

# Further Validation of the New Zealand Test of Adult Reading (NZART) as a Measure of Premorbid IQ in a New Zealand Sample

Irene Tatjana Lichtwark, *Monash University, Australia & University of Waikato*

Nicola Jayne Starkey, *University of Waikato*

Suzanne Barker-Collo, *University of Auckland*

Premorbid IQ estimates are used to determine decline in cognitive functioning following trauma or illness. This study aimed to: 1) further validate the New Zealand Adult Reading Test (NZART) in a New Zealand population and compare its performance to the UK developed National Adult Reading Test, and 2) develop regression formulae for the NZART to estimate Wechsler Adult Intelligence Scale-IV (WAIS-IV) IQ scores. The 67 participants (53 females; 16 Māori), aged 16 to 90 years old (mean age = 46.07, *SD* 23.21) completed the WASI-IV, the NART and the NZART. The NZART predicted Verbal Comprehension Index (VCI) scores slightly better than the NART ( $r = .63$  vs.  $r = .62$ ) and explained 33% of the variance in FSIQ scores. Reasons for developing regression formulae for the NZART are discussed, regression formulas for the NZART based on the WAIS-IV are included and suggestions of alternate ways of determining premorbid IQ are made.

**Keywords:** National Adult Reading Test (NART); New Zealand Adult Reading Test (NZART); premorbid IQ; cross cultural differences; IQ estimation, WAIS-IV

Neuropsychological assessments are frequently conducted to evaluate changes in cognitive functioning resulting from brain injury or degenerative decline (Lezak, 2004). To evaluate cognitive decline, comparisons between current and premorbid functioning need to be made, however as baseline assessment results are rarely available estimates of premorbid functioning are often used. The suitability of these measures depends on the resilience of the underlying cognitive domains to neurological disorders and their strong correlation with current Intelligence Quotient (IQ) in the absence of neurological disorder (Lezak, 2004). As such, one of the most commonly used methods for estimating premorbid IQ focuses on current skills which are relatively resistant to brain injury, such as word reading.

Reading skills, and in particular single word reading skills, have been shown to be relatively free from the

influence of age and gender (Crawford, Parker, & Besson, 1988), psychiatric diseases such as depression and schizophrenia (Crawford, 1992), and are resilient to mild to moderate dementia (McGurn et al., 2004), and mild to moderate traumatic brain injury (TBI) (Watt & O'Carroll, 1999). Reading skills are also highly correlated with IQ, a measure of the individuals' overall level of cognitive functioning (Crawford, Stewart, Cochrane, Parker, & Besson, 1989).

One of the most widely used premorbid assessment tools based on current single word reading skills is the National Adult Reading Test (NART) (Nelson & Willison, 1991), which has been standardised against the Wechsler Adult Intelligence Scale-Revised Edition (WAIS-R). The WAIS scales are viewed as the 'gold standard' of IQ tests to which assessments of premorbid function are compared (Lezak, 2004; Silverman, 2010). The NART was

developed in Great Britain to assess the severity of dementia. It consists of 50 single words which are irregular in their grapheme-to-phoneme translation. Thus, if the reader is unfamiliar with a word they are likely to pronounce it incorrectly (Nelson & Willison, 1991). The NART error score is inserted into a regression formula provided in the test manual to obtain an estimated premorbid IQ score.

The NART has a high split half reliability ( $r = .93$ ), high inter-rater reliability ( $r = .96$  to  $r = .98$ ) and high test-retest reliability ( $r = .98$ ) (Nelson & Willison, 1991). Criterion validity of the NART is reported to be good, explaining between 61% and 72% of the variance in Verbal IQ (VIQ), 55% to 66% of the variance of Full Scale IQ (FSIQ) scores, and 32-33% of the variance of Performance IQ (PIQ) (Crawford, Deary, Starr, & Whalley, 2001; Crawford, Stewart, Cochrane, Parker, et al., 1989; Sharpe & O'Carroll, 1991). The NART has been used extensively in Great Britain and other countries, either in its original form (Barker-Collo et al., 2008; McGurn et al., 2004; Watt & O'Carroll, 1999) or modified to better suit linguistic differences. The North American Adult Reading Test (NAART) (Blair & Spreen, 1989), the American National Reading Test (AMNART) (Gladsjo, Heaton, Palmer, Taylor, & Jeset, 1999) and the AUSNART (Hennessy & Mackenzie, 1995) in Australia are examples of modified versions of the NART. The addition of demographic variables (e.g., age, gender, years of education and occupation (Barona, Reynolds, & Chastain, 1984) to the NART-based regression formulae has been found

to increase the formulae's accuracy of premorbid IQ prediction (Bright, Jaldow, & Kopelman, 2002; Crawford et al., 1988; Crawford, Stewart, Cochrane, Foulds, et al., 1989; Watt & O'Carroll, 1999).

