 5 short regular words to read aloud. This word list and the NART (see below) were presented in the largest print possible (each letter was approximately 5 mm in height) to reduce errors due to poor vision. Use of this practice reading test enabled any NART errors to be attributed to incorrect responses rather than difficulties with task performance (see Appendix-B)

2.2.2. National Adult Reading Test, Nelson, 1982

The test consists of 50 short irregular words which the subject must pronounce, and the number of errors is recorded (see Appendix-C). Nelson (1982) reported regression equations in the NART manual, derived from the NART standardization sample (N = 120 neurologically intact subjects from the U.K.), into which NART error scores are entered to derive a predicted WAIS IQ score. These equations were not used in this study, however, since the WAIS-R was used in the present study, and due to the impracticality of using U.K.-derived equations with a North American sample. That is, one cannot assume that the relationship between word pronunciation and IQ remains the same in the U.K. as in North America. Instead, estimated premorbid WAIS-R

FSIQ and VIQ (henceforth termed NART FSIQ and NART VIQ, respectively) were determined for dementing individuals using regression equations derived from the control subjects in this study (see 3.3.5. Comparison of premorbid IQ measures in the dementing sample below). Predicted WAIS-R PIQ was not determined from NART error scores for dementing subjects, since Crawford (1989) and Blair and Spreen (1989) have reported that the NART is a poor predictor of WAIS/WAIS-R PIQ. However, for the sake of comparison the NART error scores were regressed against WAIS-R PIQ using data from the control subjects in this study, to determine the amount of variance in WAIS-R PIQ predicted by the NART.

2.2.3. Wechsler Adult Intelligence Scale-Revised, Wechsler, 1981

Due to the impracticality of required repeat visits for completion of the lengthy study protocol, only the 7 Wechsler subtests used in Nelson's standardization of the NART (Nelson, 1982), were used in the present study. These subtests included Picture Completion, Digit Span, Picture Arrangement, Vocabulary, Block Design, Arithmetic, and Similarities, always presented in that order. Following Nelson (1982), estimated WAIS-R FSIQ, PIQ and VIQ were obtained by prorating the seven subtests.

Since the Vocabulary subtest is considered to be the best "hold" subtest (i.e., least "dementia-sensitive") of the WAIS-R (Lezak, 1983) and previous studies have compared its accuracy as a premorbid IQ estimator with the NART (e.g., Hart et al., 1986; Crawford et al., 1988a), predicted Vocabulary WAIS-R FSIQ and VIQ (henceforth termed Vocabulary FSIQ and Vocabulary VIQ, respectively) were determined for the dementing sample. Although regression equations for the conversion of Vocabulary age-scaled scores to predicted WAIS IQs have been generated in the literature (Nelson & McKenna, 1975), these equations were derived on a U.K. sample using the WAIS. Therefore, as in the case of the NART (see 2.2.2. National Adult Reading Test above), control subjects' age-scaled Vocabulary scores were regressed against WAIS-R FSIQ and VIQ scores. Dementing subjects' Vocabulary age-scaled scores were then entered into the generated regression equations to determine their predicted premorbid Vocabulary FSIQ and VIQ scores. Further, as in the case of the NART, Vocabulary estimated PIQ was not determined for dementing subjects since Crawford (1989) reported that Vocabulary age-scaled scores are poor predictors of WAIS PIQ. However, Vocabulary age-scaled scores

were regressed against WAIS-R PIQ in the control sample (N=20) to determine the amount of the WAIS-R PIQ variance predicted by the Vocabulary subtest age-scaled score.

Three further measures were obtained from subtests of the WAIS-R. These included the V-P Split, age-scaled Vocabulary subtest score and the age-scaled Block Design subtest score. The V-P Split was calculated since previous reports in the literature have suggested that dementing subjects tend to demonstrate a VIQ-PIQ discrepancy in favour of VIQ (e.g., Miller, 1977). The V-P Split was determined by simply subtracting the prorated WAIS-R PIQ from the prorated WAIS-R VIQ to obtain a difference score for each subject. Vocabulary and Block Design age-scaled scores (Vocab and Bl.D., respectively), were entered into Coolidge's formula (Vocab $\geq$ 2Bl.D. = dementia) which purports to differentiate dementing from nondementing individuals (Coolidge et al., 1985).

2.2.4. Demographic Questionnaire

The demographic questionnaire (see Appendix D) included information about age, sex, race, education, occupation, and history of psychiatric care, alcohol dependence, stroke and brain injury. Wherever possible, this information was taken from patients' hospital charts, to reduce inaccuracy due to memory deficiencies. Where chart information was lacking, dementing patients' relatives were contacted to fill in the necessary information. Control subjects were asked to provide the information themselves.

Due to the number of studies which have recently investigated the use of demographic variables as premorbid IQ indicators (e.g., Barona et al., 1984; Crawford et al., 1989b), and given the recent evidence that NART performance may be related to demographic factors (Crawford et al., 1988b), four demographic variables (age, sex, occupation, and education, as used by Crawford et al., 1989b) were regressed against WAIS-R FSIQ, VIQ, and PIQ in this sample of control subjects (N=20). This was carried out to determine the amount of variance demographic variables predicted in measures of IQ (i.e., WAIS-R FSIQ, VIQ and PIQ), in comparison with the amount of variance predicted in the same measures by NART error scores and Vocabulary age-scaled scores (see Results section below). Further, the demographic variables were regressed against WAIS-R IQ measures in combination with NART and in combination with Vocabulary

age-scaled scores, to determine whether or not they significantly added to the amount of predicted variance in WAIS-R IQ. (Predicted demographic IQ scores were not calculated in the present study since Crawford (1989) suggested that misleading high or low correlations may be obtained if demographic predictors including multiple variables are regressed on current IQ measures using small samples).

2.2.5. Abbreviated Dementia Scale, Qureshi & Hodkinson, 1974

This scale contains 10 items including tests of information, memory and concentration, and allows the administrator to "stage" dementia approximately. The items were administered in an interview format. A recent study comparing the shortened version (Qureshi & Hodkinson, 1974) with the lengthier 34-item version (Blessed, Tomlinson, & Roth, 1968) indicated adequate correlations between the two (Pearson's r=0.91-0.96, Thompson & Blessed, 1987), thus the briefer version was adopted for this study (see Appendix-E). A cut-off score of seven on this scale was applied to control subjects as a precaution against including undiagnosed dementing individuals in the control category. That is, control subjects were only accepted into the present study if they scored seven or greater on this scale, but dementing individuals were accepted even if they scored greater than seven. This was in accordance with Thompson and Blessed's (1987) findings that some patients with a definite diagnosis of dementia were still able to score greater than seven on the scale.

2.2.6. Revised Hachinski Index, Rosen et al., 1980

This scale contains a list of 8 signs or symptoms used to distinguish between the two dementing conditions of Dementia Alzheimer's Type (DAT) and Multi Infarct Dementia (MID). Each sign or symptom is assigned a score of 1 or 2, yielding a scale range from 0-12. A score from 0-2 is taken as evidence of DAT, and a score of 4-10 is taken as evidence of MID (Rosen et al., 1980). A score in between these ranges (i.e., 3) is undifferentiated. Rosen and his colleagues (1980) attempted to determine the accuracy of the original 13 symptom checklist (Hachinski, Iliff, Zilhka, DuBoulay, McAllister, Marshall, Russell, & Symon, 1975). They found that only eight signs or symptoms were primarily characteristic of vascular dementia as determined by neuropathological analysis of the brains of deceased subjects. Therefore, the revised version was adopted and completed by the referring physician to provide a more valid means of defining the subject sample (see Appendix-F).

