Development of the New Zealand Adult Reading Test (NZART): Preliminary Findings

Nicola J. Starkey, *University of Waikato*
TeeJay Halliday, *University of Waikato*

The National Adult Reading Test (NART) is used to estimate premorbid intelligence. To establish New Zealand norms for the original NART and develop a New Zealand NART, sixty-three participants were administered Wechsler's Abbreviated Scale of Intelligence (WASI), the NART and the New Zealand Adult Reading Test (NZART). Regression equations were developed to estimate full scale, verbal and performance IQ from the NART and the NZART. Premorbid IQ estimations did not differ between Māori and NZ European participants. The NZART correlated highly with current IQ, was most accurate within the average range and 70% of participants were placed into the correct IQ category. This preliminary study indicates that with further validation the NZART may useful for estimating premorbid IQ in NZ.

Accurate diagnosis of cognitive impairment following the onset of illness (e.g., dementia) or injury (e.g., traumatic brain injury) relies upon knowledge of the individual's premorbid level of cognitive functioning (Lezak, Howieson, & Loring, 2004). However, as cognitive testing is uncommon in healthy individuals, this information is rarely available (Hebben & Milberg, 2002). One alternative is to compare an individual's performance to that of a normative sample, but the subtle nature of many cognitive deficits and different patterns of cognitive strengths and weaknesses within each individual mean that such comparisons are not ideal. In particular, comparisons with normative data may under-estimate cognitive decline in previously high functioning individuals and over-estimate deficits in those with lower levels of premorbid intellectual functioning (Green et al., 2008; Lezak et al., 2004). Taking these issues into account, Lezak et al. (2004) suggest that "only an individual comparison provides a meaningful basis for assessing deficit".

To overcome this, estimates of premorbid intellectual functioning are frequently employed. To be accurate, these estimates must be reliable, correlate highly with current functioning in healthy individuals and be resistant to cognitive decline (Crawford, 1989; Lezak et al., 2004; Sharpe & O'Carroll, 1991). Such estimates can be obtained by entering historical and observational data (e.g., previous occupation, income, education) into a regression equation (e.g., Barona's formula; Barona, Reynolds, & Chastain, 1984), or on some aspect of their present ability which is resistant to cognitive deterioration (see Reynolds, 1997 for a review of estimation of premorbid IQ). As a New Zealand derived regression equation to predict premorbid IQ is not yet available, local researchers and clinicians use estimates based on current ability (Barker-Collo et al., 2008; Franzen, Burgess, & Smith-Seemillar, 1997).

The use of regression equations to estimate current IQ has some problems, namely regression towards the mean and limited range of scores (Strauss, Sherman, & Spreen, 2006). In practical terms this means that the prediction accuracy is poor for IQ scores which deviate significantly from the mean. For those with true IQ scores above the population mean, generally the estimate will be too low, whilst for those with a true IQ below the population mean, the estimate will be too high (Graves, 2000; Veiel & Koopman, 2001). In addition, as these equations are developed based on group data, the derived IQ for any individual only provides an estimate of functioning, rather than an exact IQ (Hawkins, 1995). Given these limitations, it is recommended that estimated IQ data derived from regression equations are supplemented with clinical observations and information about the person's educational and occupational achievements before conclusions are reached regarding impaired functioning as a result of illness or injury (Strauss et al., 2006).

Vocabulary based tests are a popular means of estimating premorbid ability, due to their high correlation with education / general ability and evidence suggesting that well learned verbal skills are retained even in those with cognitive deterioration (Crawford, Stuart, Cochrane, Parker, & Besson, 1989; Crawford, Stewart, Garthwaite, Parker, & Besson, 1988; Lezak et al., 2004; Nelson & McKenna, 1975; Strauss et al., 2006). Tests requiring only a one or two word answer, such as word reading, are more accurate than those requiring lengthy responses such as word definitions (Blair & Spreen, 1989; Nelson & McKenna, 1975; Ruddle & Bradshaw, 1982). These observations formed the basis for the development of

the National Adult Reading Test (NART; Nelson, 1982) which has subsequently become one of the most widely used tests to estimate premorbid intelligence (Crawford, 1992; Crawford, Allan, Cochrane, & Parker, 1990; Crawford, Deary, Starr, & Whalley, 2001).

The NART consists of a written list of 50 irregular words presented in order of increasing difficulty, which the participant is asked to read aloud. The words included on the test do not follow normal grapheme-phoneme and/or stress rules, so an individual would need to be familiar with the word in order to pronounce them correctly (Nelson & Willison, 1991). Thus, accurate word pronunciation relies upon the individual's premorbid recognition of words rather than current level of functioning. The NART was originally standardised against the WAIS and more recently the WAIS R to provide regression equations which predict Full Scale IQ (FSIQ), Performance IQ (PIQ) and Verbal IQ (VIQ; Nelson & Willison, 1991). The NART shows a high correlation with general intelligence (r =.71 to .82; Crawford, Stewart, et al., 1989; Sharpe & O'Carroll, 1991) and is most accurate at predicting the variance in VIQ (60 to 72%), followed by FSIQ (55 to 66%) and least accurate for PIQ (32 to 33%) (Crawford, Parker, Stewart, Besson, & De Lacy, 1989; Nelson & O'Connell, 1978). Subsequent studies have revealed that NART scores generally show poor correlations with PIQ (e.g., Blair & Spreen, 1989; Crawford, Stewart, et al., 1989; Nelson, 1982) reflecting the lack of verbal content in this index. In addition, Gladsjo, Heaton, Palmer, Taylor, and Jeste (1999), revealed that NART-2 scores made a unique contribution to estimated premorbid IQ over that of demographic variables for VIQ (20%) and FSIQ (14%), however the contribution to PIQ was small (5%). As a consequence, it has been suggested that reading based tests should be used to estimate VIQ and FSIQ but not PIQ (e.g., Lucas, Carstairs, & Shores, 2003). The NART has high internal consistency (0.90 to 0.93: Crawford, Stewart, et al., 1988; Nelson, 1982), test-retest reliability across a 10-day period (0.98) and inter-rater reliability (0.89 to 0.99: Crawford, Parker, et al., 1989; O'Carroll, 1987).

Studies with the NART revealed that it accurately estimated premorbid IQ in a range of disorders including Alzheimer's dementia (Hart, Smith, & Swash 1986; Nebes, Martin, & Horn, 1984; O'Carroll, Baikie, & Whittick, 1987), schizophrenia (Crawford et al., 1992; Smith, Roberts, Brewer, & Pantelis, 1998), depression (Crawford, Besson, Parker, Sutherland, & Keen, 1987), and traumatic brain injury (TBI; Moss & Dowd, 1991; Riley & Simmonds, 2003). Studies indicate that word reading ability does decline as dementia becomes more severe (Cockburn, Keene, Hope, & Smith, 2000; Fromm, Holland, Nebes, & Oakley, 1991), but in these cases the level of dementia is generally established and cognitive decline is readily apparent (Crawford, 1992). Furthermore, the NART is as accurate as estimates based on demographic variables even in these circumstances (Bright, Jaldow, & Kopelman, 2002). The usefulness of the NART for those with brain injury has also been questioned due to large discrepancies between NART scores (within 12 months of injury) and premorbid IQ estimates based on demographics (Freeman, Godfrey, Harris, & Partridge, 2001). However, the accuracy of IQ scores based on demographic data for those with TBI has been questioned, as people who sustain a TBI may not be representative of the general population (Hawkins, 1995; Riley & Simmonds, 2003), for example, those who experience a TBI tend to have a poorer employment records (e.g., Putnam & Adams, 1992), higher rates of substance abuse (Rimel, Jane, & Bond, 1990) and poor academic performance (Haas, Cope, & Hall, 1987) compared to those with similar demographic backgrounds. Thus, the discrepancy between the NART and demographically predicted IQ scores (described by Freeman et al., 2001) may be due to the demographic equation over-estimating the IQ scores, rather than the inaccuracy of the NART. Furthermore, errors on the NART have been shown to be significantly higher within 12 months of injury compared to over 12 month post-injury (Riley & Simmonds, 2003) which may also provide some explanation for these discrepancies. Thus, in those with TBI, it is recommended that the NART should not be used within 12 months of injury (Strauss et al., 2006).