Other reading-based tests of premorbid functioning are also available, for example the Test of Premorbid Functioning (TOPF) (The Psychological Corporation, 2009) and its predecessor the Wechsler Test of Adult Reading (WTAR) (Wechsler, 2001). The WTAR, co-normed with the WAIS-III has been used widely due to its large U.S. normative database. Comparisons between the NART and the WTAR have found that both tests performed very similarly (Mathias, Bowden, & Barrett-Woodbridge, 2007). Unlike the NART-based tests country specific modifications of the WTAR (or the more recently released TOPF) are not yet available and the estimated IQ scores are calculated from regression equations based on US samples, which may not be suitable for use outside of the US. It should also be noted that the NART and its various adaptations are in the public domain and hence can be used at minimal cost. In contrast, the TOPF must be purchased as part of the 'Advanced Clinical Solutions' package (The Psychological Corporation, 2009) at considerable cost.

As indicated by the NART adaptations described above, a language-based tool such as the NART cannot necessarily be used in different countries without modification (Franzen, Burgess, & Smith-Seemiller, 1997). Two issues in particular need to be addressed: First, word use and understanding may differ between English speaking nations. Second, the pronunciation of words may differ between English speaking nations (Franzen et al., 1997). Ignoring these issues could lead to an underestimation of premorbid cognitive function and thus the extent of any impairment (Franzen et al., 1997; Odgen, Cooper, & Dudley, 2003).

This issue was highlighted in a New Zealand-based study by Freeman Godfrey, Harris, and Partridge (2001) who explored the reliability of the NART as a tool to estimate premorbid IQ in people with TBI. The NART was administered to 80 participants with

TBI, 27 orthopaedic patients and 80 community participants. Regression formulae (incorporating demographic factors) developed by Crawford, Stewart, Cochrane, and Foulds (1989) were used to estimate the premorbid IQ for each participant. Findings indicated that 30% of the TBI group, 18% of the orthopaedic group and 11% of the community group had scores which indicated some level of cognitive impairment. In contrast, the rate of impairment in the community group was much higher than the rate of 1% reported in a Scottish population using the same NART-based regression formulae (Crawford, Stewart, Cochrane, Foulds, et al., 1989). This suggests that the regression formulae and perhaps the NART itself may not be a reliable means of estimating premorbid IQ in a New Zealand population.

Further questions regarding the suitability of the NART for use in New Zealand were raised in two other studies. Odgen et al (2003) assessed 20 Māori and 20 non-Māori participants with several neuropsychological tests, some of which had been modified to better suit Māori participants (seven words in the Māori language were added to the WAIS-R Vocabulary subtest and a Design Fluency Test was included). Māori participants scored significantly lower than non-Māori on the non-modified tests while there were no significant differences between Māori and non-Māori on the modified tests. Additionally, Māori participants scored similarly to New Zealand European participants on the modified vocabulary and visuo-spatial skills related tests. These authors concluded that more culturally suitable assessment tools should be available to lessen the disadvantages that Māori people face in testing situations (Odgen et al., 2003). Even though the NART was not used in this study the results still highlight the importance of adjusting a measure to the population sampled.

Barker-Collo et al. came to a similar conclusion in 2008 after exploring the accuracy of the NART with a non-clinical sample of 89 New Zealanders, 14 of whom were of Māori descent (Barker-Collo et al., 2008). These researchers compared participants' scores on the WAIS-III, the NART

and Spot-the-Word Test (STW). STW is a test of word recognition used to indicate premorbid ability where the participants are asked to identify the true word in each of 60 word pairs (the other word in each pair is a made up word without meaning) (Baddeley, Hazel, & Nimmo-Smith, 1992). For the New Zealand European participants the NART and STW scores correlated highly with the WAIS-III FSIQ scores ( $r_{\text{NART}} = .70, p < 0.01, r_{\text{STW}} = .70, p < 0.01$ ), while for Māori participants there was no significant correlation between NART and WASI-III FSIQ scores. Interestingly, the WAIS-III FSIQ correlated highly with the STW scores for Māori participants ( $r_{\text{STW}} = .91, p > 0.01$ ). Despite this high correlation the STW was only able to estimate 52% of the current IQs correctly. Barker-Collo et al. concluded that the NART was particularly unsuitable for people of Māori descent, probably as a result of differing word familiarity, and called for the development of a New Zealand version of the NART or at least for New Zealand specific regression formulae (Barker-Collo et al., 2008). Even though the study only included 14 Māori participants, the findings are in line with those of others and suggest that further research is warranted.

Awareness of the issues with the assessment of premorbid IQ in New Zealand led to the development of the New Zealand Adult Reading Test (NZART) by Starkey and Halliday (2011). The NZART consists of 60 single words which are irregular in their grapheme-to-phoneme translation (the NZART is 10 words longer than the NART to avoid possible ceiling effects). To better reflect the word familiarity of New Zealanders, the word order was changed, and 28 words of the NART were replaced with words more commonly used in New Zealand, for example 'Meringue' and 'Whenua'. To increase the cultural suitability of the test, the NZART contained three Te Reo Māori words that are familiar to most New Zealanders. All words selected for the inclusion in the NZART were in the New Zealand Oxford Dictionary, were irregular in their grapheme-to-phoneme encoding and were likely to be known by most New Zealanders. To validate the NZART Starkey and Halliday

(2011) administered the NZART, the WASI (a short form of the WAIS-III), and the NART to 63 participants. The participants' ages ranged from 17 to 61 years old (mean age = 25.05 years;  $SD = 9.35$ ), 48 (75.2%) were female, and half ( $n = 33$ ; 50.8%) of the participants were of European descent, 21 (33.3%) were of Māori descent, and 6 (9.6%) were of other descent. All but three participants had completed high school and almost half of all participants had university level degrees. Estimated IQ scores were calculated using the NART error scores and the original NART formulae ( $NART_{GB}$ ). In addition, regression analyses were undertaken to develop a New Zealand equation for the NART ( $NART_{NZ}$ ), and a regression equation for the NZART.