2.3. Procedure

Subjects' doctors distributed a consent form for research candidates to sign. An information sheet accompanied the consent form explaining the purpose of the project and outlining candidates' option to refuse or withdraw participation at any time (see Appendix A). In the case of those subjects who were seriously handicapped in their ability to give informed consent (as decided by the attending physician), a relative or caretaker who was legally authorized to give informed consent on behalf of the individual was contacted by the physician. Completed consent forms were returned to the principle investigator prior to the beginning of the testing period. All assessments were carried out by the author. Subjects attending day hospitals were tested in a quiet, well-lit interview room, while outpatients and nonpatients were tested in their homes. The testing protocol was as follows: All subjects were first administered the abbreviated Dementia Scale. Next, subjects were asked to fill in demographic information which was later confirmed from their records. If the records did not contain the appropriate information, or if no records existed, C's were simply asked for the information. In the case of D.I.'s, caretakers provided the information following the testing period. Subjects were then presented with the practice reading test, followed by the list of NART words. Participants were asked to read through them at their own pace, according to directions set out in the NART manual (Nelson, 1982, p.5). Each person was given a blank card to place over the words not yet read, to ensure there would be no missed or repeated words (as suggested by Hart et al., 1986). In the case where subjects were not able to move the card by themselves, the tester moved it for them. The subjects' responses were recorded on a portable cassette recorder for later scoring. (The tester also scored the NART at the time of testing as a precaution against possible mechanical difficulties with the tape recorder). Subjects were administered the WAIS-R according to directions set out in the manual (Wechsler, 1981, pp. 59-86). An attempt was made to administer all tests on the same day, with allowance of short breaks if the test administrator sensed subject uneasiness, poor attention, or fatigue. In three cases (2 D.I., 1 C.) it was impossible to complete the protocol in one sitting, and the testing for these subjects was completed within a two week period, as recommended by Hart et al. (1986). The revised Hachinski Index was completed for all dementing individuals by the attending physician as soon as possible following the testing period. The NART was scored independently from audiotapes by two individuals; one clinical

psychologist experienced with scoring of the NART and one inexperienced scorer. The former was blind to condition by randomization of tapes and test booklets using a random numbers table. NART and WAIS-R were coded separately to avoid bias in scoring due to knowledge of one score or the other.

RESULTS

2.1. General Information

Table 1 provides summary statistics on each of the demographic/controlling variables measured for both Dementing and Control subjects. Table 2 provides summary statistics on each of the cognitive variables for Dementing and Control subjects. All statistics used to compare groups (D.I. vs. C.) were computed with the amount of variance caused by the demographic variables accounted for. This was carried out to avoid the methodological flaw pointed out by Crawford (1989), that demographic variables (i.e., education, occupation, and age) are themselves related to premorbid IQ test performance (Crawford, 1989), and therefore predict a substantial proportion of IQ variance. It was decided not to match the groups on demographic variables due to the distortion of true variances in the real world which can occur by using this method (Kerlinger & Pedhazur, 1973, pp. 82-83), and therefore the demographic differences between the two groups were controlled for in each statistical procedure.

Table 1
Demographic Information By Group: Means, Standard Deviations, Simple t-scores and Multiple Regression t-scores For Education and Age, and Number in Each Category of Occupation and Sex.

Variable	D.I.[a]	C.[b]	t-score	Reg-t[c]
Education	10.55 ± 5.12	11.90 ± 5.49	0.80	1.34
Age	78.95 ± 7.33	72.90 ± 4.18	-3.21**	2.89**
Occupation				
1(professional)	3	4		
2(intermediate)	2	7		
3(skilled)	4	6		
4(semi-skilled)	8	2		
5(unskilled)	2	1		
missing values	1	0		
Sex				
male	3	6		
female	17	14		

[a] Dementing Individuals [b] Control Subjects

[c] t-score from multiple regression part correlations, demographic variables partialled out

**p<.01

Table 2
Cognitive Measures By Group: Means, Standard Deviations, Simple t-scores, and Multiple Regression t-scores

Variable	D.I.[a]	C.[b]	t-score	Reg-t[c]
Dementia Scale (0-10 scale)	5.02 ± 2.19	9.12 ± 0.89	7.77***	-5.53****
NART Errors	35.15 ± 10.70	27.30 ± 10.18	-2.38*	1.94
WAIS-R FSIQ	74.45 ± 9.14	96.95 ± 13.15	6.28***	-5.29****
WAIS-R VIQ	77.65 ± 8.74	97.95 ± 15.12	5.20***	-4.60***
WAIS-R PIQ	72.50 ± 10.58	95.60 ± 12.04	6.45***	-4.31***
V-P Split	5.15 ± 9.03	2.35 ± 14.17	-0.75	-0.36
Vocab Score	6.70 ± 2.58	10.40 ± 3.50	3.81***	-3.03**
Bl.D Score	4.75 ± 2.12	9.05 ± 2.58	5.75***	-3.60**

[a] Dementing Individuals
[b] Control Subjects
[c] t-score from multiple regression part correlations, demographic variables partialled out

*$p<.05$ **$p<.01$ ***$p<.001$ ****$p<.0001$

2.2. Descriptive Information

2.2.1. Group comparison on demographic variables

Multiple regression analyses were performed to determine the amount of variance accounted for by demographic variables in determining group membership (D.I. vs. C., see Table 1). Of the four demographic variables measured (age, sex, education, and occupation), age was the only one which significantly differentiated between groups when the variance associated with the other demographic variables was controlled for. Dementing subjects were significantly older than control subjects ($t=2.89$, $p<.01$). As mentioned above, the age difference between groups was statistically controlled for such that it would not interfere with other statistical comparisons between groups.

2.2.2. Group comparison on dementia scale

As expected, a multiple regression analysis revealed that dementia score significantly predicted group membership, even with the demographic variables controlled for ($t=-5.53$, $p<.0001$). In other words, dementing subjects scored significantly lower on the dementia scale. Although age significantly predicted dementia score ($t=-3.29$, $p<.01$), this effect disappeared when Type (i.e., demented individuals vs. controls) was entered in the equation, ($t=-1.58$, $p>.05$). This indicates that although age was positively correlated with dementia score in this sample, the relationship can be explained by the significant age difference between dementing subjects and control subjects.

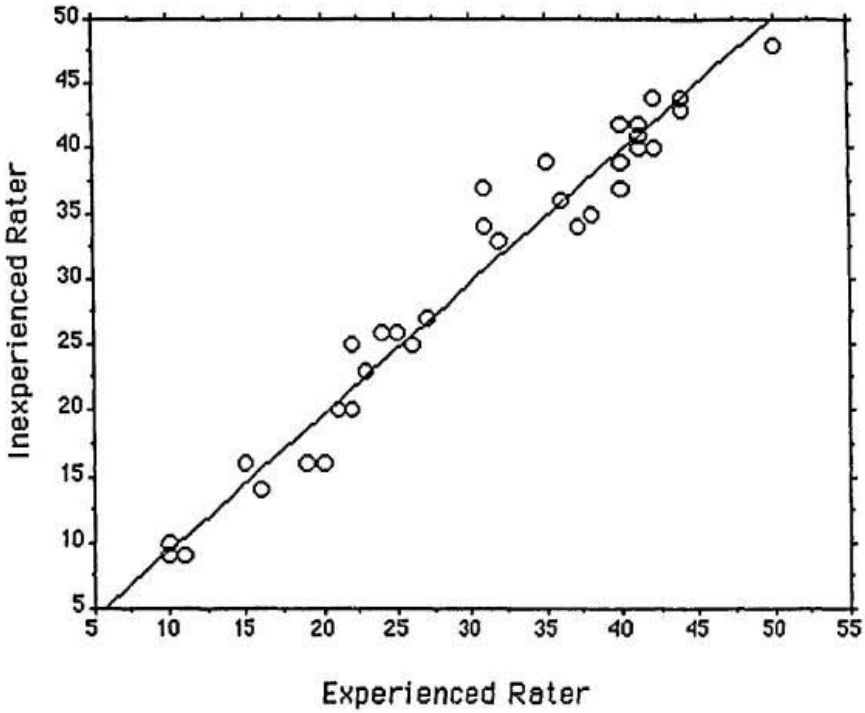
2.3. Main Analyses

2.3.1. Inter-rater reliability of the NART

Due to the possible difficulty in understanding regional accents, to the author's inexperience in scoring the NART, and to the lack of complete blindness of the author, inter-rater reliability of NART error scores was determined between the two raters who independently scored the audiotape of subjects' performance on the NART. One was experienced and blind to group status, and one was inexperienced and partially blind to group status. The two sets of ratings of NART error scores were significantly correlated ($r=0.96$, $p<.001$, see Figure 1), and

Figure 1

Inter-rater Reliability for NART Error Scores, Experienced vs. Inexperienced Rater.



there was no significant difference between the scores($t=0.38$, $p>.05$). Due to the extremely high inter-rater reliability, the author's scores were used for the remaining analyses.

2.3.2. Correlations between NART errors and WAIS-R IQ

Analyzing the entire sample as a whole, NART errors were significantly correlated with all measures of WAIS-R IQ, were most highly correlated with WAIS-R verbal IQ measures ($r=-0.75$, $p<.001$), followed by WAIS-R FSIQ ($r=-0.73$, $p<.001$), and least highly correlated with WAIS-R performance IQ measures ($r=-0.57$, $p<.001$, see correlation matrix for the total sample in Table 3). As expected, these correlations increased when control subjects were examined in absence of dementing subjects (see correlation matrix for control subjects in Table 4), and decreased when dementing subjects were examined in absence of control subjects (see correlation matrix for dementing subjects in Table 5). For example, the correlation between NART errors and WAIS-R FSIQ for control subjects was $r=-0.76$, $p<.001$, with the groups combined was $r=-0.73$, $p<.001$, and for dementing subjects was $r=-0.70$, $p<.001$.