From this brief review, it is clear that there are limitations associated with using the NART to estimate premorbid IQ. These include the accuracy of the estimated IQ scores for those with high or low IQ, its suitability for those with brain injury, and its inaccuracy in predicting PIQ. However, the NART is less vulnerable than many other tests to the effects of cognitive decline (e.g., Mini Mental State Examination, Wechsler Vocabulary test) and provides a useful measure of premorbid functioning which, when used in conjunction with other clinical observations can assist in determining the extent of cognitive decline.

The NART (Nelson & Willison, 1991) was developed for use in the United Kingdom (UK) and its popularity quickly led to its use in other English speaking populations. However, findings from a study with an English speaking population in South Africa revealed lower correlations between NART scores and WAIS IQ scores compared to UK samples, indicating that adaptations may be needed for use in other countries (Struben & Tredoux, 1989 cited in Lezak et al., 2004). As a result many NART equivalents have been developed for English and non-English speaking countries including the US and Canada (North American Adult Reading Test [NAART]; Blair & Spreen, 1989), America (American National Adult Reading Test [AMNART & ANART]; Hopkins Adult Reading Test [HART]; Gladsjo et al., 1999; Schretlen et al., 2009), Italy (Columbo et al., 2000 cited in Schrauf, Weintraub, & Navarro, 2006; Isella et al., 2005), France (Mackinnon & Mulligan, 2005), Spain (Del Ser, Gonzalez-Montalvo, Martinez-Espinosa, Delgado-Villapalos, & Bermejo, 1997), Argentina (Burin, Jorge, Arizaga, & Paulsen, 2000), Japan (Matsuoka, Uno, Kasai, Koyama, & Kim, 2006), Norway (Vaskinn & Sundet, 2001), Sweden (Rolstad, Nordlund, Gustavsson, Eckerstr□m, & Klang, 2008) and Australia (AUSNART; Hennessy & Mackenzie, 1995).

As yet, a New Zealand (NZ) version of the NART has not been developed,

but two studies have reported on the utility of the UK NART in NZ. Freeman et al. (2001) examined whether NART predicted scores were similar to those based on demographic variables for community participants, a TBI group and orthopaedic controls. Their findings indicated that the demographic equation and the NART predicted scores were similar in the community participants, but NART predicted scores were significantly less accurate in the TBI group (30% showed a large discrepancy between their scores). However, this study did not include any assessment of current ability, so it did not determine the actual accuracy of the NART estimations in the community group.

The second NZ based study (Barker-Collo et al., 2008) compared the accuracy of estimated IQ's from the NART and the Spot the Word Test (from the Speed and Capacity of Language Processing assessment [SCOLP]; Baddeley, Hazel, & Nimmo-Smith, 1993) to current IQ from the WAIS III. The NART FSIQ estimates showed a significant correlation with WAIS III FSIQ for participants of European descent (r = .70), but not for those of Māori descent. In contrast, the Spot the Word Test showed high correlations in both ethnic groups (European r = .70, Māori r = .90). In spite of these high correlations, the Spot the Word Test accurately predicted only 52% of IQ classifications indicating that even this test is not an ideal predictor of premorbid IQ in a New Zealand population.

These findings also raise issues around cultural bias in testing. Although the Māori sample in the latter study was small (n = 13), these participants obtained lower WAIS III IQ scores compared to the NZ Europeans even though the samples were similar in age, education and gender. This bias has been noted in earlier studies using verbal tests (e.g., Ogden & McFarlane-Nathan, 1997). However, replacing unfamiliar words with NZ relevant words has been shown to significantly improve performance (Barker-Collo, 2001; Barker-Collo, Clarkson, Cribb, & Grogan, 2002; Ogden & McFarlane-Nathan, 1997), indicating that existing tests may be successfully modified for use in NZ. In keeping with this researchers have suggested that a New Zealand version of the NART should be developed (Barker-Collo et al., 2008).

Taken together, the preceding review highlights the need for an accurate method of estimating premorbid IQ and suggests that the NART may be a successful tool for this. Thus, this preliminary study aims to: 1) develop a New Zealand version of the NART (a New Zealand Adult Reading Test; NZART); 2) develop regression equations based on a New Zealand sample for the NART and the NZART; 3) determine the accuracy of the NART and the NZART in predicting current IQ.

Method

The study is presented in two parts: Part 1 outlines the development of the NZART, while part 2 describes the evaluation of the NART and the NZART. The study received ethics approval from the Department of Psychology Research and Ethics Committee, University of Waikato.

Part 1

Participants.

An opportunity sample of 20 participants (7 male, 13 female; age range 21 to 46 years) were recruited from the University of Waikato campus. Fourteen participants were of European descent whilst the remainder self-identified as Māori. All participants were New Zealand born and their first language was English.

Measures.

New Zealand relevant words were identified by selecting 72 irregular words (see Table 1) from the New Zealand Oxford Dictionary (Kennedy & Deverson, 2004). The words were selected on the premise that they did not follow the normal grapheme-phoneme and/or stress rules, and that they were likely to be known to, or be encountered by New Zealanders. Initially a larger number of words were identified, and the suitability of each word was discussed at a meeting with 10 colleagues (staff and students from the University) and their presence in contemporary NZ literature was evaluated. These group discussions focused on the familiarity of each word,

Table 1. Initial word selection for the New Zealand Adult Reading Test (NZART). Those selected for the final list are shown first.

Words Retained				Words Discarded[a]			
1.	Caveat	20.	Facetious	39.	Manoeuvre	58.	Wyvern
2.	Debris	21.	Risqué	40.	Whanau	59.	Phlegm
3.	Chameleon	22.	Amygdaloid	41.	Recipe	60.	Inadequate
4.	Torque	23.	Ochre	42.	Gauge	61.	Paradigm
5.	Choir	24.	Meringue	43.	Champagne	62.	Tertiary
6.	Indict	25.	Hippocrates	44.	Paroxysm	63.	Allele
7.	Lingerie	26.	Indices	45.	Tsar	64.	Synapse
8.	Fatigue	27.	Caecum	46.	Eunuch	65.	Eucharist
9.	Impugn	28.	Chassis	47.	Reign	66.	Apophthegum
10.	Crochet	29.	Whenua	48.	Ctenophore	67.	Epistle
11.	Kaitiaki	30.	Insatiable	49.	Guerrilla	68.	Cheyenne
12.	Cognac	31.	Reify	50.	Mortgage	69.	Ci Devant
13.	Sieve	32.	Tacit	51.	Mousse	70.	Bourgeois
14.	Māori	33.	Grotesque	52.	Ménage	71.	Marquess
15.	Epitome	34.	Corps	53.	Unique	72.	Epergne
16.	Colonel	35.	Subpoena	54.	Yacht		
17.	Cologne	36.	Talipes	55.	Touché		
18.	Chaos	37.	Vivace	56.	Legate		
19.	Tourniquet	38.	Inertia	57.	Chaise		

[a]Words 39-63 suitable but not selected. Words 64-72 not selected for reasons shown in Table 2

frequency of use in NZ, and ease of pronunciation. From these discussions, consensus was reached regarding the 'final' list of 72 words which were recorded in written form.

The National Adult Reading Test (NART 2nd Ed; Nelson & Willison, 1991) consists of a list of 50 phonetically irregular words, presented in order of increasing difficulty (for a full list of words see source). Participants are asked to read these words aloud. The words are relatively short to minimise the possible adverse effects of stimulus complexity that may occur in subjects with dementia. A NART error score is inserted into a regression equation to predict a FSIQ score. Verbal and Performance IQ scores can also be predicted using alternative equations.

Procedure.

Initially, the 72 NZ words (displayed in Table 1) and the 50 NART words were recorded in written form, together with their correct pronunciation according to the International Phonetic Alphabet. To ensure the researcher understood the correct pronunciation of the words and to standardize scoring, a linguistics professor from Waikato University recorded the correct pronunciation of each of the words on to a dictaphone. Subsequently, a written list was compiled of the 72 NZ words randomly interspersed among the 50 NART words (the NART words remained in the order presented in the manual), resulting in a list comprising 122 words.