The new NART equations explained 42%, 49% and 17% of the variance of Full Scale IQ, Verbal IQ and Performance IQ (FSIQ, VIQ and PIQ), respectively. The regression formulae for the NZART were able to explain 46%, 55 % and 19% of the variance of FSIQ, VIQ and PIQ respectively. In addition, the NZART formula was more accurate in estimating IQ score category for FSIQ, VIQ and PIQ across all IQ levels.

While these findings and those of Barker-Collo et al. (2008) are promising, there are still several issues to be addressed. Firstly, the regression formulae for the NART and the NZART need validation in a separate sample across a wider age range. Secondly, the regression estimates in these earlier studies were based on the WAIS-III (Barker-Collo et al., 2008) or the WASI rather than the recently released WAIS-IV which has a different index structure to the WAIS-III and WASI. The WAIS-III provided Verbal and Performance IQ scores in addition to Full Scale IQ. In contrast the WAIS-IV allows calculation of four index scores (Verbal Comprehension Index, Perceptual Reasoning Index, Working Memory Index and Processing Speed Index) as well as the full scale IQ score. The Full Scale IQ (FSIQ) and Verbal Comprehension Index (VCI) were used as the estimates of current IQ in this study as previous work has shown that the NART scores are more accurate at predicting FSIQ and scores based on

verbal ability rather than performance scores (Crawford, Parker, Stewart, Besson, & De Lacey, 1989; Franzen et al., 1997). A further reason for this choice is that clinicians would find estimates of FSIQ most useful.

Therefore the present study was undertaken to validate the NZART in a separate sample, with a wider age range, and to develop regression formulae for the NZART based on the WAIS-IV. It was expected that the NZART estimated IQ scores would be more accurate than those based on the NART.

## Method

### Participants

To be eligible for the study, participants had to have been born in New Zealand, speak English as their first language and have a health history free of neurological conditions (e.g., stroke, TBI). Of the 75 participants who volunteered for the study, 8 were excluded because of a history of mild stroke or TBI. The demographic characteristics of the sample are shown in Table 1. The sample was predominantly female (76.8 %) and included participants across a wide age range (16-90 years). Overall, 75.4% of the participants self-identified as New Zealand European and 21.7% were of Māori descent. The male participants had spent slightly more years in formal education than the female participants, and most participants' occupations were classified in the skilled labourer category (e.g., trades people and nurses). Female participants had spent more time in New Zealand on average than the male participants.

### Materials

All participants completed the Australia and New Zealand adaptation of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV), the NART and the NZART. They also completed a short demographic questionnaire which requested information about their age, marital status, ethnicity, education and general health.

**WAIS-IV (Wechsler, 2008).** The WAIS-IV was used as the measure of current IQ. The 10 core subtests of Block Design, Similarities, Digit Span, Matrix Reasoning, Vocabulary, Arithmetic, Symbol Search, Visual

Puzzles, Information and Coding were administered in accordance with the instructions in the test manual (Wechsler, 2008). The Full Scale IQ and Verbal Comprehension Index scores were used in the current study.

### The National Adult Reading Test (NART) (Nelson & Willison, 1991).

The NART is a word reading test developed in Great Britain to estimate the premorbid cognitive functioning in people who are suspected of suffering from cognitive deterioration (Nelson & Willison, 1991). As noted earlier, the test includes 50 irregularly spelt words and scoring is based upon the total number of pronunciation errors. The possible error scores range from 0 (all words correct) to 50 (no words correct).

### New Zealand Adult Reading Test (NZART) (Starkey & Halliday, 2011).

The NZART is the New Zealand version of the NART and was developed as a test of premorbid functioning suitable for the use in New Zealand (Starkey & Halliday, 2011). As noted earlier, the test includes 60 irregularly spelt words (ten additional words were included to avoid ceiling effects) and scoring is based upon pronunciation of each word. Again, the NZART provides a possible error score from 0 (all words correct) to 60 (no words correct).

### Procedure

The 75 participants for this study were students of the University of Waikato, or members of the community. The study was advertised on posters around the University, on the electronic learning platform of the University as well as the Māori network at the University. Older participants (over 40 years of age) were recruited through contacting community groups such as exercise classes, gardening clubs, bowling clubs and through word of mouth. Ethical approval for the study was obtained from the School of Psychology Human Ethics Committee, University of Waikato.