2.3.3. Relationships between WAIS-R IQ, NART error scores, and dementia score

Simply from the correlation matrix for dementing subjects, dementia score was significantly correlated with NART error score, WAIS-R FSIQ, VIQ, and PIQ (see Table 5). However, using multiple regression analyses with dementia score as the dependent variable and NART errors and demographic variables (the latter partialled out) as independent variables, NART errors did not significantly predict dementia score for either control subjects ($t=1.67$, $p>.05$) or dementing subjects ($t=-0.02$, $p>.05$). In comparison, multiple regression analyses with dementia score as the dependent variable and WAIS-R FSIQ and demographic variables (the latter partialled out) as independent variables, revealed that WAIS-R FSIQ did not significantly predict dementia score for control subjects ($t=0.31$, $p>.05$), but did significantly predict dementia score for dementing subjects ($t=3.01$, $p<.01$).

Table 3
Correlations Between All Variables For Total Sample

	NARTERR	FSIQ	VIQ	PIQ	VPSpl	Vocab	BL.D	Demen	Sex	Edu	Occ	Age
NARTERR	1.00	-0.73***	-0.75***	-0.57***	-0.23	-0.84***	-0.54***	-0.47**	-0.31*	-0.83***	0.79***	0.20
FSIQ	-	1.00	0.96***	0.88***	0.09	0.89***	0.85***	0.77***	-0.17	0.56***	-0.60***	-0.48**
VIQ	-	-	1.00	0.73***	0.35*	0.92***	0.73***	0.67***	-0.09	0.59***	-0.59***	-0.40**
PIQ	-	-	-	1.00	-0.38**	0.70***	0.92***	0.81***	-0.26	0.43**	-0.55***	-0.59***
VPSpl	-	-	-	-	1.00	0.27*	-0.29*	-0.20	0.23	0.21	-0.05	0.27*
Vocab	-	-	-	-	-	1.00	0.66***	0.60***	-0.17	0.70***	-0.69***	-0.37*
Bl.D	-	-	-	-	-	-	1.00	0.72***	-0.19	0.34*	-0.53***	-0.45**
Dement	-	-	-	-	-	-	-	1.00	-0.29*	0.32*	-0.44**	-0.49**
Sex	-	-	-	-	-	-	-	-	1.00	-0.21	0.37*	0.09
Edu	-	-	-	-	-	-	-	-	-	1.00	-0.74***	-0.14
Occ	-	-	-	-	-	-	-	-	-	-	1.00	0.18
Age	-	-	-	-	-	-	-	-	-	-	-	1.00

*p<.05 **p<.01 ***p<.001

Table 4
Correlations Between All Variables For Control Subjects

	NARTERR	FSIQ	VIQ	PIQ	VPSpl	Vocab	BL.D	Demen	Sex	Edu	Occ	Age
NARTERR	1.00	-0.76***	-0.81***	-0.39*	-0.53**	-0.89***	-0.51*	-0.09	-0.10	-0.81***	0.70***	0.16
FSIQ	-	1.00	0.96***	0.70***	0.43*	0.88***	0.75***	0.31	0.12	0.70***	-0.56**	-0.31
VIQ	-	-	1.00	0.48*	0.66**	0.90***	0.59**	0.22	0.22	0.70***	-0.55**	-0.24
PIQ	-	-	-	1.00	-0.34	0.54**	0.89***	0.49*	-0.06	0.47*	-0.48*	-0.44*
VPSpl	-	-	-	-	1.00	0.50*	-0.13	-0.18	0.29	0.35	-0.18	0.12
Vocab	-	-	-	-	-	1.00	-0.61**	0.19	0.11	0.81***	-0.69***	-0.39*
Bl.D	-	-	-	-	-	-	1.00	0.35	-0.12	0.43*	-0.56**	-0.24
Dement	-	-	-	-	-	-	-	1.00	-0.03	0.33	-0.36	-0.20
Sex	-	-	-	-	-	-	-	-	1.00	0.12	0.28	-0.26
Edu	-	-	-	-	-	-	-	-	-	1.00	-0.70***	-0.30
Occ	-	-	-	-	-	-	-	-	-	-	1.00	0.25
Age	-	-	-	-	-	-	-	-	-	-	-	1.00

*p<.05 **p<.01 ***p<.001

Table 5
Correlations Between All Variables For Dementing Subjects

	NARTERR	FSIQ	VIQ	PIQ	NARTIQ	VocIQ	VPSpl	Vocab	BL.D	Demen	Sex	Edu	Occ	Age
NARTERR	1.00	-0.70***	-0.69***	-0.59**	-1.00	-0.78***	0.03	-0.77***	-0.35	-0.45*	0.72***	-0.90***	0.83***	-0.02
FSIQ	-	1.00	0.91***	0.85***	0.70***	0.84***	-0.12	0.84***	0.65**	0.72***	-0.37	0.65**	-0.61**	-0.25
VIQ	-	-	1.00	0.58**	0.69***	0.88***	0.29	0.89***	0.37	0.61**	-0.38	0.67**	-0.60**	-0.11
PIQ	-	-	-	1.00	0.59**	0.57**	-0.61**	0.57**	0.80***	0.69***	-0.37	0.52**	-0.53*	-0.44*
NARTIQ	-	-	-	-	1.00	0.78***	-0.03	0.77***	0.35	0.45*	-0.72***	0.90***	-0.83***	0.02
VocIQ	-	-	-	-	-	1.00	0.18	1.00	0.28	0.54**	-0.45*	0.67**	-0.64**	-0.02
VPSpl	-	-	-	-	-	-	1.00	0.19	-0.58**	-0.23	0.07	0.04	0.04	0.40*
Vocab	-	-	-	-	-	-	-	1.00	0.28	0.54**	-0.44*	0.67**	-0.64**	-0.02
Bl.D	-	-	-	-	-	-	-	-	1.00	0.52*	-0.05	0.26	-0.35	-0.21
Dement	-	-	-	-	-	-	-	-	-	1.00	-0.42*	0.41*	-0.34	-0.25
Sex	-	-	-	-	-	-	-	-	-	-	1.00	-0.63**	0.42*	0.21
Edu	-	-	-	-	-	-	-	-	-	-	-	1.00	-0.82***	0.03
Occ	-	-	-	-	-	-	-	-	-	-	-	-	1.00	-0.06
Age	-	-	-	-	-	-	-	-	-	-	-	-	-	1.00

*p<.05 **p<.01 ***p<.001

2.3.4. Distinctions between groups by WAIS-R IQ measures and NART errors

Several multiple regression analyses using Type (i.e., D.I. vs. C.) as the dependant variable, revealed that all current measures of intelligence (WAIS-R FSIQ, WAIS-R VIQ, and WAIS-R PIQ) significantly distinguished between the groups, when the variance associated with the demographic variables was controlled for (see Table 2). In comparison, there was no significant difference between groups on the V-P Split, which is a measure intended to distinguish between groups. Similarly, there was no significant difference between the performance of dementing and control subjects on the NART when demographic variables were controlled for, which is not intended to distinguish between groups. However, it is interesting to note that performing a simple t-test without taking demographic variability into consideration revealed significant differences between groups on NART performance (see Table 2).

2.3.5. Comparison of premorbid IQ measures in the dementing sample

Data from control subjects was used to calculate regression equations to predict premorbid WAIS-R FSIQ and premorbid WAIS-R VIQ scores for dementing individuals, from NART errors and Vocabulary age-scaled scores. Equations to predict NART IQs were generated by regressing NART errors against WAIS-R FSIQ and WAIS-R VIQ in the control sample:

NART FSIQ = 123.92 - 0.99 x NART error score, (Standard Error of Estimate, S.E.E. = 8.70).

NART VIQ = 130.76 - 1.20 x NART error score, (S.E.E. = 9.13).

Although the present regression equations were derived on a small sample (N=20), they appear comparable to the original regression equations cited in the NART manual (Nelson, 1982), which were:

NART FSIQ = 127.7 - 0.826 x NART error score, (S.E.E. = 7.6).

NART VIQ = 129.0 - 0.919 x NART error score, (S.E.E. = 7.6).

Similarly, regression equations to predict Vocabulary IQs were generated by regressing Vocabulary age-scaled scores against WAIS-R FSIQ and WAIS-R VIQ in the control sample:

Vocabulary FSIQ = 62.67 + 3.30 x Vocabulary subtest age-scaled score, (S.E.E. = 6.49).