Participants were given a brief overview of the research in written and oral forms. They were assured that their results would remain confidential and they had the right to withdraw from the research at any time. Those who wished to participate provided written informed consent.

After this, each participant was asked to read aloud each of the 122 words at a pace they were comfortable with. Participant's verbal responses were recorded on a Dictaphone for subsequent scoring of correct and incorrect pronunciation. After completing the word reading, participants were asked about their familiarity with the words, how easy they found it to decode words they did not know and what words they found to be easy or hard to pronounce (Schrauf et al., 2006). Notes were made of participant's answers and were used in the selection of a final word list.

Results

The list of NZ words is presented in Table1. The first 38 words were deemed suitable for inclusion in the final version of the NZART. The procedure for discarding words was identical for words from the NART and from the NZ dictionary. Initially, NART words which could not be found in the NZ dictionary (see Table 2), and words which had poor inter-rater reliability were discarded (Rolstad et al., 2008). Subsequently, items that could be decoded (even though they had not been encountered before), and those unfamiliar to all participants were also discarded (Nelson & Willison, 1991). Finally, the suitability and familiarity of the remaining words were discussed with a small group of clinical psychology students and staff (n = 10). This discussion identified 5 more words (see Table 2) whose pronunciation was ambiguous, which were also discarded, leaving 99 words (Nelson & Willison, 1991).

In order to reduce the word list further, a procedure based on Fromm et al. (1991) was adopted. Firstly, correlations were conducted between the accuracy of pronunciation for each item and the WASI Vocabulary score (Burin et al., 2000; Matsuoka et al., 2006). Correlations for all of the 99 words were significant ($p < .01$). After this, for each word, the total number of correct pronunciations was recorded (as there were 20 participants, this figure ranged from 0 to 20 for each word). The words were then ranked based on the number of correct pronunciations (i.e., easiest to hardest) and then grouped based on the number of correct responses (i.e., the words that all participants pronounced correctly were grouped together, and so on). For each group of words, half were randomly selected to be included in the final word list (e.g., if 6 words were pronounced correctly by 10 participants, 3 of these words were selected). Where there was only one word in a group, this was also selected for inclusion. Thus, words selected spanned the entire range of difficulty. This resulted in a final word list of 60 words (see Table 3), which form the New Zealand Adult Reading Test (NZART).

Table 2. Words deemed unsuitable for the NZART and reason for non-selection.

	Word Source	
Reason for discarding	NART	NZ Dictionary
Not in NZ dictionary	Assignate Cellist	
Low inter-rater reliability (Crawford et al., 1989)	Prelate Aeon Puerperal Sidereal Aver	
Words that could be decoded	Radix Capon Banal	Synapse Epistle Eucharist Apophthegum
Unlikely to be encountered in NZ	Drachm Gaoled Campanile	Epergne
Ambiguous pronunciation	Beatify	Cheyenne Ci Devant Bourgeois Marquess

Part 2.

Participants.

Of the 63 participants who volunteered to take part in the study, 48 (75.2%) were female and 15 (23.8%) were male, their age ranged from 17-61 years (mean 25.05; sd = 9.35). The majority of participants had completed high school level qualifications (n = 34, 54%), 14 (22.0%) had degree level qualifications, 5 (7.9%) had obtained a graduate diploma, 5 (7.9%) had Honours

level qualifications and 1 (1.6%) had completed a Masters. Only 4 (6.3%) participants had not completed high school. Half of the participants (n = 32; 50.8%) were New Zealand European, a third (n = 21; 33.3%) self-identified as Māori, 3 (4.8%) were of Pacific Island descent, 2 (3.2%) of Indian descent, and 1 (1.6%) each of Asian, British, African American, Canadian, and Other European heritage. Participants were excluded if they did not speak English as a first language, had a history of substance abuse, psychological illness, head injury, reading or eyesight problems that would affect their ability to undertake the test, and if they had not spent the majority of their life residing in New Zealand.

Measures.

Each participant completed a short demographic questionnaire, the NART 2nd Ed. (Nelson & Willison, 1991), the NZART and the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). Details of the NART and the NZART are presented above.

The WASI is an individually administered, short and reliable test of intelligence (Wechsler, 1999). The test was developed to provide researchers and clinicians with a brief and standardised means of determining IQ. Previously, short forms of the Wechsler Adult Intelligence Scale III (WAIS III) were used, but these lacked both standardisation across clinicians and normative data. The WASI contains four subtests – Vocabulary, Block Design, Matrix Reasoning and Similarities, which together yield measures of Full Scale IQ, Verbal IQ and Performance IQ. The WASI takes approximately 30 minutes to administer (compared to 80 minutes for the WAIS III). The reliability coefficients range from .92 to .98 for VIQ, from .94 to .97 for PIQ and from .96 to .98 for FSIQ and correlations with the relevant WAIS III scores are high (e.g., .92 for FSIQ; Lezak et al., 2004; Strauss et al., 2006; Wechsler, 1999). The WASI was administered and scored according to the instructions in the test manual.

Procedure.

The study was advertised via posters and flyers around the University of Waikato. Potential participants contacted the researcher via e-mail to obtain more information about the study and to arrange a suitable time and date to meet. Participants were assessed individually in a quiet room either at the University, or at another mutually convenient location.

At the start of each session, participants were provided with a written information sheet, which the researcher also explained to them verbally. They then provided written informed consent and completed a short demographic questionnaire. After this participants completed the NART, NZART and the WASI. All participants completed the reading tests prior to the WASI but the order of reading test was counterbalanced, with odd numbered participants completing the NART first and even numbered participants reading the NZART first. The NART and WASI were administered according to their respective instruction manuals, whilst the NZART was administered in the same way as the NART (i.e., participants were asked to read the word out loud, at a pace they were comfortable with and they were told that most people would not recognise all of the words). Reponses to the NART and NZART were recorded on Dictaphone and participants were encouraged to attempt all items. Each session took approximately one and a half hours to complete, and participants were offered a break if required. At the end of the session, participants were offered a $5 Warehouse voucher or 1% course credit (for first year students).

The word pronunciations were scored as correct or incorrect by comparison with those provided by the linguistics professor. Two other researchers independently rated a sample of the scoring of the pronunciations to ensure accuracy (the agreement between raters was high, $r > .9$). The WASI and the NART were scored according to published guidelines, for the NZART the total number of incorrect responses were noted. All data were entered into the Statistical Package for Social Sciences 16.0 (SPSS) for analysis.

Table 3. The final wordlist for the New Zealand Adult Reading Test (NZART), arranged in ascending order of difficulty.

	NZART Words						
1.	Debt[a]	16.	Thyme[a]	31.	Reify	46.	Tourniquet
2.	Choir	17.	Lingerie	32.	Cognac	47.	Hippocrates
3.	Aisle[a]	18.	Kaitiaki	33.	Amygdaloid	48.	Quadruped[a]
4.	Chaos	19.	Insatiable	34.	Risqué	49.	Indict
5.	Māori	20.	Courteous[a]	35.	Epitome	50.	Caveat
6.	Nausea[a]	21.	Hiatus[a]	36.	Indices	51.	Corps
7.	Grotesque	22.	Meringue	37.	Chassis	52.	Abstemious[a]
8.	Fatigue	23.	Debris	38.	Superfluous[a]	53.	Topiary[a]
9.	Cologne	24.	Inertia	39.	Leviathan[a]	54.	Idyll[a]
10.	Subtle[a]	25.	Placebo[a]	40.	Subpoena	55.	Vivace
11.	Naïve[a]	26.	Chameleon	41.	Facetious	56.	Labile[a]
12.	Psalm[a]	27.	Equivocal[a]	42.	Ochre	57.	Détente[a]
13.	Torque	28.	Crochet	43.	Impugn	58.	Caecum
14.	Sieve	29.	Tacit	44.	Zealot[a]	59.	Talipes
15.	Whenua	30.	Colonel	45.	Façade[a]	60.	Syncope[a]

[a] from the NART (Nelson & Willison, 1991).