Prospective participants registered their interest in the study during recruitment talks at various clubs and community centres by adding their name and phone number to a list or by contacting the researcher directly after reading one of the posters. The

researcher contacted the volunteers on the list and provided detailed information about the study. Of 77 volunteers, 75 (97.4%) consented to take part. Participants were assessed either in an office at the University of Waikato (60%) or at their request in a quiet place in their home (40%). After ensuring that the participants were familiar with the purpose of the study and answering any questions they might have, participants provided informed consent and 2% course credit (for University students), or gift vouchers to the value of NZ\$20 were provided to participants. The participants were reminded of their right to withdraw from the study at any time without having to give an explanation. The participants' responses were audio recorded during the administration of the NART and NZART for scoring purposes (after gaining permission from the participant). The tests were then administered according to standardised

procedures. The tests of premorbid IQ were administered separately but in the same way. The participants were given a laminated chart with the words in 2 or 3 columns printed in large, bold font (font size 20) for each test. They were asked to read the words out loud at their own pace, after the warning that some of the words might be unknown to them or difficult to pronounce. Incorrect pronunciations were noted by the assessor on a score sheet.

Because the NART and the NZART share 56% of their words there was the possibility of a learning effect which could result in a higher estimation of IQ from the second measure administered. The tests could have been combined but because the purpose of the study was to compare the two tests and the word orders differed it was decided to administer them separately. To counteract any order effects the two measures were presented either side

of the WAIS-IV administration during the test session and the order of administration was alternated between participants. Half of the participants started with the NART and the others began with the NZART. No significant differences were found between these two groups.

The participants took between 90 to 120 minutes to complete the assessments. At the end of the assessment, participants were thanked for their time and any questions they had about the study were answered. All data were entered into SPSS 20.0 for analyses.

In this study the NART-based regression formulae are from the new data supplement in the NART manual (Nelson & Willison, 1991) and will be referred to as NART<sub>GB</sub> in reference to the origin of the sample. Two NZART formulae are used: the formulae

Table 1: Demographic Characteristics of the Sample

Variable	Male <i>n</i> = 14	Female <i>n</i> = 53	Total <i>N</i> = 67
<b>Age [mean(<i>SD</i>)]</b>	41.56 (24.2)	48.58 (23.2)	46.07 (23.2)
Min- max	18 - 89	16 - 90	16 - 90
<b>Ethnicity [n (%)]</b>			
Māori	2 (14.3)	13 (24.5)	15 (21.7)
New Zealand European	12 (85.7)	40 (75.5)	52 (75.4)
<b>Years in formal education [mean(<i>SD</i>)]</b>	15.6 (3.9)	14.23 (2.9)	14.60 (3.2)
<b>Occupation [n (%)]</b>			
Student	4 (25)	8 (15.1)	12 (17.4)
Labourer/ clerical	1 (6.3)	7 (13.2)	8 (11.6)
Skilled labourer	8 (50.0)	36 (67.9)	44 (63.8)
Professional	3 (18.8)	2(3.8)	5 (7.2)
<b>Years in NZ [mean(<i>SD</i>)]</b>	38.19 (22.7)	47.97 (23.6)	45.52 (23.6)
<b>Marital stat. [n (%)]</b>			
Single	8 (50.0)	15 (28.3)	23(33.3)
Married	7 (43.8)	15 (28.3)	22 (31.9)
De Facto	1 (6.3)	10 (18.9)	11(15.9)
Sep /widowed	0	13 (24.5)	13 (18.8)

developed by Starkey and Halliday (2011), the  $NZART_{WASI}$ , and the formula developed in the current study, based on WAIS-IV ( $NZART_{WASIS}$ ).

**Results**

The results are presented in four sections. The first section describes the development of the regression equations based on the WAIS-IV. The next section summarises the performance of the participants on the WAIS-IV, NART and NZART and compares the performance across gender and ethnicity using independent *t*-tests (due to unequal group sizes it was not possible to conduct a 2-way ANOVA). The third section reports the correlation between participants' scores on the NART and NZART and the WAIS-IV. The accuracy of the NART and NZART in predicting the current WAIS-IV IQ is described in the final section. Effect sizes were calculated as Cohen's *d*, with  $d \leq .2$  as small,  $d \leq .5$  as medium and  $d \geq .5$  as large (Aron & Aron, 2006). The incidence of missing data was less than 1% and the assumption of homogeneity of the data

was met.

**Development of Regression Equations (WAIS-IV)**

Initially correlations were calculated between demographic factors and the NZART error scores. Lower NART and NZART error scores were related to more years in formal education, ( $r_{NART} = -.26, r_{NZART} = -.33,$ ) and older age ( $r_{NART} = -.37;$  all  $p$ 's < 0.05). Two new NZART regression formulae were developed to predict WASI-IV FSIQ and VCI. The NZART error score, ethnicity and education (in years) were subjected to a forced entry linear regression. Surprisingly, the demographic variables did not make a significant contribution to the FSIQ or VCI regression equations. Thus, the resulting formulae for the prediction of WAIS-IV FSIQ and VCI were based on NZART error scores only:

$$NZART_{WASIS} FSIQ = 121.56 - 0.65 \times NZART_{ERR}$$

$$NZART_{WASIS} VCI = 122.55 - 0.69 \times NZART_{ERR}$$

**Participants' Test Performance**

The participants' performance on each of the tests is presented in Table 2. Overall, WAIS-IV FSIQ and VCI scores were within the average range. It is of interest to note that males obtained significantly higher WAIS-IV VCI scores compared to females (medium effect size). The estimated IQ scores from NART and NZART were in the average range and there were no statistically significant differences between male and female participants' scores. The average number of errors made on the NART and NZART was 20.81 ( $SD = 7.37$ ) and 20.41 ( $SD = 12.01$ ) respectively. Generally the mean estimated FSIQ scores were similar to the mean of the current WAIS-IV FSIQ however the  $NZART_{GB}$  tended to underestimate FSIQ and VCI.

