Is the test of premorbid functioning a valid measure for Māori in New Zealand?

Margaret Dudley, Kelly Scott and Suzanne Barker-Collo
The University of Auckland, New Zealand

This study examined whether scores on the Test of Premorbid Functioning (TOPF) accurately predicted current performance on the Wechsler Adult Intelligence Scale (Fourth Edition; WAIS-IV) in a sample of 284 neurologically normal Māori (age range 16-90yrs). Participants were recruited from seven areas of New Zealand's (NZ) North and South Islands, as well as from different iwi (tribes) to ensure a representative sample. The hypothesis that the TOPF would not accurately predict current WAIS-IV performance was supported. TOPF scores only accounted for between 32-36% of the variance in WAIS-IV Full Scale Intelligence Quotient (FSIQ) scores as compared to the over 50% typically reported in overseas studies. The TOPF accurately predicted FSIQ categorisation for just over half the sample. These findings suggest that the TOPF is not a valid measure of premorbid functioning for Māori in New Zealand.

There is an abundance of research indicating culture influences performance on neuropsychological tests (Brickman Cabo, & Manly, 2006), with disparities in test scores between majority and minority cultures (Brickman et al., 2006; Kaufman, McLean, & Reynolds, 1988; Razani, Murcia, Tabares, & Wong, 2006). This has relevance for Māori in New Zealand (NZ) as, while Māori comprise 15% of the population (Statistics New Zealand, 2013a), they are disproportionately more likely to be referred for a neuropsychological assessment than their Pākehā European counterparts (Dudley, Wilson, & Barker-Collo, 2014), with much higher rates of traumatic brain injury (Feigin, Theadom, Barker-Collo, Starkey, McPherson, Kahan, & Ameratunga, 2013) and stroke (Harwood, 2010). In NZ, inequity persists as a result of the "import and drop" approach to neuropsychological assessment (Ogden, 2001; Ogden, Cooper, & Dudley, 2003). While tests that were developed and normed overseas are used to assess, diagnose and plan rehabilitation for Māori clients (Ogden, 2001; Ogden et al., 2003); these tests are culturally bound and inaccurate when applied cross-culturally (Ardila, 1995; Brickman, et al, 2006). A potential result of this is misdiagnosis, inappropriate rehabilitation, and inappropriate financial compensation awarded (Ogden, 2001; Ogden et al., 2003).

In the few studies conducted, Māori perform more poorly than Pākehā on tests that rely on Western education and content, and Māori Perform better than Pakeha on tests that measure visuospatial abilities or on tests that have been adapted to include culturally relevant content (Ogden & McFarlane-Nathan, 1997; Ogden et al., 2003). One aspect of neuropsychological assessment, the assessment of premorbid functioning (PF), is of particular importance. Premorbid, or pre-injury functioning is the estimate of an individuals' level of functioning prior to injury/disease onset, and provides a baseline against which their current performance is compared. In most cases PF must be estimated, and specific tests have been designed to produce these estimates. Valid and reliable tests of premorbid ability should correlate highly with intelligence and be resilient to the effects of brain damage (Crawford, Stewart, Cochrane, Foulds, Besson, & Parker, 1989; Crowell, Vanderploeg, Small, Graves & Mortimer, 2002).

Overseas studies of word reading tests typically report that at least 50-60% of the variance in Full scale IQ scores (FSIQ) is explained. For example, Crawford, Deary, Starr and Whalley, (2001) who followed up 179 individuals who had completed an IQ test at age 11 and administered the National Adult Reading Test (NART; Nelson & Willison, 1991) at age 77). The NART consists of a list of 50 unrelated, phonetically irregular words of graded difficulty which must be read aloud, with scoring based upon correct pronunciation. Performances on the NART and IQ were highly correlated (r = 0.73), accounting for 53% of variance. When applying Japanese and Spanish versions of the NART up to 70% of variance in IQ has been explained (Matsuoka, Masatake, Kasia, Koyama, & Kim, 2006; Schrauf, Weintraub, & Navarro, 2006). For example, Matsuoka et al in a normal elderly population (n = 50) compared a Japanese version of the NART (the JART) with the revised Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1997), finding that the JART explained 61% of variance in IQ scores. These authors further reported that JART-predicted IQs were not significantly different between the normal elderly and age, gender and education matched participants with Alzheimer's disease.