Vocabulary VIQ = 57.73 + 3.87 x Vocabulary subtest age-scaled score, (S.E.E. = 6.92).

These equations, although derived from a small sample (N=20), were similar to an equation for prediction of WAIS-R FSIQ using the Vocabulary age-scaled score, derived by Nelson and McKenna (1975) on a sample of 98 U.K. non-neurologically impaired subjects. This equation was used by Crawford and his colleagues (1988a) and Hart and her colleagues (1986) to predict Vocabulary estimated FSIQ and was:

Vocabulary FSIQ = 61.00 + 4.00 x Vocabulary age-scaled score, (S.E.E. = 5.6).

NART errors and Vocabulary age-scaled scores for each dementing subject were entered into the appropriate regression equation reported above to yield NART and Vocabulary predicted premorbid IQ scores. Two repeated measures Analyses of Variance (ANOVAs) were computed to determine whether or not there were significant differences between current and premorbid IQ scores for dementing individuals. The first ANOVA included WAIS-R FSIQ, NART estimated FSIQ and Vocabulary estimated FSIQ and revealed that there were significant differences between these measures ($F = 50.00$, $p<.0001$) (see Table 6). A series of post hoc Scheffe tests revealed that the mean NART estimated FSIQ ($M = 88.85$, S.D. = 10.70) was significantly greater than the mean Vocabulary estimated FSIQ ($M = 84.85$, S.D. = 8.38, $p<.05$) which was significantly greater than the mean WAIS-R FSIQ ($M = 74.45$, S.D. = 9.14, $p<.05$).

A second repeated measures ANOVA was similarly computed on WAIS-R VIQ, NART estimated VIQ and Vocabulary estimated VIQ. Again there was a significant difference between these measures for dementing individuals ($F= 20.34$, $p<.0001$, see Table 7). A series of post hoc Scheffe tests again revealed that the mean NART estimated VIQ ($M = 88.70$, S.D. = 12.94) was significantly greater than the mean Vocabulary estimated VIQ ($M = 83.65$, S.D. = 10.09, $p<.05$) which was significantly greater than the mean WAIS-R VIQ ($M = 77.65$, S.D. = 8.74, $p<.05$).

NART and Vocabulary predicted WAIS-R PIQ and demographic predicted WAIS-R FSIQ, VIQ, and PIQ were not determined for dementing subjects. This was due to indications in the literature that (a) NART errors, Vocabulary age-scaled scores, nor demographic variables are good predictors of WAIS-R PIQ, and (b) multiple variable predictors such as required to

Table 6
Repeated measures ANOVA comparing three measures of FSIQ: WAIS-R, NART, and Vocabulary

Source	SS	df	MS	F
Between Subjects	4258.13	19	224.11	
Within Subjects				
Between measures	2210.13	2	1105.07	50.00****
error(w)	839.87	38	22.10	

****$p<.0001$

Table 7
Repeated measures ANOVA comparing three measures of VIQ: WAIS-R, NART, and Vocabulary

Source	SS	df	MS	F
Between Subjects	5426.00	19	285.58	
Within Subjects				
Between measures	1224.03	2	612.02	20.34****
error(w)	1143.30	38	30.09	

****$p<.0001$

determine demographic predicted IQ should not be used in small samples due to the risk of obtaining misleading high or low correlations (Crawford, 1989). For the sake of comparison, and with these potential contra-indications in mind however, the amount of variance predicted in WAIS-R FSIQ, WAIS-R VIQ and WAIS-R PIQ by NART errors, Vocabulary age-scaled scores, and demographic variables (age, sex, education, occupation), was determined by regressing the estimated premorbid indicators against the current IQ measures. The results showed that demographic variables accounted for less of the variance in WAIS-R FSIQ (51%) than either NART errors (59%) or Vocabulary age-scaled scores (77%). Further, demographic variables predicted less of the variance in WAIS-R VIQ (54%) than either NART errors (65%) or Vocabulary age-scaled scores (80%). None of the premorbid estimators (NART errors, Vocabulary age-scaled scores, demographic variables) was a good predictor of WAIS-R PIQ (15%, 29%, and 37%, respectively). In addition, demographic variables did not add a significant amount of predicted variance to either of the regression equations predicting WAIS-R FSIQ: NART errors (only 4%) or Vocabulary age-scaled scores (<1%).

1.0.1. Another diagnostic measure used to distinguish D.I.'s from C.'s: Coolidge's Algorithm

A comparative analysis revealed that only 30% (6/20) of the dementing subjects were correctly classified using Coolidge's Algorithm (dementia is indicated if the Vocabulary age-scaled subtest score is equal or greater than twice the Block Design age-scaled subtest score, Coolidge et al., 1985). In contrast, 95% (19/20) of the control subjects were incorrectly classified using this formula. Fisher's Exact Test revealed that there was a marginally statistically significant difference between groups (dementing vs. control, $p<.05$).

Since Crawford and his colleagues (unpublished manuscript b) discovered that Coolidge's WAIS algorithm was probably more beneficial in distinguishing those suffering from Dementia of the Alzheimer's Type (DAT) from Normal and Depressed controls than in distinguishing Multi Infarct Dementing individuals (MID) from these same controls, these two groups of dementing subjects were compared on the accuracy of diagnosis using Coolidge's formula. The results indicated that 37.5% (6/16) of DATs as compared to 0% (0/4) of MIDs were correctly classified using Coolidge's algorithm. This difference between groups (DAT vs. MID) was not statistically

significant as measured by Fisher's Exact Test. The DAT group was compared alone with the normal control group to determine whether or not there was a greater statistically significant difference between control subjects and this more homogeneous dementing group than there was when all dementing individuals were categorized together. Fisher's Exact Test revealed that there was still only a marginally significant difference between groups (DAT vs. Controls, $p<.05$).

DISCUSSION

This study has examined the reliability and validity of the NART in the estimation of premorbid intelligence in Newfoundland. It is, of course, unlikely that any psychometric test will be completely unaffected by neuronal death in dementia, and therefore this thesis has been concerned with relative rather than absolute "dementia-insensitivity". With this in mind, the discussion will begin with the NART's ability to be reliably scored by two separate scorers, followed by its ability to measure the construct of intelligence. The NART will then be discussed in terms of the implications for its relative "dementia-insensitivity" (i.e., the ability for mild-moderate dementing subjects to perform similar to control subjects), and its comparability with two other commonly used measures for assessing premorbid intelligence (the Vocabulary subtest of the WAIS-R and demographic variables). Finally, the ability of two WAIS-R algorithms to distinguish the dementing from the nondementing subjects in the present study, Coolidge's algorithm (Coolidge et al., 1985) and the V-P Split (Wechsler, 1955; 1981), will be discussed.

2.1. Inter-rater reliability

In the present study, an inexperienced NART user's scoring of NART errors was compared with an experienced NART user's scoring. The result was almost perfect correlation ($r = 0.96$) between the two sets of scores, confirming Crawford and his colleagues' (1989a) finding that the NART can be reliably scored by both those experienced and those inexperienced with the test. The results further concur with a study which found a high inter-rater reliability between ten experienced NART users scoring of the NART, (O'Carroll, 1987), and another study finding high inter-rater reliability of a revised Canadian version of the NART (Blair and Spreen, 1989). This finding supports the NART's versatility as a clinical test, in that it can be readily mastered in a short period of time.

2.2. Construct validity: Measurement of intelligence

The present study attempted to discern the degree to which the NART correlated with commonly used measures of current intelligence (i.e., WAIS-R FSIQ, VIQ, and PIQ). The results supported the first experimental hypothesis in that the NART was significantly correlated with both WAIS-R FSIQ and WAIS-R VIQ, but correlated less well (though still significantly) with WAIS-R PIQ in the control sample. This finding replicated that of other studies which have found the NART to correlate well with WAIS/WAIS-R FSIQ in neurologically intact subjects (Crawford et al., 1990 a; c). The correlation between NART and WAIS-R FSIQ in this study (r = -0.76) was similar to correlations between the NART and WAIS-R FSIQ reported by Crawford and his colleagues in two studies (r = -0.72 and r = -0.79, Crawford et al., 1990 a; c, respectively). Further, NART errors predicted a substantial amount of variance in WAIS-R FSIQ and VIQ, but predicted less of the variance in WAIS-R PIQ, using the data from the control subjects. The results indicated that the amount of variance predicted in WAIS-R FSIQ in this study (59%) fell in between the amount of variance in WAIS FSIQ reported by Nelson (1982) in her standardization of the NART (55%) and Crawford and his colleagues (1989a) in their cross-validation of the NART (66%). Blair and Spreen (1989) found similar results with their revised North American NART predicting 56% of the variance in WAIS-R FSIQ. Similarly, the present results for the amount of variance NART predicted in WAIS-R VIQ mediated between Nelson's (1982) and Crawford and his colleagues' (1989a) results with the WAIS (65% versus 60% and 72%, respectively). There were no observable differences between the amount of variance predicted in PIQ between these results and the above studies, all finding the NART to predict a lower amount of variance in WAIS/WAIS-R PIQ (approximately 30%) than WAIS/WAIS-R FSIQ and VIQ.