Descriptive statistics for the Māori and NZ European participants on each of the tests are provided in Table 3. Both groups of participants obtained WAIS-IV FSIQ and VCI scores in the average range, and the IQ estimates were also

Table 2: Descriptive Statistics (Mean and Standard Deviation) for Male and Female Participants' Scores for the WAIS-IV, NART and NZART.

Measure	Overall mean (SD) <sup>a</sup> <i>N</i> = 67	Male mean (SD) <sup>a</sup> <i>n</i> = 14	Female mean (SD) <sup>a</sup> <i>n</i> = 53	<i>t</i> <i>df</i> = 65	Cohen's ( <i>d</i> )
<b>WAIS-IV</b>					
WAIS FSIQ	109 (11.64)	114 (12.61)	108 (11.17)	1.72	.52
WAIS VCI	109 (12.43)	116 (14.06)	108 (11.42)	2.44*	.64
<b>NART<sub>GB</sub><sup>b</sup></b>					
NART <sub>GB</sub> FSIQ	105 (8.80)	106 (8.20)	105 (9.03)	.05	.11
NART <sub>GB</sub> VIQ	104 (8.09)	105 (7.54)	104 (7.54)	.05	.12
<b>NZART<sub>WASI</sub><sup>b</sup></b>					
NZART <sub>WASI</sub> FSIQ	107 (9.24)	106 (8.68)	107 (9.44)	.51	.11
NZART <sub>WASI</sub> VCI	103 (10.57)	103 (9.90)	104 (10.81)	.50	.10
<b>NZART<sub>WASIS</sub><sup>b</sup></b>					
NZART <sub>WASIS</sub> FSIQ	109 (6.62)	108 (6.13)	109 (6.78)	.48	.15
NZART <sub>WASIS</sub> VCI	109 (7.56)	110 (7.80)	109 (7.56)	.22	.13

Note: <sup>a</sup> IQ scores have mean of 100, *SD* = 15.

<sup>b</sup> Scores are not scaled for age.

\*denotes statistical significance at  $p < .05$

NART<sub>GB</sub> calculated from formulae provided by Nelson and Willison (1991).

NZART<sub>WASI</sub> calculated from formulae provided by Starkey and Halliday (2011).

NZART<sub>WASIS</sub> calculated from formulae developed in this study.

Table 3: Descriptive Statistics (Mean and Standard Deviation) for Māori and New Zealand European Participants' Scores for the WAIS-IV and their estimated IQ scores from the NART and NZART.

Measure	Māori mean ( <i>SD</i> ) <sup>a</sup> <i>n</i> = 15	NZ European mean ( <i>SD</i> ) <sup>a</sup> <i>n</i> = 52	<i>t</i> <i>df</i> = 65	Cohen's <i>d</i>
<b>WAIS-IV</b>				
WAIS FSIQ	109 (11.17)	109(11.88)	.11	0
WAIS VCI	110 (8.72)	109 (13.38)	.25	.08
<b>NART<sub>GB</sub><sup>b</sup></b>				
NART <sub>GB</sub> FSIQ	102 (8.24)	106 (8.81)	1.64	.45
NART <sub>GB</sub> VIQ	102 (7.58)	105 (8.10)	1.64	.38
<b>NZART<sub>WASI</sub><sup>b</sup></b>				
NZART <sub>WASI</sub> FSIQ	108 (11.00)	110 (13.75)	.76	.22
NZART <sub>WASI</sub> VCI	102 (10.51)	104 (10.63)	.76	.19
<b>NZART<sub>WAIS</sub><sup>b</sup></b>				
NZART <sub>WAIS</sub> FSIQ	108(6.56)	109 (6.66)	.76	.15
NZART <sub>WAIS</sub> VCI	110 (8.55)	109 (7.33)	.33	.13

Note: <sup>a</sup> IQ scores have mean of 100, *SD* = 15.

<sup>b</sup> Scores are not scaled for age.

\*denotes statistical significance at  $p < .05$

NART<sub>GB</sub> calculated from formulae provided by Nelson and Willison (1991).

NZART<sub>WASI</sub> calculated from formulae provided by Starkey and Halliday (2011).

NZART<sub>WAIS</sub> calculated from formulae developed in this study.

within the average range. Interestingly, there were no statistically significant differences between the Māori and NZ European participants on any of the measures.

### Regression Formulae

The formulae that are compared in this paper are the NART<sub>GB</sub> (Nelson & Willison, 1991), NZART formulae based on the WASI (Starkey & Halliday, 2011; NZART<sub>WASI</sub>) and the NZART formulae developed as part of the current study, based on the WAIS-IV (NZART<sub>WAIS</sub>).