Results

The results section is organised in the following way. Initially the sample demographics are described. This is followed by an examination of the relationship between current IQ, NART and NZART error scores (using correlations). Regression equations based on participants' current IQ scores from the WASI are then presented for the NART and the NZART. The subsequent sections examine differences in current and predicted IQ scores between males and females, and across the two major ethnic groups in the sample (using t tests). The final section of the analyses

examines the ability of the NART (using UK and NZ equations) and the NZART to accurately predict current IQ. This was carried out in two ways; first analyses (repeated measures analysis of variance [ANOVA]) were conducted to explore differences between participants' current IQ and the NART and NZART predicted IQ scores; and second the accuracy of the NART and NZART in placing people in the correct WASI IQ category was examined.

Table 4 summarises the demographic details of the participants. The majority of the sample had completed either high school or tertiary level qualifications, with a greater proportion of females obtaining tertiary level qualifications compared to males (X^2 = 7.83, df = 2, p < .05). Most participants had a household income of less than $10,000 per year. As previously reported, a third of the participants were of Māori decent, around half were of NZ European decent and 16% of the sample were from other ethnic backgrounds. All the participants who were not of Māori or NZ European descent were female. There were no significant differences between these two groups in relation to age or income.

To examine the relationship between the reading tests and current IQ, correlations were conducted between the NART and NZART error scores and WASI IQ scores. These analyses revealed that error scores from the NART and the NZART showed significant negative correlations with the WASI IQ scores (FSIQ [NART r = -.65, NZART r = -.68]; VIQ [NART r = -.70, NZART r = -.74]; PIQ [NART r = -.41, NZART r = -.44]; all p < .01). In each case, the NZART correlated more highly with the current IQ scores. For both measures, the correlations with VIQ were higher than either FSIQ or PIQ.

Regression equations for the NART and NZART.

For the NART, regression equations (based on a UK population) are published in the manual which allow estimation of FSIQ, VIQ and PIQ from the error scores. In order to compare the effectiveness of the NART and the NZART, regression equations based on a New Zealand population were required. Thus, a series of linear regressions were conducted using the NZART and the NART error score to predict scores of WASI FSIQ, VIQ and PIQ. Linear regression is closely linked to correlation and describes the relationship between two variables (Aron, Aron, & Coups, 2009). It differs from correlation in that it allows us to predict one variable from another (in this case premorbid IQ from NART errors). The relationship between the two variables can be expressed as a mathematical equation for a straight line ($y = a + bx$) where y is the predicted score (premorbid IQ), a is the regression constant (where the line cuts the x axis), b is the regression coefficient (slope of the line) and x is the persons score on the predictor variable (NART errors). Clearly it would not be useful to generate a regression equation for each individual's data, instead, one regression equation is produced which minimises errors between the actual and predicted scores (using least squares criterion; Howell, 1997). The equations generated from the data collected for this study are presented below. For the NART, the regression accounted for 42% of the variance of FSIQ scores, 49% of VIQ and 17% of PIQ, and produced the following equations for estimation of IQ:

Predicted FSIQ = 128.78 - (1.033 x NART error) (S.E. est. = 9.31)

Predicted VIQ = 128.02 - (1.162 x NART error) (S.E est. = 9.06)

Predicted PIQ = 121.99 - (.598 x NART error) (S.E. est. = 10.24)

The NZART accounted for slightly more of the variance in each of the IQ measures (46% for FSIQ, 55% for VIQ, and 19% for PIQ). The following equations were generated:

Predicted FSIQ = 124.18 - (.903 x NZART error) (S.E. est. = 8.99)

Predicted VIQ = 123.07 - (1.025 x NZART error) (S.E est. = 8.56)

Predicted PIQ= 119.616 - (.535 x NZART error) (S.E. est. = 10.09)

Effects of gender and ethnicity on IQ scores.

Following this, IQ estimates were derived from the error scores for the NART and the NZART for each participant. For the NART, two estimations of each IQ were calculated based on the UK and NZ equations presented above. Table 5 presents the current IQ scores (from the WASI), NART estimated IQ scores (from UK and NZ equations) and the NZART estimated scores. From Table 5 it can be seen that the current mean IQ scores were in the average range. Male and female participants obtained higher scores on PIQ compared to VIQ. However, males obtained slightly higher scores than females for each of the current IQ measures. Interestingly,

Table 4. Demographic characteristics of the study participants [number (percentage)].

Measure		Male	Female	Total
		(*n* = 15)	(*n* = 48)	(*n* = 63)
Age [mean (*sd*)]		24.93 (8.10)	25.08 (9.78)	25.05 (9.35)
Education				
	Below high school	3 (20%)	1 (2.1%)	4 (6%)
	High school	9 (60%)	25 (52%)	34 (54%)
	Degree or over	3 (20%)	22 (46%)	25 (40%)
Income				
	Under 10,000	7 (47%)	25 (52%)	32 (51%)
	10,000-20,000	5 (33%)	12 (25%)	17 (27%)
	20,000-30,000	3 (20%	3 (6%)	6 (10%)
	30,000+		3 (6.3%)	3 (5%)
	Not reported		5 (10.4%)	5 (8%)
Ethnicity				
	Māori	6 (40%)	15 (31%)	21 (33%)
	NZ European	9 (60%)	23 (31%)	32 (51%)
	Other		10 (21%)	10 (16%)

Table 5. Means and standard deviations of current IQ (WASI), NART and NZART predicted IQ scores for male and female participants.

Measure		Male (*n* = 15)	Female (*n* = 48)	Total (*n* = 63)
WASI				
	FSIQ[a]	104.27 (9.89)	100.52 (12.74)	101.41 (12.15)
	VIQ[a]	98.47 (14.19)	96.83 (12.26)	97.22 (12.64)
	PIQ[a]	109.67 (8.16)	105.04 (11.77)	106.14 (11.14)
NART UK				
	Errors	26.33 (7.27)	26.54 (7.84)	26.49 (7.65)
	FSIQ[a]	97.95 (9.01)	97.69 (9.72)	97.75 (9.48)
	VIQ[a]	97.38 (8.20)	97.14 (8.93)	97.20 (8.72)
	PIQ[a]	98.83 (8.00)	98.60 (8.62)	98.66 (8.41)
NART NZ				
	FSIQ[a]	101.57 (7.50)	101.36 (8.09)	101.41 (7.90)
	VIQ[a]	97.42 (8.44)	97.18 (9.11)	97.24 (8.89)
	PIQ[a]	106.24 (4.34)	106.12 (4.69)	106.14 (4.57)
NZART				
	Errors	22.60 (8.45)	26.02 (9.28)	25.21 (9.14)
	FSIQ[a]	103 77 (7.63)	100.68 (8.37)	101.42 (8.25)
	VIQ[a]	99.91 (8.66)	96.40 (9.51)	97.23 (9.37)
	PIQ[a]	107.53 (4.52)	105.70 (4.96)	106.13 (4.89)

FSIQ = full scale IQ; NART UK = National adult reading test UK version; NART NZ National adult reading test NZ; NZART = New Zealand adult reading test; PIQ = performance IQ; VIQ = verbal IQ; WASI = Wechsler abbreviated scale of intelligence.
[a]IQ scores are standardised with a mean of 100 and standard deviation of 15

NART UK estimated FSIQ and PIQ scores were lower than current FSIQ and PIQ scores, while the NART NZ and NZART derived scores were more similar to current IQ scores. Independent t tests revealed no statistically significant differences between males and females for any of these measures, therefore these data were pooled for subsequent analyses.

The next part of the analysis examined differences in current and predicted IQ scores across ethnic groups. This focused on the two majority ethnic groups, Māori and NZ European, as there were too few participants in each of the other groups. There were approximately a third more NZ European participants compared to those of Māori descent. There were no significant differences between the groups with regard to age (Māori mean age = 28.43, *sd* = 11.48; NZ European mean age = 23.97, *sd* = 8.45), level of education or income. Three (14%) of the Māori participants left school prior to completing high school certificate, 8 (38%) completed their high school education and the majority (*n* = 10; 48%) had degree level qualifications. Of the NZ Europeans, 1 (3.1%) had not completed high school, the majority had high school certificate (*n* = 19; 59%) whilst the remainder (n = 12; 38%) had tertiary qualifications. Nearly half (*n* = 9; 43%) of the Māori participants had a household income of less than $10,000 per year, 6 (29%) earned between $10-20,000, 4 (19%) had an income between $20-30,000, and 2 (9%) did not provide information. The majority (*n* = 19; 59.4%) of the NZ European participants earned less than $10,000, 7 (22%) had a household income of $10-20,000, 2 (6%) were in the $20-30,000 bracket, whilst 3 (9.4%) reported the highest income level (over $30,000). One NZ European did not provide the information.