In summary, the results from this study concur with the results of other researchers in the field, both British and Canadian, despite the small sample size. That is, the NART correlated well with and predicted a substantial amount of variance in WAIS-R FSIQ and VIQ, but correlated less well with and predicted less of the variance in PIQ. The results therefore suggest that the NART is a valid measure of current intelligence (at least WAIS/WAIS-R FSIQ and VIQ).

2.3. Construct validity: "Dementia-insensitivity"

This study attempted to evaluate the ability of the NART to "hold" in dementia (i.e., its ability to remain relatively unaffected by the dementing process). This was done in two ways, first by establishing the NART's correlation with the dementia score (i.e., was there an association between NART performance and dementia severity?), and second, by estimating the NART's ability to distinguish between dementing and control subjects.

The results from this study supported the second hypothesis. Although the NART significantly correlated with the dementia score in dementing subjects when a simple correlation between the two measures was calculated, this correlation became nonsignificant when the demographic variables were controlled for. In comparison, the WAIS-R FSIQ significantly correlated with dementia score regardless of whether or not the demographic variables were controlled for. In other words, the more severe the degree of dementia, the lower the WAIS-R FSIQ, while NART performance was relatively unaffected regardless of severity of dementia. This result was in keeping with two previous studies investigating the correlation between NART and dementia severity (Crawford et al., 1988, cited in Crawford, 1989, O'Carroll & Gilleard, 1986). The results did not support the finding of Wood and his colleagues (1984), who found the NART to be significantly correlated with dementia severity. Since the latter researchers cautioned their results due to the possibility that some individuals with low standing intelligence may have been included in the dementing group, the collective evidence supports the view that NART performance is at least relatively unaffected by dementia severity in mild-moderate dementing subjects.

Hypothesis three was supported by these findings. That is, although there was a significant difference between dementing and control subjects' NART performance using a simple t-test, this became nonsignificant when demographic variables were partialled out using multiple regression analysis. In contrast, dementing subjects scored lower than control subjects on all measures of current IQ (i.e., WAIS-R FSIQ, VIQ, and PIQ), regardless of whether demographic variables were partialled out or not. In other words, NART performance was shown to be relatively "dementia-resistant" in that dementing individuals did not perform significantly different from control subjects on this test, after controlling for demographic differences between groups.

This finding replicates that of four groups of researchers comparing NART performance and WAIS performance in dementing and nondementing individuals (Crawford et al., 1988a, Nebes et al., 1984, Nelson & O'Connell, 1978, Schlosser & Ivison, 1989). None of these studies found significant differences between dementing and cognitively intact control groups on NART performance, but found significant differences between the groups on WAIS IQ (Nelson & O'Connell, 1978) and memory tests (Nebes et al., 1984, Schlosser & Ivison, 1989). However, the results conflicted with three studies: Brayne & Beardsall (1990) and Hart and her colleagues (1986), found their dementing samples scored lower on the NART than their normal elderly control samples, and Stebbins and his colleagues (1987) found that their moderately and severely dementing subjects scored lower on the NART than their mildly dementing subjects.

One possible explanation for the present findings of a nonsignificant difference in NART performance between dementing and control subjects when demographic variables were controlled for, is that the dementing subjects were mildly and moderately impaired (i.e., scored an M of 5.02 with an SD of 2.19 on the Abbreviated Dementia Scale, Qureshi & Hodkinson, 1974, which is close to the midpoint). That is, it is probable that neither the NART, nor any cognitive test for that matter, "holds" in the more severe cases of dementia. However, the clinical utility of the instrument would be in the milder cases of dementia where differential diagnosis is a clinical problem. It is rarely necessary to employ psychometric instruments to aid in diagnosis when the cognitive and behavioral deficits associated with severe dementia are readily observable. Therefore, since the NART "held" relatively well in the present and the above studies in mild-moderate dementing individuals, (i.e., was not correlated with dementia severity, and there was no significant difference in NART performance between mild-moderate dementing individuals and non-cognitively impaired subjects), the evidence suggests the NART is a valuable clinical tool to aid in the early diagnosis of dementia, both in Canada and the U.K..

2.4. Comparability with existing premorbid intelligence measures

2.4.1. Vocabulary subtest of the WAIS-R

The fourth experimental hypothesis was also borne out in this study. That is, when dementing individuals' NART estimated premorbid intelligence was compared with Vocabulary estimated premorbid intelligence, the NART estimated a significantly higher WAIS-R FSIQ and WAIS-R VIQ. Further, both the NART and the Vocabulary subtest WAIS-R FSIQ and VIQ estimates were significantly larger than the current IQ measures (WAIS-R FSIQ and VIQ). It should also be noted that comparison of unconverted NART and Vocabulary subtest scores between dementing and control subjects also indicated a superiority of the NART. That is, there was no significant difference between dementing and control subjects in NART performance when demographic variables were controlled for, while dementing subjects performed significantly more poorly than control subjects on the Vocabulary subtest of the WAIS-R whether or not demographic variables were controlled for. The NART therefore appears to be a more "dementia-resistant and therefore a superior estimate of premorbid IQ than the Vocabulary subtest of the WAIS-R. In addition, since both the Vocabulary subtest and the NART estimated higher WAIS-R FSIQ and VIQ scores than the observed current performance on these measures, both the NART and Vocabulary subtest are likely to be better estimators of premorbid functioning than simply the observed performance on WAIS-R current measures of intelligence.

The results of this study concur with the results of other researchers who have compared the NART with the Vocabulary subtest of the WAIS in the estimation of premorbid WAIS IQ (Crawford et al., 1987; 1988a; Hart et al., 1986). That is, the above researchers have found the NART to estimate a significantly higher WAIS FSIQ than the Vocabulary subtest of the WAIS. The recent evidence therefore supports the NART as a superior predictor of premorbid intelligence than the Vocabulary subtest of the WAIS/WAIS-R.

It was impossible to derive predicted NART and Vocabulary subtest IQ scores for the control subjects since the regression equations were developed using data from these control subjects. This problem was unavoidable since British WAIS norms would have been inappropriate in this study (the WAIS-R was used in the present investigation, and it would have been unwise to assume that the relationship between Vocabulary and/or word reading ability and the WAIS/WAIS-R was the same in Canada as in Britain). Further, North American norms were

unavailable, as to the writer's knowledge this was the first study of its kind to be carried out in Canada. However, one possible criticism of this study is that the NART estimated WAIS-R FSIQ and VIQ, which were significantly higher than Vocabulary estimated WAIS-R FSIQ and VIQ, might simply represent an overestimation of these measures. That is, since there was no control group for comparison, it is difficult to know if the premorbid estimates of dementing individuals' IQs exceeded or fell short of premorbid IQ levels obtained by control subjects. However, if one visually compares the mean NART estimated WAIS-R FSIQ for dementing subjects (M = 88.85, SD = 10.70) with the mean observed WAIS-R FSIQ for control subjects (M = 96.95, SD = 13.15), it appears that if anything, the NART predicted WAIS-R FSIQ may have been an underestimate. Similarly, comparing NART estimated WAIS-R VIQ for dementing subjects (M = 88.70, SD = 12.94) with observed WAIS-R VIQ for control subjects (M = 97.95, SD = 15.12), the NART estimate again appears to be, if anything, lower. Therefore, it is unlikely that the NART overestimated the premorbid IQ of dementing subjects in this study, and since the NART estimated a significantly higher WAIS-R FSIQ and VIQ than the Vocabulary subtest, it can be concluded that in the present study, NART was the procedure of choice for estimating the premorbid IQ of dementing subjects.

Premorbid WAIS-R PIQs were not calculated in this study since Crawford (1989) and Blair and Spreen (1989) have reported that neither the NART nor the Vocabulary subtest of the WAIS/WAIS-R is a good predictor of PIQ, which was confirmed in the present study.