The equations used are as follows:

#### FSIQ

$$\text{NART}_{\text{GB}} \text{ FSIQ} = 130.6 - 1.24 \times \text{NART}_{\text{ERR}}$$

$$\text{NZART}_{\text{WASI}} \text{ FSIQ} = 124.18 - 0.903 \times \text{NZART}_{\text{ERR}}$$

$$\text{NZART}_{\text{WASI}} \text{ FSIQ} = 121.56 - 0.65 \times \text{NZART}_{\text{ERR}}$$

#### VCI

$$\text{NART}_{\text{GB}} \text{ VCI} = 127.4 - 1.14 \times \text{NART}_{\text{ERR}}$$

$$\text{NZART}_{\text{WASI}} \text{ VCI} = 123.07 - 1.025 \times \text{NZART}_{\text{ERR}}$$

$$\text{NZART}_{\text{WAIS}} \text{ VCI} = 122.55 - 0.69 \times \text{NZART}_{\text{ERR}}$$

### Accuracy of the Premorbid IQ Estimations

The scaled scores obtained from the WAIS-IV can be allocated to one of seven categories: (Wechsler, 2008, p. 126): Extremely low (69 and below), Borderline (70-79), Low Average (80-89), Average (90-109), High Average (110-119), Superior (120-129), and Very Superior (130 and above). These

WAIS-IV categories are a common way for clinicians to describe the level of performance of a client in qualitative terms.

To explore the accuracy of each of the formulae in predicting a participant's current WAIS-IV FSIQ and VCI scores, the participant's current IQ category was compared to the category in which they were placed by each of the regression formulae. This was conducted separately for FSIQ and the VCI. The number and percentage of participants who were correctly categorised by their estimated IQ score were calculated (Table 4; note only five of the seven categories are displayed because there were no participants in the 'Extremely Low' and 'Borderline' groups).

For people with extreme scores (Low Average or Very Superior) the estimated IQ scores were largely inaccurate. However, estimated IQs were more accurate for the middle categories (Average, High Average and Superior). All three formulae were most accurate at categorising FSIQ scores in the Average range however the NZART<sub>WASI</sub> and NZART<sub>WAIS</sub> were twice as accurate as the NART<sub>GB</sub> in accurately categorising participants with FSIQ scores in the High Average range. The prediction of VCI was more accurate in the lower categories (from Low Average to High Average) than the higher categories. The NART<sub>GB</sub> was most accurate at estimating VCI in the Average category, placing 90% of participants correctly. The NZART<sub>WASI</sub> formula was also most accurate for the Average VCI category (placing 77% of participants correctly). In contrast, NZART<sub>WAIS</sub> was most accurate in the High Average VCI category (placing 73% of participants correctly).

## Discussion

The aim of the present study was to validate the NZART in a separate New Zealand sample with a wide age range. We also sought to develop regression formulae for the NZART based on the WAIS-IV. It was hypothesised that the NZART estimated IQs would be more accurate than those calculated with the use of the NART.

The hypothesis was only partly supported by the findings because the NART-based formula was most

Table 4: The accuracy of NART and NZART-based formulae for placing participants into the correct FSIQ and VCI category.

	Low Average (% correct) <i>n</i> = 2	Average (% correct) <i>n</i> = 30	High Average (% correct) <i>n</i> = 20	Superior (% correct) <i>n</i> = 13	Very Superior (% correct) <i>n</i> = 2
FSIQ (N = 67)					
NART <sub>GB</sub>	0	22 (73.3)	5 (25.0)	4 (30.8)	0
NZART <sub>WASI</sub>	0	18 (60.0)	9 (45.0)	2 (15.4)	0
NZART <sub>WAIS</sub>	0	19 (63.3)	11 (55.0)	1 (7.7)	0
VCI (N = 67)	<i>n</i> =5	<i>n</i> =27	<i>n</i> =23	<i>n</i> =6	<i>n</i> =6
NART <sub>GB</sub>	1 (20.0)	25 (92.6)	10 (43.5)	0	0
NZART <sub>WASI</sub>	3 (60.0)	21 (77.8)	12 (52.2)	0	0
NZART <sub>WAIS</sub>	5 (100.0)	16 (59.3)	17 (73.9)	1 (16.7)	0

Note: NART<sub>GB</sub> calculated from formulae provided by Nelson and Willison (1991).

NZART<sub>WASI</sub> calculated from formulae provided by Starkey and Halliday (2011).

NZART<sub>WAIS</sub> calculated from formulae developed in this study.

accurate for participants in the Average and Superior FSIQ categories while the NZART-based formulae were most accurate for those in the Low Average and High Average VCI categories. Overall, the estimations were more accurate at predicting VCI than FSIQ. Given the estimates of premorbid IQ and the subtests that make up the VCI are based on verbal skills it is not surprising that reading based tests are more accurate in estimating VCI compared to FSIQ.

The percentage of the variance explained in current IQ scores by the NART and the NZART was low, around 33% for the NZART and only 26% for the NART. These figures are lower than previous reports for the NART (42%) and the NZART (46%) when compared to current WASI scores (Starkey & Halliday, 2011) or when NART estimates were compared to current WAIS-III FSIQ scores (49%) (Barker-Collo et al., 2008). Interestingly, the proportion of variance explained in these latter two studies is also lower than that reported in other overseas validation studies - typically around 50-60% of the variance of FSIQ scores is explained (Crawford, Parker, et al., 1989). These discrepancies may be a result of differences between the groups of participants and the small

sample sizes. Earlier NZ studies were focused on younger, university educated participants, rather than including older community-based participants (Barker-Collo et al., 2008; Starkey & Halliday, 2011).