Table 6 presents the participants' scores from the WASI, NART UK, NART NZ and NZART for the Māori and NZ European participants. Overall, the NZ European group obtained higher IQ scores on the WASI, NART and NZART, and made fewer errors on NART and the NZART. Independent t tests revealed that the NZ European participants obtained significantly higher current FSIQ and VIQ scores compared

Table 6. A comparison of test scores across the Māori and NZ European participants, including effect size and confidence intervals (data are presented as Mean (*sd*).

Measure		Māori (*n* = 21)	NZ European (*n* = 32)	*t* (*df* = 51)	Cohen's *d* ± 95% CI
WASI					
	FSIQ	97.24 (10.67)	104.97 (10.61)	2.59*	.73 ± .57
	VIQ	93.05 (10.51)	100.38 (12.29)	2.25*	.63 ± .56
	PIQ	103.81 (8.04)	108.91 (10.54)	1.89	.53 ± .56
NART UK					
	Errors	28.76 (6.24)	25.88 (8.65)	1.32†	.37 ± .56
	FSIQ	94.94 (7.73)	98.52 (10.72)	1.32	.37 ± .56
	VIQ	94.61 (7.11)	97.90 (9.86)	1.32	.37 ± .56
	PIQ	96.16 (6.86)	99.34 (9.51)	1.32	.37 ± .56
NART NZ					
	FSIQ	99.06 (6.44)	102.05 (8.93)	1.32†	.37 ± .56
	VIQ	94.60 (7.25)	97.95 (10.07)	1.32	.37 ± .56
	PIQ	104.79 (3.73)	106.51 (5.17)	1.32	.37 ± .56
NZART					
	Errors	27.90 (7.85)	23.56 (9.85)	1.70	.48±.56
	FSIQ	98.98 (7.08)	102.90 (8.90)	1.70	.48±.56
	VIQ	94.47 (8.04)	98.91 (10.09)	1.70	.48±.56
	PIQ	104.69 (4.20)	107.01 (5.27)	1.70	.48±.56

* $p<.05$; [a]IQ scores are standardised with a mean of 100 and standard deviation of 15. † *t* values are the same because they are derived from the same error scores. 95% CI = 95% confidence intervals of effect size; Cohen's *d* = effect size; FSIQ = full scale IQ; NART UK = National adult reading test UK version; NART NZ National adult reading test NZ; NZART = New Zealand adult reading test; PIQ = performance IQ; VIQ = verbal IQ; WASI = Wechsler abbreviated scale of intelligence.

Figure 1. Means and Standard deviations of full scale IQ scores from the WASI, NART UK, NART NZ and NZART across current IQ category. * = $p < .05$ compared to current (WASI) IQ

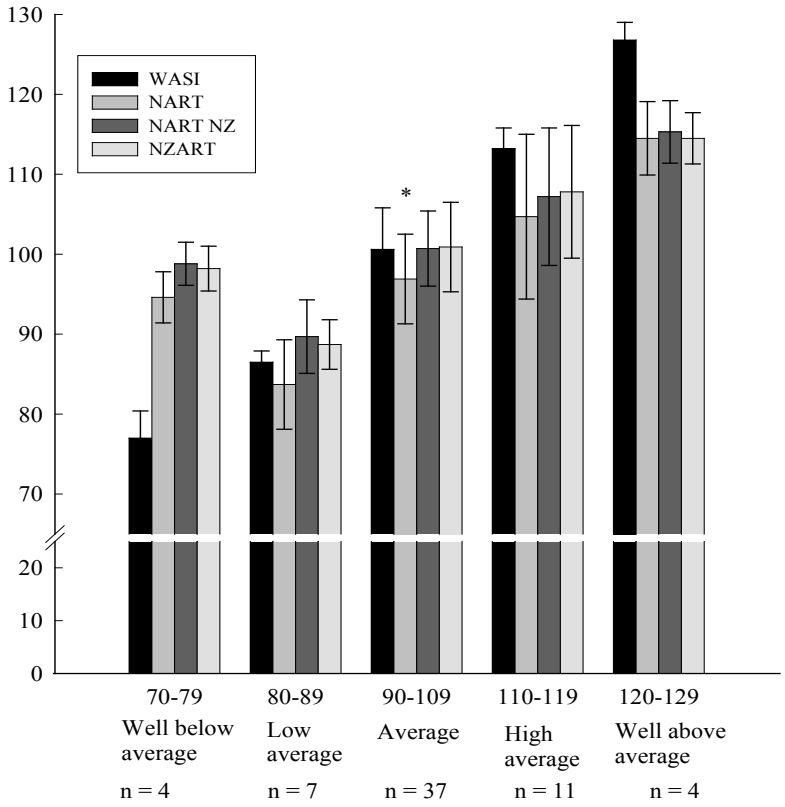
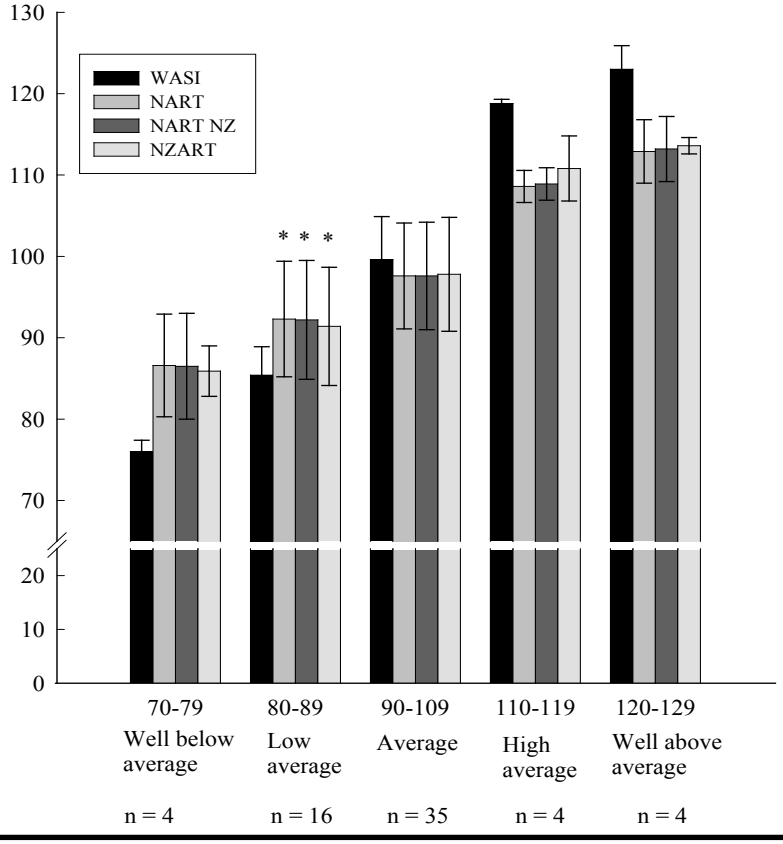


Figure 2. Means and Standard deviations of verbal IQ scores from the WASI, NART UK, NART NZ and NZART across current IQ category. * = $p < .05$ compared to current (WASI) IQ



to the Māori participants. There were no statistically significant differences on any of the other measures, although the effect sizes were medium for most of the comparisons (see Table 6).