2.4.2. Demographic variables

The fifth hypothesis was not borne out by these data. That is, demographic variables did not significantly increase the amount of predicted variance in WAIS-R FSIQ, when combined in a regression analysis with either the NART (added only 4%) or the Vocabulary subtest of the WAIS-R (added <1%). This finding was in contrast to that of Crawford and his colleagues (1989d), who found that demographic variables plus NART significantly increased the amount of predictive accuracy in WAIS-R FSIQ in comparison with NART alone (i.e., demographic variables added 7% of the predicted variance in WAIS FSIQ). It should be noted, however, that the present results concurred with the results of Blair and Spreen (1989), who found demographic variables added only 3% when combined with the Canadian NART-R to predict WAIS-R FSIQ. However, the

results of the prediction of premorbid IQ by the NART, the Vocabulary subtest of the WAIS-R, and demographic variables, must be interpreted with caution due to the small sample size of the present study (see 4.6 Criticisms of the present study below).

It is difficult to make comparisons between the studies investigating demographic variables in combination with psychometric techniques to predict premorbid IQ, since Blair and Spreen (1989) used a revision of the NART and different demographic variables from the present study and that of Crawford and his colleagues (1989d). Further, the present study as well as the Blair and Spreen (1989) study used the WAIS-R as the comparative measure, while Crawford and his colleagues (1989d) used the WAIS. However, the differences are interesting and suggest the possibility that the relationship between demographic variables and IQ may be different in continents separated by the Atlantic.

2.5. WAIS-R algorithms to distinguish dementing from nondementing subjects: Coolidge's algorithm and the V-P Split

The sixth hypothesis was only partially borne out by these data. That is, Coolidge and his colleagues' (1985) algorithm (Vocab $\geq$ 2 Bl.D = dementia) was able to correctly classify a statistically but not a clinically significant number of dementing and control subjects. In addition, as expected from previous research, dementing subjects did not score significantly higher on the V-P Split (VIQ > PIQ) than control subjects.

The results of this study indicated that 70% (14/20) of dementing subjects would have been misclassified as nondementing using this algorithm. Further, the algorithm was only able to correctly classify some of the DAT patients (37.5% or 6/16), and none of the MID patients (0/4). In comparison, the formula misclassified only one (5%) of the twenty control subjects (i.e., only one control subject's profile would have indicated dementia).

These findings are in contrast to those of Coolidge and his colleagues (1985), who found the algorithm to correctly classify 74% of their dementing subjects, and Crawford and his colleagues (unpublished manuscript b), who found the algorithm to correctly classify 68% of their DAT subjects and 45% of their MID subjects. However, the present results concur with some aspects of Crawford and his colleagues' (unpublished manuscript b) findings in that first, few

normal elderly controls were incorrectly classified as dementing, and second, the algorithm was less accurate in identifying MID patients than DAT patients.

One possible explanation for the lower rate of classification for dementing subjects in the present study might be that the sample size was rather small. However, if the formula is to be clinically useful, it must work with individual cases presenting with a query of dementia. That is, the formula would be clinically useless if it classified a statistically significant proportion of a large sample, but was unable to predict dementia in the individual case. Another possible explanation for differences between this and Coolidge and his colleagues (1985) and Crawford and his colleagues (unpublished manuscript b) studies, is that the present study used the WAIS-R while other studies used the WAIS. Since the WAIS-R is the updated version of the test, it is important to determine the utility of Coolidge's algorithm in the more widely used WAIS-R.

Thus, the results suggest that Coolidge's algorithm may be subject to the same criticisms aimed at previous research into WAIS algorithms (Vogt & Heaton, 1977), in that it might only have diagnostic utility when the patient's profile is positive for dementia (i.e., V $\geq$ 2 Bl.D), if at all. That is, dismissal of a dementing process cannot be made if the profile is not positive for dementia, and a positive profile should only alert the clinician to further investigation rather than suggesting definite diagnosis.

As mentioned above, there was no significant difference between dementing and control subjects on the V-P Split. Further, both groups (dementing and control) received positive V-P Splits ($\underline{M}$ = 5.15, $\underline{SD}$ = 9.03 for dementing subjects and $\underline{M}$ = 2.35, $\underline{SD}$ = 14.17 for control subjects), indicating that in the present cognitively intact elderly sample, scores were in the direction of VIQ > PIQ rather than PIQ > VIQ. This result suggests that previous research aimed at determining V-P Split norms in nondementing subjects, which disregarded the direction of verbal performance discrepancies (Field, 1960; Grossman, 1983; Naglieri, 1982; Wechsler, 1981), may have underestimated the extent of the V-P Split (VIQ > PIQ) in the normal elderly. Further, the results of this study concurred with Hart and her colleagues (1986) who found no significant difference between their dementing and control subjects on the V-P Split. Thus, although this Wechsler algorithm is frequently used clinically to aid in diagnosis of dementing disorders, the present study and previous research (Hart et al., 1986) suggest the V-P Split is

unimpressive as an indicator of dementia, and it is worrisome how commonly the V-P Split is used clinically, given the dearth of evidence to support its diagnostic validity.

2.6. Criticisms of the present study

The present study should be considered as a pilot study since the sample size was rather small (N = 40 or N = 20 per group, dementing versus control). This point is particularly important when considering the derived predicted premorbid intelligence of dementing individuals. The regression equations for determining premorbid IQ scores were developed by regressing variables (NART errors, Vocabulary subtest scores, demographic variables) independently or in combination against WAIS-R current IQ measures in the control sample (N = 20). Crawford (1989) cautioned that misleading high or low correlations may be obtained when using small samples, especially with reference to multiple demographic variables. Therefore, the equations generated in this study to predict premorbid intelligence must not be used clinically, even though they appear comparable to regression equations generated in previous research. In order to generate regression equations that can be used for differential diagnosis in clinical practice, or to state any firm conclusions about the NART's ability to predict higher premorbid IQs than the Vocabulary subtest, or the demographic variables' inability to significantly increase the amount of predicted IQ variance when combined with NART or the Vocabulary subtest, a large scale standardization study must be completed in Canada.

A further criticism is that the present study used a prorated WAIS-R (the seven subtests employed in the NART standardization study, Nelson, 1982). Crawford and his colleagues (1989a) have cautioned against this, since they found the amount of predicted WAIS variance accounted for by NART errors, increased when the full WAIS was used in comparison with a prorated WAIS (Nelson, 1982). Therefore the results of this study may be an underestimation of the NART's ability to estimate premorbid WAIS-R intelligence.

Another criticism of the present study involves the nature of the sample. The original intention was to collect only dementing and medical daypatients and outpatients, in attempts to include only mildly and moderately dementing patients and their best control counterparts. However, difficulties in data collection prevented the collection of a 'clean' sample. That is, one

dementing inpatient and eight control nonpatients were included in the study. The dementing inpatient was severely demented as measured by the Abbreviated Dementia Scale (Qureshi & Hodkinson, 1974) and thus may have been impaired in her ability to perform at premorbid levels on the NART, as is suggested by the research including severely dementing individuals in their samples (Hart et al., 1986; Stebbins et al., 1987). If this was the case, inclusion of her data may have resulted in an underestimate of the NART's ability to predict premorbid intelligence in this study.

The inclusion of 8 nonpatient control subjects in the present sample, may have resulted in a control sample consisting of better adjusted subjects than the dementing sample. An individual in good physical/mental health might be expected to perform better on cognitive tasks than an individual in poor physical/mental health. Further, the 8 nonpatients were identified from a list of elderly individuals who expressed an interest in participating in research, many of whom had participated in prior research studies. Therefore they may have been more task motivated and less anxious in the testing situation. Had only medical day- and outpatients been used in the control sample, there may have been smaller discrepancies between the groups (dementing vs. control) on the WAIS-R.

Further with regard to the nature of the sample, both dementing and control subjects seem to have had better educations and to have held more prestigious occupations than the general Newfoundland elderly population. Had a more representative sample been used, it is possible that there would have been differential NART performance between the groups (dementing vs. control). That is, it is possible that dementing Newfoundlanders with limited education might lose their ability to read more readily than well-educated dementing Newfoundlanders. Therefore, our results can only be generalized to better educated Newfoundlanders who were employed in more prestigious occupations than the average for their age cohort.

Finally, due to the lack of experienced NART users in the Memorial University of Newfoundland Department of Psychology, both the author and her supervisor were the sole scorers of the NART. Since both scorers were well aware of the experimental hypotheses, this may have biased the scores in the predicted direction. This problem was addressed by keeping the experienced NART user blind to group (dementing vs. control), however, it is possible that he could guess to which group some subjects belonged, thus making him not completely blind.