It was surprising that the NZART did not perform much better than the NART because the NZART was developed in New Zealand and words were selected carefully to ensure they were suitable in terms of familiarity (Barker-Collo et al., 2008), and cultural suitability (Odgen et al., 2003). It also incorporated some frequently encountered Te Reo Māori words. One possible explanation for these findings may relate to differences between the two NZART samples. The development sample for the NZART consisted of mainly young student with a mean age of about half of that of the present study (25 compared to 46 years of age). Use of language changes over the years and different age groups are familiar with different words. The mean age gap of 21 years between the two samples could well have a great influence on word familiarity. This seemed to be the case for the Te Reo Māori words. The three Te Reo Māori words used in the NZART were in positions 5 (Māori), 15 (Whenua) and 18 (Kaitiaki). Words number 15 and 18 had fewer correct

pronunciations than their surrounding words. This indicates that participants in our sample, of which about half were over 45, had more difficulties pronouncing these words correctly than the participants of the development sample.

Overall the sample's current WAIS-IV FSIQ and VCI were in the Average range and males obtained significantly higher VCI scores compared to females however this difference was not apparent in any of the estimated IQ scores. There were no significant differences between Māori and New Zealand European participants on any of the measures. Although these findings differ from earlier studies (Barker-Collo et al., 2008; Barker-Collo, Kelly, Riddick, & de Jager, 2011; Odgen et al., 2003; Starkey & Halliday, 2011), these results should be interpreted with caution because our sample size was small and all Māori participants were current students or graduates of a University (and had a long history of formal education). This is not representative for the general Maori population where only 3.1% reported to have a Bachelor's degree in the 2001 census compared to 7.3 % of the New Zealand European population (Statistics New Zealand, 2001).

## Limitations

Although the current study attempted to address issues with earlier work by ensuring we recruited participants from the community and across a wide age range, our sample size was still small and not representative of the general population in terms of age, years in formal education, ethnicity, gender, and location. Further research should endeavour to recruit a larger study sample which is representative of the general population. This would increase the validity of such a study and would address some of the issues raised in this study pertaining to age and ethnicity. In addition, efforts need to be made to recruit Māori participants in greater numbers possible by basing a future study at a Marae (a communal Māori meeting place).

A further limitation of our study is inherent to the method of estimating premorbid IQ with the use of regression formulae; 'regression towards the mean' is a well-known phenomenon and shared by all studies based on regression formulae. Typically, low current IQ scores get over estimated while high current IQ scores get under estimated. For a review please see Veiel and Kooperman (2001) and Graves (2000).

## Future Directions

The tests to estimate premorbid IQ used in this study did not predict participants' current IQ well and only explained around 30% of the variance in IQ scores. All these tests are based on word recognition and the 'hold/don't hold' paradigm which was proposed in 1954 (Lezak, 2004). These tests are based on the assumption that word recognition is one of the best preserved abilities following injury or illness.

It is possible that the assumptions on which these tests are based are now no longer valid. The exposure to irregular words might not be as common as it was 60 years ago as a result of changes in people's reading habits. A well educated person 60 years ago would have spent a considerable amount of time reading books and as a consequence would have maintained their familiarity with these irregular words. Since the appearance of the personal computer a lot of people, educated and otherwise, spend more time reading short, simply worded

messages or snippets of information on their computer or mobile phone screens. Irregularly spelt words are no longer very common or their spelling has often been simplified (for example 'night' to 'nite'). As a consequence of this, familiarity with irregular words might be decreasing across the population and the ability to pronounce these words may no longer be an indicator of IQ. Further studies using other tests of premorbid IQ such as the TOPF (The Psychological Corporation, 2009) with a large representative sample would be helpful to fully explore the utility of these tests in the New Zealand context.

Another approach to estimating premorbid IQ which does not rely on the use of regression formulae and is therefore not in danger of the regression towards the mean phenomenon, is the Best Performance Method (Lezak, 2004). This method relies on the combination of observation, test results, interviews, assessment of past achievements, school records, employment records and so on to find the patient's best performance level (Lezak, 2004). The assessing clinician decides which information to gather and include and based on these findings will build up a profile of the person before the brain damage which can be compared to the current level of cognitive functioning. The advantage of this method is that the assessment is not restricted to the cognitive abilities tested by the IQ test and can be easily adapted for any individual circumstances. On the other hand the quality of the assessment is very vulnerable to clinician's subjectivity. This method has often been criticised by researchers because in some studies only the best test score of an IQ test subtest was used as a predictor of premorbid IQ which was not Lezak's (2004) original intention. She proposed that these estimates be based on a clustering of scores of abilities.

Because the vocabulary-based tests in our study did not perform as well as expected we recommend that future research looks for alternative methods such as the Best Performance Method or the development of more complex regression equations (such as the Barona estimates) to better estimate premorbid IQ in New Zealand. In reality it is unlikely that a single test can accurately reflect the premorbid functioning of an

individual – however such test scores can provide additional evidence to help piece together a picture of the pre-injury functioning of a client.