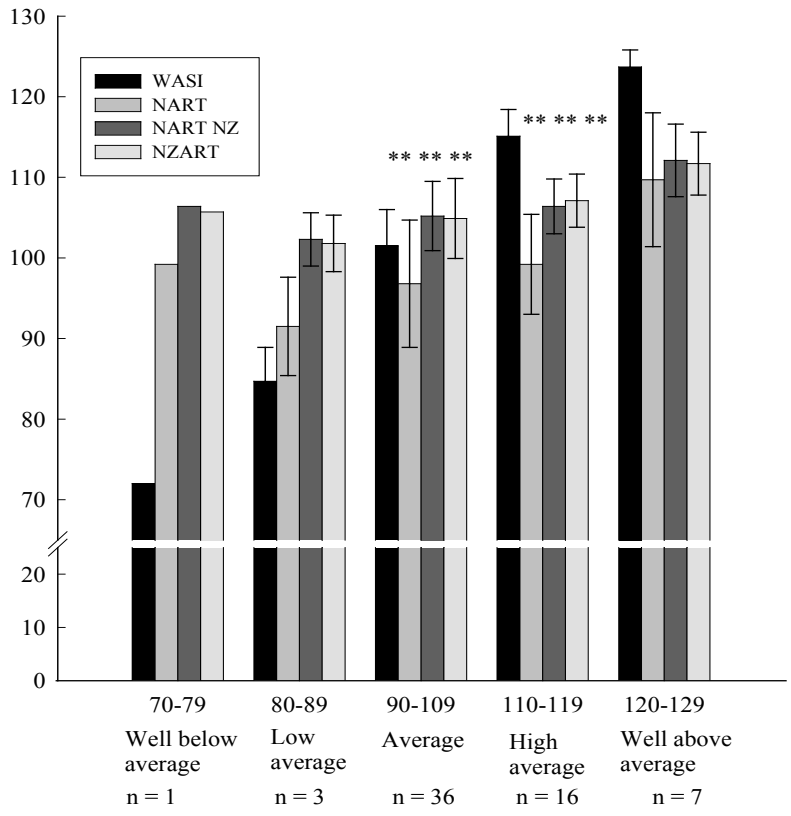
Accuracy of the premorbid IQ estimations.

In order to determine the accuracy of the NART and NZART in predicting current IQ, WASI data were grouped according to categories commonly used to describe IQ scores (i.e., well below average [70-79]; low average [80-89]; average [90-109]; high average [110-119]; well above average [120-129]). The mean IQ scores (for each measure) for participants in each of these categories were then calculated. Initially, these analyses were performed separately for Māori and NZ European, but as findings did not differ between the groups, data were pooled and are presented here for the whole sample (Figures 1-3).

Figure 1 presents the FSIQ scores across each of the above categories. This figure shows that the NART and NZART substantially over-estimated FSIQ in the 70-79 range. For the low average and average ranges, estimations were much more accurate, particularly for the NART NZ and the NZART. Both the tests under-estimated current IQ in the above average categories. Figure 2 presents the data for VIQ. In the lower two categories the NART and NZART over-estimated VIQ scores. In the middle range, the estimated scores were similar to the current IQ, whilst higher IQ scores tended to be under-estimated. There were few clear differences between the tests. The data for PIQ is presented in Figure 3. Once again, the NART and the NZART over-estimated PIQ in the lower categories (IQs from 70-89). Estimation in the average category was more accurate with the NART UK derived estimation falling slightly below current PIQ and estimations based on NZ regression equations over-estimating these scores. In the high average and well above average groups, PIQ scores were under-estimated by the NART (UK & NZ) and the NZART.

To determine if there were significant differences between the current IQ in each category and the three estimated IQ scores, a series of one-way repeated

Figure 3. Means and Standard deviations of performance IQ scores from the WASI, NART UK, NART NZ and NZART across current IQ category. ** = $p < .01$ compared to current (WASI) IQ.



measures ANOVAs were conducted for the IQ scores within each category (where the sample size was greater than 10). When the overall ANOVA was significant, Bonferroni corrected post hoc tests were conducted to further explore differences between current IQ and the three IQ estimates.

For FSIQ, analyses were conducted for the average (n = 37) and high average (n = 11) groups. The ANOVA was significant for the average group, $F(3, 108) = 10.61$, $p < .001$, $\eta^2 = .23$, with the NART UK FSIQ score being significantly lower than the WASI FSIQ ($p < .01$). In the high average group, ANOVA revealed a significant difference across the IQ scores, $F(3,30) = 6.36$, $p < .005$, $\eta^2 = .39$, however this was due to differences between the estimated IQ scores, rather than between the current and estimated IQs. For the low average VIQ category (n = 16), the ANOVA was significant, $F(3, 45) = 10.93$, $p < .001$, $\eta^2 = .42$, all three estimated scores were significantly higher than current VIQ ($p < .05$), however there were no significant differences between the IQ scores in the average VIQ category (n = 35). For PIQ, ANOVAs were conducted for the average (n = 36) and high average (n = 16) groups. The ANOVA for the average group was significant, $F(3, 105) = 39.54$, $p < .001$, $\eta^2 = .53$; the NART UK estimate was significantly lower and the NART NZ and NZART estimates were significantly higher than current PIQ (all $p < .01$). For the high average group, the analysis revealed a significant effect, $F(3, 45) = 56.18$, $p < .001$, $\eta^2 = .79$, as all three estimated scores were significantly lower than current PIQ (all $p < .05$).

For the final part of the analysis, frequency counts were used to determine the number and percentage of participants within each category that were correctly categorised by the NART (UK & NZ) and NZART estimated IQ scores (see Table 7). Both the NART and NZART were most accurate at classifying IQs within the average range and were least accurate for those in the extreme categories. The NART UK correctly classified a greater number of participants in the low average range while the NART NZ and the NZART were more accurate for those in the high average category. Overall, the NART NZ was most accurate for classifying participants on the basis of FSIQ, whilst the NZART was more accurate for VIQ and PIQ.

Discussion

The first part of the study focused on the development of a New Zealand version of the NART. This resulted in a 60 item reading list comprised of words from the New Zealand dictionary and the original NART, together forming the NZART. The latter part of the study focused on the accuracy of the NART and the NZART in predicting current IQ in a healthy New Zealand sample.

Overall, IQ scores did not did not differ between males and females and over half of the participants obtained current IQ scores within the average range. NZ European participants produced significantly higher WASI FSIQ and VIQ scores than Māori, but there were no significant differences on the NART or NZART. These differences were not explained by educational background or income. These findings are in keeping with those of Barker-Collo et al. (2008) who found that Māori obtained significantly lower WAIS III FSIQ, VIQ and PIQ scores than NZ Europeans, but no differences on the NART. Similarly, Ogden, Cooper, and Dudley (2003) found that Māori participants' performance on vocabulary based tests was significantly poorer than NZ Europeans.

With regard to the accuracy of estimating IQ, the NZART explained a slightly higher proportion of the variance in IQ scores compared to the NART (4% for FSIQ, 6% for VIQ, 2% for PIQ). Overall, the NART and the NZART explained 42% and 46% of the variance in WASI FSIQ, which is somewhat lower than reported in other studies. For example, in a New Zealand sample, Barker Collo et al. (2008) found that the NART explained 49% of the variance in WAIS III FSIQ. Overseas researchers report a range of values: 70% of the variance in IQ scores was explained by Japanese and Spanish versions of the NART (Matsuoka et al., 2006; Schrauff et al., 2006); 66% in a UK sample (using the NART; Crawford et al., 1989), 56% using the NAART in North America (Uttl, 2002), 58% using

Table 7. The proportion (and percentage) of cases correctly categorised by the NART (UK & NZ) and NZART according to current (WASI) IQ category for all participants.