2.7. Conclusions and recommendations

The results of this study have supported the NART's inter-rater reliability and its validity as a relatively "dementia-insensitive" measure of intelligence (estimated WAIS-R FSIQ and VIQ). That is, there were no significant differences between dementing and control subjects' performance on the NART and the NART was not correlated with dementia severity, when demographic variables were statistically controlled for. These results are in concordance with much of the results reported in the literature. Most of the research on this instrument has been carried out in the U.K. using the outdated WAIS IQ as the comparative current intelligence measure. Thus, it appears as though the NART may be a valid estimator of premorbid WAIS-R intelligence in Canada (or more specifically, Newfoundland). Further, the results of this study indicate that the NART may be the procedure of choice in estimating premorbid intelligence (as compared with the Vocabulary subtest of the WAIS-R and demographic variables). However, this research should be viewed as a pilot study as the sample size was rather small and a prorated WAIS-R was used.

Future research should focus on using the full WAIS-R index of intelligence with a larger sample size. Canadian NART norms should be established and equations specific to Canadians should be developed for determination of NART premorbid IQ. Blair and Spreen (1989) have developed a North American revised NART, which may be the 'NART of choice' in North America, since North American and U.K. word pronunciation may differ. Future research should also examine the possibility of comparing performance on the NART to the Wechsler Memory Scale (WMS), since Schlosser and Ivison (1989) found the NART to correlate well ($r = -0.67$) with this test, and since memory decline is often one of the earliest indications of dementia. Nevertheless, this study indicates that pronunciation of short irregular words may currently be the best clinical tool avail ble for determining the premorbid intelligence of Canadians.

References

American Psychiatric Association (1987). The Diagnostic and Statistical Manual of Mental Disorders, Third Edition (Revised). Washington: American Psychological Association.

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(4), 169-173.

Barona, A., Reynolds, C.R., & Chastain, R.L. (1984) A demographically based index of premorbid intelligence for the WAIS-R. Journal of Consulting and Clinical Psychology, 52(5), 885-887.

Binks, M.G., & Davies, A.D.M. (1984). The early detection of dementia. A baseline from healthy community dwelling old people. In D.B. Bromley (Ed.) Gerontology: Social and Behavioral Perspectives. Beckenham, Kent:Croon Helm, 7-13.

Blair, J.R. & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. The Clinical Neuropsychologist, 3(2), 129-136.

Blessed, G., Tomlinson, B.E. & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. British Journal of Psychiatry, 114, 797-811.

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

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

Coolidge, F.E., Peters, B.M., Brown, R.E., & Harsch,T.L. (1985). Validation of a WAIS algorithm for the early onset of dementia. Psychological Reports, 57, 1299-1302.

Coolidge, F.L., Rakoff, R.J., Schwellenbach, L.D., Bracken, D.D., & Walker, S.H. (1986). WAIS profiles in mentally retarded adults. Journal of Mental Deficiency Research, 30, 15-17.

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

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

Crawford, J.R., Allan, K.M., Cochrane, R.H.B., & Parker, D.M. (unpublished manuscript a). Assessing the validity of NART estimated premorbid IQs in the individual case.

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

Crawford, J.R., Cochrane, R.H.B., Besson, J.A.O., Parker, D.M., & Stewart, L.E. (1990b). Premorbid IQ estimates obtained by combining the NART and demographic variables: construct validity. Personality and Individual Differences, 11(2), 209-210.

Crawford, J.R., Hart, S., & Nelson, H.E. (1990). Improved detection of cognitive impairment with the NART: An investigation employing hierarchical discriminant function analysis. British Journal of Clinical Psychology, 29, 239-241.

Crawford, J.R., Morrison, F.M., Jack, A.M., Cochrane, R.H.B., Allan, K.M., & Besson, J.A.O. (1990c). WAIS vs WAIS-R in matched U.K. samples. Personality and Individual Differences, 11(4), 427-428.

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

Crawford, J.R., Parker, D.M., Besson, J.A.O., & Beavan, D.J. (unpublished manuscript b). Differentiation of Dementia Alzheimer Type using a WAIS derived formula.

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

Crawford, J.R., Stewart, L.E., Cochrane, R.H.B, Foulds, J.A., Besson, J.A.O., & Parker, D.M. (1989b). Estimating premorbid IQ from demographic variables: Regression equations derived from a UK sample. British Journal of Clinical Psychology, 28, 275-278.

Crawford, J.R., Stewart, L.E., Cochrane, R.H.B., Parker, D.M., & Besson, J.A.O. (1989c). Construct validity of the National Adult Reading Test: A factor analytic study. Personality and Individual Differences, 10(5), 585-587.

Crawford, J.R., Stewart, L.E., Garthwaite, P.H., Parker, D.M., & Besson, J.A.O. (1988b). The relationship between demographic variables and NART performance in normal subjects. British Journal of Clinical Psychology, 27, 181-182.

Crawford, J.R., Stewart, L.E., Parker, D.M., Besson, J.A.O., & Cochrane, R.H.B. (1989d). Estimation of premorbid intelligence: Combining psychometric and demographic approaches improves predictive accuracy. Personality and Individual Differences, 10(7), 793-796.

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

Eppinger, M.G., Craig, P.L., Adams, R.L., & Parsons, O.A (1987). The WAIS-R index for estimating premorbid intelligence: Cross-validation and clinical utility. Journal of Consulting and Clinical Psychology, 55(1), 86-90.

Field, J.G. (1960). Two types of tables for use with Wechsler's intelligence scales. Journal of Clinical Psychology, 16, 3-7.

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

Grossman, F.M. (1983). Percentage of WAIS-R standardization sample obtaining Verbal-Performance discrepancies. Journal of Consulting and Clinical Psychology, 51(4), 641-642.

Guertin, W.H., Ladd, C.E., Frank, G.H., Rabin, A.I., & Hiester, D.S. (1966). Research with the Wechsler intelligence scales for adults: 1960-1965. Psychological Bulletin, 66(5), 385-409.

Hachinski, V.C., Iliff, L.D., Zilhka, E., Du Boulay, G.H., McAllister, V.L., Marshall, J., Russell, R.W.R., & Symon, L. (1975). Cerebral blood flow in dementia. Archives of Neurology, 32, 632-637.

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

Karzmark, P., Heaton, R.K., Grant, I., & Matthews, C.G. (1985). Use of demographic variables to predict full scale IQ: A replication and extension. Journal of Clinical and Experimental Neuropsychology, 7(4), 412-420.

Kerlinger, F.N. & Pedhazur, E.J. (1973). Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston, Inc.

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

Klesges, R.C., Sanchez, V.C., & Stanton, A.L. (1981). Cross validation of an adult premorbid functioning index. Clinical Neuropsychology, 3(1), 13-15.

Law, J.G., Price, D.R., & Herbert, D.A. (1981). Study of Quick Test, WAIS, and premorbid estimates of intelligence for neuropsychiatric patients. Perceptual and Motor Skills, 53(3), 919-922.

Lezak, M.D. (1983). Neurological Assessment (Second Edition). New York: Oxford University Press, Inc.

Matarazzo, J.D., & Herman, D.O. (1985). Clinical uses of the WAIS-R: Base rates of differences between VIQ and PIQ in the WAIS-R standardization sample. Handbook of Intelligence: Theories, Measurement and Applications, 899-932.

Miller, E. (1977). Abnormal Ageing: The Psychology of Senile and Presenile Dementia. London: John Wiley & Sons.

Naglieri, J.A. (1982). Two types of tables for use with the WAIS-R. Journal of Consulting and Clinical Psychology, 50(2), 319-321.

Nebes, R.D., Martin, D.C., & Horn, L.C. (1984). Sparing of semantic memory in Alzheimer's Disease. Journal of Abnormal Psychology, 93(3), 321-330.

Nelson, H.E. (1982). National Adult Reading Test for the Assessment of Premorbid Intelligence in Patients with Dementia: Test Manual, Windsor: NFER Publishing Co.

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 premorbid intelligence levels using the New Adult Reading Test. Cortex, 14, 234-244.

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

O'Carroll, R.E., 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.

Offices of Population, Censuses and Surveys (1980). Classification of Occupations. London: HMSO.

Qureshi, K.N. & Hodkinson, H.M. (1974). Evaluation of a ten question mental test in the institutionalised elderly. Age and Ageing, 3, 152-157.

Rosen, W.G., Terry, R.D., Fuld, P.A., Katzman, R., & Peck, A. (1980). Pathological verification of ischemic score in differentiation of dementias. Annals of Neurology, 7, 486-488.

Roszkowski, M. & Spreat, S. (1982). Scatter as an index of organicity: A comparison of mentally retarded individuals experiencing and not experiencing concomitant convulsive disorders. Journal of Behavioral Assessment, 4(4), 311-316.