Ogden et al. (2003) were the first to look at the premorbid estimation in a Māori sample, examining the Spot the Word (STW) test (Baddley, Emslie, & Nimmo-Smith, 1993).; a test in which the individual must identify which is the real word from a series of 60 pairs of real and made-up words (Baddeley, Hazel, & Nimmo Smith, 1992) . While the authors did not report the relationship between STW scores and intelligence quotient (IQ) scores they did note was that STW scores were significantly correlated with scores on the Vocabulary subtest, suggesting that they both measure verbal IQ. There have been three other studies on premorbid estimation in NZ: Baker-Collo, Bartle, Clarke, van Toledo, Vykopal, and Willetts (2008) who compared the STW and the NART; Starkey and Halliday (2011) who compared the NART and a newly developed NZ Adult Reading Test (NZART); and Lichtwark et al. (2013) who also compared the NART and NZART.

Barker-Collo et al. (2008) compared NART and STW estimates with Wechsler Adult Intelligence Scale (3rd ed; WAIS-III) scores in a sample of 89 NZ adults (75 NZ European; 14 Māori. For the NZ European participants, NART and STW scores correlated significantly with WAIS-III full scale

intelligence quotient scores (FSIQ) (rNART = .70, p<0.01; rSTW=0.70, p<0.01). For Māori participants, there was only a significant correlation between STW and FSIQ scores (rSTW=0.91, p<0.01) and not between NART and FSIQ scores (rNART=.27). This led the authors to conclude that the STW may be a more accurate PF measure for Māori. They hypothesised that cultural bias and differing word familiarity may explain the difference between the NART's ability to predict NZ European and Māori scores on the WAIS-III. The authors noted that replication with a larger sample was needed and called for the development of a NZ based version of the NART. While the NART and STW scores correlated significantly with WAIS-III full scale IQ scores for NZ Europeans, they only accounted for 49% of the variance.

The NZART is based on the same concept as the NART, but the words included are more appropriate to the NZ vernacular (examples include meringue and whenua). Starkey and Halliday (2011) compared the NZART and NART with scores on the Wechsler Abbreviated Scale of Intelligence (WASI); an abbreviated version of the WAIS-III. The sample consisted of 63 participants; 32 NZ European, 21 Māori, and 10 other. The authors initially conducted separate analyses for the NZ European and Māori data, but pooled them together when it was apparent that the findings did not differ. Overall, the NART and NZART explained 42% and 46% of the variance in the WASI FSIQ scores. As with the above study, this is lower than is reported in other international studies. The limitations of the study were its use of the WASI and not a full neuropsychological assessment battery such as the WAIS-III, and that the sample was not representative of the NZ population. The authors stated that additional work to develop the NZART would be worthwhile.

Lichtwark et al. (2013) sought to validate the NZART with a more representative sample, to again compare its performance to the NART and develop regression formulae for the NZART to predict Wechsler Adult Intelligence Scale (4th ed; WAIS-IV) IQ scores. The sample consisted of 67 participants (52 NZ European; 15 Māori). NZ European and Māori data were again analysed together. The percentage of variance explained in the IQ scores by the NZART and NART was lower than the two studies discussed; 33% for the NZART and only 26% for the NART. Lichtwark et al. (2013) called into question the practical and clinical utility of the NART and NZART as a result. The authors again noted that despite their endeavours, the sample size was small and not representative of the NZ population.

Overall, the findings suggest that the STW is an accurate PF measure for Māori, though only in a small sample (Barker-Collo et al., 2008), but that the NART and NZART have limited to no use with the NZ population as a whole (Lichtwark et al., 2013; Starkey & Halliday, 2011). Unfortunately all three studies recruited from the general NZ population and included only very small samples of Māori. As a result, both Starkey and Halliday (2011) and Lichtwark, Starkey and Barker-Collo (2013) highlight the need for further research in this area. The most recent permutation of the Wechsler intelligence scales (the WAIS-IV) was developed alongside a new PF test, the Test of Premorbid Function (TOPF; Wechsler, 2011), which is similar to the NART in presenting individuals with a list of 70 words that have atypical grapheme to phoneme translations which must be read aloud.