	Current (WASI) IQ category (range)					
	Well below average (70-79)	Low average (80-89)	Average (90-109)	High average (110-119)	Well above average (120-129)	Overall correct category (n = 63)
NART UK						
FSIQ	0/4	5/7	33/37	5/11	1/4	44
	-	(71)	(89)	(45)	(50)	(70)
VIQ	0/4	3/16	30/35	1/4	0/4	34
	-	(19)	(86)	(25)	(0)	(54)
PIQ	0/1	1/3	26/36	1/16	0/7	28
	-	(33)	(72)	(6)	-	(44)
NART NZ						
FSIQ	0/4	4/7	35/37	6/11	1/4	46
	-	(57)	(95)	(55)	(25)	(73)
VIQ	0/4	3/16	30/35	1/4	0/4	34
	-	(19)	(86)	(25)	-	(54)
PIQ	0/1	0/3	31/36	2/16	0/7	33
	-	-	(86)	(13)	-	(52)
NZART						
FSIQ	0/4	4/7	35/37	6/11	0/4	45
	-	(57)	(95)	(55)	-	(71)
VIQ	0/4	4/16	29/35	2/4	0/4	35
	-	(25)	(83)	(50)	-	(56)
PIQ	0/1	0/3	30/36	4/16	0/7	34
	-	-	(83)	(25)	-	(54)

FSIQ = full scale IQ; NART UK = National adult reading test UK version; NART NZ National adult reading test NZ; NZART = New Zealand adult reading test; PIQ = performance IQ; VIQ = verbal IQ; WASI = Wechsler abbreviated scale of intelligence.

the AUSNART (Hennessy & Mackenzie, 1995) and 44% using the NART UK in Australia (Mathias, Bowden, & Barrett-Woodridge, 2007). These studies differ from each other in two important ways; the form of NART used and how current IQ is measured (i.e., using a full or shortened version of the WAIS to assess IQ) and both factors may be important in explaining discrepancies in the findings. Bearing in mind the NZART accounted for less than 50% of the variance in current IQ, future studies may improve this by developing a contextual version of the NZART. Lucas et al. (2003) placed AUSNART words in short sentences to provide context (based on the Cambridge Contextual Reading Test; Beardsall, 1998; Beardsall & Huppert, 1994) which improved the performance of all participants, and the accuracy of VIQ prediction.

The NART and the NZART overestimated IQ at the lower end (under 80) and under estimated IQs at the higher end (over 110). This is a common feature of tests of this type, with range restriction also being well documented (see Lezak et al., 2004; Mathias et al., 2007; Strauss et al., 2006). As a result, these tests provide the most accurate estimations for IQs within the average range (Strauss et al., 2006). In keeping with this, in the current study, the NART and NZART accurately classified around 95% of the participants in the average range (IQs between 90 and 109). However, the NZART was more accurate than the NART NZ in the high average range. Overall, the NART (UK & NZ) and the NZART accurately classified a similar proportion of participants for FSIQ and VIQ (around 70% and 50% respectively) However, the NART NZ and NZART were both more accurate at correctly classifying PIQ compared to the NART UK (52% and 54% for NART NZ and the NZART compared to 44% for the NART UK). For FSIQ, previous studies have indicated that the NART accurately classified 49% of NZ participants (Barker-Collo et al., 2008), a much lower proportion than reported here. This may be due to differences between the current IQ level of the two samples; our sample fell predominantly within the average range, with some in the well below average category. In contrast, the participants in the Barker-Collo et al. (2008) study were all average or above, in fact 18 of their sample were in the superior range (IQ of over 130). Given the NART is most accurate at estimating IQ in the average range this may explain the greater

proportion of accurately categorised cases in the current study.

Turning to the limitations of the study, the participants were self-selected, relatively young, and highly educated. In addition, the majority of participants were female and a third were of Māori descent. Thus, the sample was not representative of the NZ population, which limits the generalisability of the current findings. Further evaluation of the NART and the NZART should be undertaken in a sample, selected from the electoral role, which is representative of the NZ population. In particular, studies need to be conducted with the age groups that are most likely to use the test (e.g., older adults), with a wider range of educational backgrounds. After this, the ability of the NZART to identify cognitive decline in various clinical populations (e.g., early dementia, TBI) needs to be evaluated.

A second limitation of the study relates to the use of the WASI to evaluate current IQ rather than the full WAIS III. Although the WASI correlates well with the WAIS III, administration of the full WAIS would provide a more accurate evaluation of current IQ. Given the recent publication of the WAIS IV, future studies should focus on developing regression equations for the NART NZ and the NZART based on the most current version of the WAIS. Furthermore, given the recent release of the Test of Premorbid Functioning (TOPF; Holdnack & Drozdick, 2009), with regression equations based on the WAIS IV, it would be interesting to determine which of the premorbid IQ tests is most suitable for use in the NZ context.

Overall, this preliminary study revealed that the NZART explained a greater proportion of the variance in IQ than either the NART UK or NART NZ. Furthermore, it accurately estimated the IQ category of around 70% of participants overall, and appears particularly accurate within the average range. Given the limitations outlined above, the NZART is clearly not yet suitable for clinical use however the current findings suggest that additional work to develop the NZART would be worthwhile.

Author Note

TeeJay Halliday is now at Karaka Special Treatment Unit, Waikeria Prison, P.O. Box 225, Te Awamutu, 3800, New Zealand.

This research was supported by a University of Waikato Masters Scholarship awarded to TeeJay Halliday. The authors would like to acknowledge the academic support and advice provided by Dr Jo Thakker.

References

Aron, A., Aron, E. N., & Coups, E. J. (2009). *Statistics for Psychology 5th Edition*, New Jersey: Pearson Prentice Hall.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot the Word Test: A robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology, 32*(1), 55-65.

Barker-Collo, S. L. (2001). The 60 item Boston Naming Test: Cultural bias and possible adaptations for New Zealand. *Aphasiology, 15*(1), 85-92.

Barker-Collo, S., Clarkson, A., Cribb, A., & Grogan, M. (2002). The impact of American content on the California Verbal Learning Test performance: A New Zealand illustration. *Clinical Neuropsychologist, 16*(3), 290-299.

Barker-Collo, S., Bartle, H., Clarke, A., van Toledo, A., Vykopel, H., & Willetts, A. (2008). Accuracy of the National Adult Reading Test and Spot the Word estimates of premorbid intelligence in a non-clinical New Zealand sample. *New Zealand Journal of Psychology, 37*(3), 53-61.

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

Beardsall, L. (1998). Development of the Cambridge Contextual Reading Test for improving the estimation of premorbid verbal intelligence in older persons with dementia. *British Journal of Clinical Psychology, 37*, 229-240.

Beardsall, L., & Huppert, F. A. (1994). Improvement in NART word reading in demented and normal older persons using the Cambridge Contextual Reading Test. *Journal of Clinical & Experimental Neuropsychology, 16*(2), 232-242.

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

Bright, P., Jaldow, E., & Kopelman, M. D. (2002). The National Adult Reading Test as a measure of premorbid intelligence: A comparison with demographic variables. *Journal of the International Neuropsychological Society, 8*, 847-854.

Burin, D. I., Jorge, R. E., Arizaga, R. A., & Paulsen, J. S. (2000). Estimation of premorbid intelligence: the Word Accentuation Test-Buenos Aires version. *Journal of Clinical and Experimental Neuropsychology, 22*, 677-685.

Cockburn, J., Keene, J., Hope, T., & Smith, P. (2000). Progressive decline in NART score with increasing dementia severity. *Journal of Clinical & Experimental Neuropsychology, 22*, 508-517.

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.

Crawford, J. R. (1992). Current and premorbid intelligence measures in neuropsychological assessment. In J. R. Crawford, D. M. Parker & W. W. McKinlay (Eds.), *A Handbook of Neuropsychological Assessment*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Crawford, J. R., Allan, K. M., Cochrane, R. H. B., & Parker, D. M. (1990). Assessing the validity of NART estimated premorbid IQ's in the individual case. *British Journal of Clinical Psychology, 29*, 435-436.

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

Crawford, J. R., Besson, J. A. O., Parker, D. M., Sutherland, K. M., & Keen, P. L. (1987). Estimation of premorbid intellectual status in depression. *British Journal of Clinical Psychology, 26*, 312-314.

Crawford, J. R., Deary, I. J., Starr, J., & Whalley, L. J. (2001). The NART as an index of prior intellectual functioning: A retrospective validity study covering a 66-year interval. *Psychological Medicine, 31*, 451-458.

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

Crawford, J. R., Stewart, L. E., Cochrane, R. H. B., Parker, D. M., & Besson, J. A. O. (1989). Construct validity of the

National Adult Reading Test: A factor analytic study. *Personality and Individual Differences, 10*, 585-587.