Roszkowski, M. & Spreat, S. (1983). Assessment of effective intelligence: Does scatter matter? The Journal of Special Education, 17(4), 453-460.

Ruddle, H.V. & Bradshaw, C.M. (1982). On the estimation of premorbid intellectual functioning: Validation of Nelson & McKenna's formula, and some new normative data. British Journal of Clinical Psychology, 21, 159-164.

Russell, E.W. (1972). WAIS factor analysis with brain damaged subjects using criterion measures. Journal of Consulting and Clinical Psychology, 39(1), 133-139.

Savage, R.D., Britton, P.G., Bolton, N., & Hall, E.H. (1973). Intellectual Functioning in the Aged. London: Methuen & Co. Ltd.

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(6), 785-792.

Searle, R.T. (1984). A community based follow-up of some suspected cases of early dementia: An interim report. In D.B. Bromley (Ed.) Gerontology: Social and Behavioral Perspectives. Beckenham, Kent: Croon Helm. 20-28.

Silverstein, A.B. (1987). Accuracy of estimates of premorbid intelligence based on demographic variables. Journal of Clinical Psychology, 43(5), 493-495.

Stebbins, G.T., Wilson, R.S., Gilley, D.W., Bernard, B.A., & Fox, J.H. (1987). Estimation of premorbid intelligence. Journal of Clinical and Experimental Neuropsychology, 10, 63-64.

Thompson, P. & Blessed, G. (1987). Correlation between the 37-item mental test score and abbreviated 10-item mental test score by psychogeriatric day patients. British Journal of Psychiatry, 151, 206-209.

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

Wechsler, D. (1944) The Measurement of Adult Intelligence (3rd edition). Baltimore: Williams and Wilkins.

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

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

Whalley, L.J. (1989). Drug treatments of dementia. British Journal of Psychiatry, 155, 595-611.

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

Wilson, R.S., Rosenbaum, G., Brown, G., Rourke, D., Whitman, D., & Grisell, J. (1978). An index of premorbid intelligence. Journal of Consulting and Clinical Psychology, 46(6), 1554-1555.

Wood, N., Copeland, J.R.M., Forshaw, D.M., Muthu, M.S, Abed, R., Sharma, V.K., & Dewey, M.E. (1984). The early detection of dementia: Initial findings from a longitudinal study. In D.B. Bromley (Ed.) Gerontology: Social and Behavioral Perspectives. Beckenham, Kent: Croon Helm. 14-19.

Zarit, S.H. & Zarit, J.M. (1983). Cognitive impairment. Chapter in P.M. Lewinsohn and L. Teri (Eds.) Clinical Geropsychology: New Directions in Assessment and Treatment. New York: Pergamon Press.

Appendix-A

Information Sheet and Consent Forms for Dementing and Control Subjects

Research Project on Cognitive and Reading Ability and Demographic Background

We are currently interested in investigating reading ability and cognitive functioning in people with memory and orientation problems. Psychologists have suggested that reading ability remains relatively unaffected in individuals with impaired cognitive functioning. If this is true, then measures of reading ability might provide an estimate of previous intellectual functioning of patients with current difficulties. This type of measure is extremely useful to psychologists and doctors who wish to determine how severe the impairment is and with what speed the process is occurring. Early detection of impairment might be beneficial in terms of treatment.

We are carrying out a project where we plan to look at the performance of individuals on tests of reading ability and cognitive functioning. This involves the person being assessed for a total of approximately 1-2 hours. The individual will be given as many breaks as required to make the testing period as comfortable as possible. The first test involves memory questions and asking the person some questions on current information and some concentration tasks such as counting backwards. The second test involves reading aloud a list of words into a tape recorder. The third test involves a number of different tasks, some requiring a spoken response, and some requiring completion of puzzles and other nonverbal tasks. The items start off fairly easy and increase in difficulty. In addition to the testing period, patient records will be consulted for demographic information and details concerning the reasons for hospital care. The results of this investigation will remain confidential, being seen only by the primary investigator (Karen Sharpe, M.Sc. candidate) and her supervisor (Dr. Ronan O'Carroll), of Memorial University Psychology Department. All the assessments will be carried out by Karen Sharpe, who will be happy to answer any questions you may have about the project (telephone number 737-8496).

Consent Form For Dementing Subjects

Consent Form For The Research Project on Cognitive and Reading Ability

I understand that this research involves the completion of several tests of cognitive ability, reading, and memory, and that personal information will be used. I realize that patient records will be utilized to gain some of the required information. The results will be treated confidentially. I understand that the complete assessment will take approximately 2 hours and that short breaks will be given wherever necessary. I realize that the patient is free to withdraw from the project at any point. Finally I understand that the results may not be of direct benefit to the patient but that they may be of some value in the assessment and treatment of patients in the future.

I, _____, the undersigned, agree to

(my relative's or ward's, _____)

participation in the research study described above.

_____ _____

(Signature of Participant) (date)

_____ _____

(Signature of Witness) (date)

To be signed by investigator:

To the best of my ability I have fully explained to the subject the nature of this research study. I have invited questions and provided answers. I believe that the subject fully understands the implications and voluntary nature of the study.

_____ _____

(Signature of Investigator) (date)

Consent Form For Control Subjects

Consent Form For The Research Project on Cognitive and Reading Ability

I understand that this research involves the completion of several tests of cognitive ability, reading, and memory, and that some personal information will be used. I realize that my hospital records will be used to gain some of the required information. The results will be treated confidentially. I understand that the complete assessment will take approximately 2 hours and that short breaks will be given wherever necessary. I realize that I am free to withdraw from the project at any point. Finally I understand that the results may not be of direct benefit to me but that they may be of some value in the assessment and treatment of patients in the future.

I, _____, the undersigned, agree to participate in the research study described above.

_____ _____
(Signature of Participant) (date)

_____ _____
(Signature of Witness) (date)

To be signed by investigator:

To the best of my ability I have fully explained to the subject the nature of this research study. I have invited questions and provided answers. I believe that the subject fully understands the implications and voluntary nature of the study.

_____ _____
(Signature of Investigator) (date)

Appendix-B

Practice Reading Test

BUN

SAP

CARD

DARK

RING

Appendix-C

National Adult Reading Test, Nelson, 1982

CHORD

ACHE

DEPOT

AISLE

BOUQUET

PSALM

CAPON

DENY

NAUSEA

DEBT

COURTEOUS

RAREFY

EQUIVOCAL

NAIVE

CATACOMB

GAOLED

THYME

HEIR

RADIX

ASSIGNATE

HIATUS

SUBTLE

PROCREATE

GIST

GOUGE

SUPERFLUOUS

SIMILE

BANAL

QUADRUPED

CELLIST

FACADE

ZEALOT

DRACHM

AEON

PLACEBO

ABSTEMIOUS

DETENTE

IDYLL

PUERPERAL

AVER

GAUCHE

TOPIARY

LEVIATHAN

BEATIFY

PRELATE

SIDEREAL

DEMESNE

SYNCOPE

LABILE

Appendix-D

Demographic Questionnaire

Date of Birth: day____month____year____

Sex: male____female____

Race: white____black____other____

Mothertongue: English____French____Other____

Number of Years Full-time Education:_____

Age on Leaving School:_____

Number of Years of Night School:_____

Occupation/Previous Occupation/
Husband's Occupation (if never
gainfully employed):_____

Community Type: rural____urban____

Current Medication?_____

Average Weekly Intake of Alcohol:_____

Have you ever had treatment for a
head injury? yes____no____

If yes, please give brief details_____
_____
_____

Have you ever had a stroke
before? yes____no____

Have you ever seen a psychologist or a
psychiatrist before? yes____no____

If yes, please give a brief description
(STRICTLY CONFIDENTIAL)_____
_____

Appendix-E

Abbreviated Dementia Scale, Qureshi & Hodkinson, 1984

Information-Memory-Concentration Test

Information Test

Age ----------1

Time (hour) ----------1

Year ----------1

Place-Name ----------1

Recognition of persons (cleaner, doctor,
nurse, patient, relative; any 2 available)----------1

Memory

(1) personal

Date of Birth ----------1

(2) non-personal

*Date of World War 1 ----------1

Monarch ----------1

(3) Address (5-minute recall)

42 West Street ----------1

Concentration

Counting 20-1 ----------2 1 0

*1/2 for approximation within 3 years

Appendix-F

Revised Hachinski Index

Rosen et al, 1980

HACHINSKI INDEX

Ischemic Score

Abrupt Onset	2
Stepwise Deterioration	1
Somatic Complaints	1
Emotional Incontinence	1
History of Strokes	2
History or Presence of Hypertension	1
Focal Neurological Symptoms	2
Focal Neurological Signs	2