So does this study provide evidence to support the validity of the NART or the NZART for use in New Zealand? Disappointingly, it seems the answer is no - the fact that less than a third of the variance in current IQ could be explained by these premorbid IQ tests brings into question the practical utility and clinical validity of these measures.

## Author Note

Irene Lichtwark is now at the School of Psychology and Psychiatry, Monash University, Melbourne, Australia

This research was supported by a Masters Research Scholarship from the University of Waikato to Irene Lichtwark.

## References

- Aron, A., & Aron, E. N. (2006). *Statistics for Psychology* (4th ed.). Upper Saddle River, New Jersey: Pearson Education Inc.
- Baddeley, A., Hazel, E., & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language- Processing Test*. London: Pearson Assessment.
- Barker-Collo, S., Bartle, H., Clarke, A., van Toledo, A., Vykopal, 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, 53-61.
- Barker-Collo, S., Kelly, T., Riddick, E., & de Jager, A. (2011). New Zealand regression formula for premorbid estimation using the National Adult Reading Test. *New Zealand Journal of Psychology*, 40, 47- 55.
- Barona, A., Reynolds, C. R., & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS-R. *Journal of Clinical Neuropsychology*, 8, 169- 173.
- 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.
- Crawford, J. R. (1992). *Current and premorbid intelligence measures in neuropsychological assessment*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.
- 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., & Besson, J. A. O. (1988). Estimation of premorbid intelligence in organic conditions. *British Journal of Psychiatry*, 153, 178- 181.
- Crawford, J. R., Parker, M. D., Stewart, L. E., Besson, J. A. O., & De Lacey, G. (1989). Prediction of WAIS IQ with the National Adult Reading Test: Cross-validation and extension. *British Journal of Clinical Psychology*, 28, 267- 273.
- Crawford, J. R., Stewart, L. E., Cochrane, R. H. B., Foulds, J. A., Besson, J. A. O., & Parker, D. M. (1989). 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. (1989). Construct validity of the National Adult Reading Test: a factor analytic study. *Personality and Individual Differences*, 10(5), 585-587. doi: 10.1016/0191-8869(89)90043-3
- Franzen, M., D, Burgess, E. J., & Smith-Seemiller, L. (1997). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology*, 12, 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.
- Gladso, J. A., Heaton, R. K., Palmer, B. W., Taylor, M. J., & Jeset, D. V. (1999). Use of oral reading to estimate premorbid intellectual and neuropsychological functioning. *Journal of the International Neuropsychological Society*, 5, 247-254. doi: doi:null
- Graves, R. E. (2000). Accuracy of regression equation prediction across the range of estimated premorbid IQ. *Journal of Clinical and Experimental Neuropsychology*, 22, 316- 324.
- 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. pp. 183-188). Bowen Hills: Australian Academic Press.
- Lezak, M., D. (2004). *Neuropsychological assessment*. Oxford: Oxford University Press.
- 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, 49-56. doi: 10.1080/00050060600827599
- McGurn, B., Starr, J. M., Topfer, J. A., Pattie, A., Whiteman, M. C., Lemmon, H. A., . . . Deary, I. J. (2004). Pronunciation of irregular words is preserved in dementia, validating premorbid IQ estimation. *Neurology*, 62, 1184-1186.
- Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test ( NART)* (2nd ed.). Berkshire, Great Britain: NFER- Nelson Publishing Company Ltd.
- Odgen, J., A., Cooper, E., & Dudley, M. (2003). Adapting neuropsychological assessments for minority groups: A study comparing white and Maori New Zealanders. *Brain Impairment*, 4, 122-134.
- Sharpe, K., & O'Carroll, R. (1991). Estimating premorbid intellectual level in dementia using the National Adult Reading Test: A Canadian study. *British Journal of Clinical Psychology*, 30, 381- 384.
- Silverman, W. (2010). Stanford-Binet and WAIS IQ differences and their implications for adults with intellectual disability (aka mental retardation). *Intelligence*, 38, 242.
- Starkey, N. J., & Halliday, T. (2011). Development of the New Zealand Adult Reading Test (NZART): Preliminary findings. *New Zealand Journal of Psychology*, 40, 129- 141.
- Statistics New Zealand. (2001). *New Zealand Census Data 2001*. Wellington, New Zealand: New Zealand Government.
- The Psychological Corporation. (2009). *Advanced Clinical Solutions for WAIS-IV and WMS-IV*. San Antonion, TX: Pearson.
- Veiel, H. O. F., & Kooperman, R. F. (2001). The bias in regression- based indices of premorbid IQ. *Psychological Assessment*, 13, 356- 368.
- Watt, K. J., & O'Carroll, R. (1999). Evaluating methods for estimating premorbid intellectual abilities in closed head injury. *Journal of Neurology*, *Neurosurgery and Psychiatry*, 66, 474-779.
- Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR*.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale- Fourth Edition: Technical and interpretative manual*. San Antonio: The Psychological Corporation.

### Corresponding Author:

Irene Lichtwark,  
School of Psychology and Psychiatry,  
Monash University,  
Melbourne,  
Australia.  
irene.lichtwark@monash.edu

© This material is copyright to the New Zealand Psychological Society. Publication does not necessarily reflect the views of the Society.