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

Del Ser, T., Gonzalez-Montalvo, J. I., Martinez- Espinosa, S., Delgado-Villapalos, C., & Bermejo, F. (1997). Estimation of premorbid intelligence in Spanish people with the Word Accentuation Test and its application to the diagnosis of dementia. *Brain and Cognition, 33*, 343-356.

Franzen, M. D., Burgess, E. J., & Smith-Seemiller, L. (1997). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology, 12*(8), 711-738.

Freeman, J., Godfrey, H. P. D., Harris, J. K. J., & Partridge, F. M. (2001). Utility of a demographic equation in detecting impaired NART performance after TBI. *British Journal of Clinical Psychology, 40*, 221-224.

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

Gladsjo, J. A., Heaton, R. K., Palmer, B. W., Taylor, M. J., & Jeste, D. V. (1999). Use of oral reading to estimate premorbid intellectual and neuropsychological functioning. *Journal of the International Neuropsychological Society, 5*(3), 247-254.

Graves, R. E. (2000). Accuracy of regression equation prediction across the range of estimated premorbid IQ. *Journal of Clinical and Experimental Neuropsychology, 22*(3), 316-324.

Green, R. E. A., Melo, B., Christensen, B., Ngo, L-A., Monette, G., & Bradbury, C. (2008). Measuring premorbid IQ in traumatic brain injury: An examination of the validity of the Wechsler Test of Adult Reading (WTAR). *Journal of Clinical and Experimental Neuropsychology, 30*(2), 163-172.

Haas, J. F., Cope, D. N., & Hall, K. (1987). Premorbid prevalence of poor academic performance in severe head injury. *Journal of Neurology, Neurosurgery and Psychiatry, 50*, 52-56.

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

Hawkins, K. A. (1995). Limitations to the validity of the Barona regression formula and similar demographically-based methods of estimating pre-injury intellectual functioning. *Behavioral Sciences & the Law, 13*, 491-503.

Hebben, N., & Milberg, W. (2002). *Essentials of Neuropsychological Assessment*. New York: John Wiley & Sons, Inc.

Hennessy, M., & Mackenzie, B. (1995). AUSNART: The development of an Australian version of the NART. In J. Fourez & N. Page (Eds.), *Treatment issues and long-term outcomes: Proceedings of the 18th Annual Brain Impairment Conference* (pp. 183-188). Bowen Hills: Australian Academic Press.

Holdnack, J. A., & Drozdick, L. W. (2009) *Advanced Clinical Solutions for the WAIS IV and WMS IV: Clinical and Interpretive Manual*. Texas: Pearson.

Howell, D. C. (1997). *Statistical Methods for Psychology (4th Ed.)*. California: Wadsworth Publishing Company.

Isella, V., Villa, M. L., Forapani, E., Piamarta, F., Russo, A., & Appollonio, I. M. (2005). Ineffectiveness of an Italian NART-equivalent for the estimation of verbal learning ability in normal elderly. *Journal of Clinical and Experimental Neuropsychology, 27*, 618-623.

Kennedy, G., & Deverson, T. (2004). *The New Zealand Oxford Dictionary*. Melbourne: Oxford University Press.

Lezak, M. D., Howiseson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment (4th Ed.)*. New York: Oxford University Press, Inc.

Lucas, S. K., Carstairs, J. R., & Shores, E. A. (2003). A comparsion of methods to estimate premorbid intelligence in an Australian sample: Data from the Macquarie University Neuropsychological Normative Study (MUNNS). *Australian Psychologist, 38*(3), 227-237.

Mackinnon, A., & Mulligan, R. (2005). The estimation of premorbid intelligence in French speakers. *Encephale, 31*, 31-43.

Mathias, J. L., Bowden, S. C., & Barrett-Woodbridge, M. (2007). Accuracy of the Wechsler Test of Adult Reading (WTAR) and National Adult Reading Test (NART) when estimating IQ in a healthy Australian sample. *Australian Psychologist, 42*(1), 49-56.

Matsuoka, K., Masatake, U., Kasia, K., Koyama, K., & Kim, Y. (2006). Estimation of premorbid IQin individuals with Alzheimer's disease using Japanese ideographic script (Kanji) compound words: Japanese version of National Adult Reading Test. *Psychiatry and Clinical Neurosciences, 60*, 332-339.

Moss, A. R., & Dowd, T. (1991). Does the NART hold after head injury? A case report. *British Journal of Clinical Psychology, 30*, 179-180.

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

Nelson, H. E. (1982). *National Adult Reading Test*. Windsor, UK: NFER-Nelson.

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

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

Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test (NART) (2nd ed.)*. Berkshire: NFER-NELSON Publishing Company.

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

Ogden, J. A.,Cooper, E., & Dudley, M. (2003). Adapting neuropsychological assessments for minority groups: A study comparing White and Māori New Zealanders. *Brain Impairment, 4*(2), 122-134.

Ogden, J. A., & McFarlane-Nathan, G. (1997). Cultural bias in the neuropsychological assessment of young Māori men. *New Zealand Journal of Psychology, 26*(2), 2-12.

Putnam, S. H., & Adams, K. M. (1992). Regression-based prediction of long-term outcome following multidisciplinary rehabilitation of traumatic brain injury. *The Clinical Psychologist, 6*(4), 383-405.

Reynolds, C. R. (Ed.). (1997). Estimation of premorbid neuropsychological ability [Special issue]. *Archives of Clinical Neuropsychology, 12*(8).

Riley, G. A., & Simmonds, L. V. (2003). How robust is performance on the National Adult Reading Test following traumatic brain injury? *British Journal of Clinical Psychology, 42*, 319-328.

Rimel, R. W., Jane, J. A., & Bond, M. R. (1990). Characteristics of the head-injured patient. In M. Rosenthal, E. R. Griffith, M. R. Bond, & J. D. Miller

(Eds.), *Rehabilitation of the adult and child with traumatic brain injury* (2nd ed., pp. 8-16). Philadelphia: F.A. Davies.

Rolstad, S., Nordlund, A., Gustavsson, M. H., Eckerström, C., & Klang, O. (2008). The Swedish National Adult Reading Test (NART-SWE): A test of premorbid IQ. *Scandinavian Journal of Psychology, 49*, 577-582.

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

Schrauf, R. W., Weintraub, S., & Navarro, E. (2006). Is adaptation of the Word Accentuation Test of premorbid intelligence necessary for use among older, Spanish-speaking immigrants in the United States? *Journal of the International Neuropsychological Society, 12*, 391-399.

Schretlen, D. J., Buffington, A. L. H., Meyer, S. M., & Pearlson, G. D. (2005). The use of word reading to estimate "premorbid" ability in cognitive domains other than intelligence. *Journal of the International Neuropsychological Society, 11*, 784-787.

Schretlen, D. J., Winicki, J. M., Meyer, S. M., Testa, S. M., Pearlson, G. D., & Gordon, B. (2009). Development of psychometric properties, and validity of the Hopkins Adult Reading Test (HART). *The Clinical Neuropsychologist, 23*, 926-943.

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

Smith, D., Roberts. S., Brewer, W., & Pantelis, C. (1998). Test-retest reliability of the National Adult Reading Test (NART) as an estimate of premorbid IQ in patients with schizophrenia. *Cognitive Neuropsychiatry, 3*, 71-80.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. New York: Oxford University Press, Inc.

Uttl, B. (2002). North American Adult Reading Test: Age Norms, Reliability and Validity. *Journal of Clinical & Experimental Neuropsychology, 24* (8), 1123-1137.

Vaskinn, A., & Sundet, K.(2001). Estimering av premorbid IQ: En norsk versjon av Nataional Adult Reading Test/Estimating premorbid IQ: A Norwegian version of National Adult Reading Test. *Tidsskrift for Norsk Psykologforening, 38*(12), 1133-1140.

Veiel, H. O. F., & Koopman, R.F. (2001). The bias in regression-based indices of premorbid IQ. *Psychological Assessment, 13*(3), 356-368.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio: The Psychological Corporation.

Corresponding Author:
Nicola J. Starkey
School of Psychology,
University of Waikato
Private Bag 3105
Waikato
New Zealand
nstarkey@waikato.ac.nz