We are aware of only one other study that has looked at the applicability of the TOPF in NZ (Lichtwark, 2011; unpublished master's thesis); the results of which support the general trend that word reading tests are not accurate PF measures in NZ. The studies discussed thus far have recruited from a general NZ population and have only included very small samples of Māori. The focus of the current research is on the TOPF when used to estimate premorbid abilities in Māori. Current functioning is most commonly measured by performance on the Wechsler Adult Intelligence Scale (WAIS) and its revised editions (Lezak, Howieson, Bigler, & Tranel, 2012). The focus of the current research will therefore be on the accuracy of the TOPF in predicting scores on the most recent version of the WAIS battery, the WAIS-IV, for a sample of Māori in NZ.

Method

Participants

Participants were 284 adults who self-identified as Māori. Participants were excluded from the study if there was any indication that their cognitive functioning might be compromised by any history of psychiatric, neurological, developmental, behavioural or medical conditions; the same exclusion criteria used for the standardisation sample of the WAIS-IV (Wechsler, 2008). Participants were recruited from seven areas in NZ to ensure a representative sample from urban and rural locations, as well as from different iwi (tribes) from both the North and South Island consistent with the proportion of Māori living on each island based on the NZ Census statistics (Statistics New Zealand, 2013a). Purposive sampling was used to ensure an even split between males and females and to ensure roughly even spread across the age range. Participant age ranged from 16 to 90 years; and were grouped into seven age brackets; 16 to 20 years, 21 to 30 years, 31 to 40 years, 41 to 50 years, 51 to 60 years, 61 to 70 years, 71+. All participants were fluent speakers of English. Almost half (45.8%) had completed a tertiary qualification, and their mean annual income ($22,500) was similar to the 2013 NZ Census (Statistics New Zealand, 2013b). A summary of demographic information about the sample is presented in Table 1.

Measures

All participants completed the Multi-dimensional Model of Māori Identity and Cultural Engagement (MMM-ICE) to assess identity and cultural engagement as Māori (Houkamau & Sibley, 2010). Participants were also administered the TOPF (Wechsler, 2011) and the Australia and NZ adaptation of the WAIS-IV (Wechsler, 2008).

Multi-dimensional Model of Māori Identity and Cultural Engagement (MMM-ICE)

The MMM-ICE is a 47 item questionnaire which takes approximately 20-30 minutes to complete. Its 47 items are focussed on what it means to be Māori (e.g., "I stand up for Māori rights" and "I can sense when I am in a Tapu space").

Table 1. Participant Demographics

Characteristic	N	%
Gender		
Male	140	49.3
Female	144	50.7
Age (years)		
16-20	40	14.1
21-30	41	14.4
31-40	43	15.1
41-50	40	14.1
51-60	40	14.1
61-70	40	14.1
71+	40	14.1
Education (years completed)		
≤5 (primary school)	5	1.8
6-7 (intermediate school)	12	4.2
8-12 (high school)	137	48.2
≥ 13 (tertiary)	130	45.8
Household Income		
$0 - $10,000	32	11.3
$11,000 - $20,000	77	27.1
$21,000 - $30,000	46	16.2
$31,000 - $40,000	28	9.9
$41,000 - $50,000	33	11.6
$51,000 - $60,000	5	1.8
$61,000 - $70,000	23	8.1
$71,000 +	40	14.1

Participants indicate how much they agree or disagree with each statement using a 7-point scale, from 1 (strongly disagree) to 7 (strongly agree). The questionnaire is both a valid and reliable self-report measures of six dimensions of identity and cultural engagement, which include: group membership evaluation (e.g., "I love the fact that I am Māori"), socio-political consciousness ("I stand up for Māori rights"), cultural efficacy and active identity engagement (e.g., " I know how to act the right way when I am on the marae"), spirituality (e.g., "I can sense it when I am in a Tapu place"), interdependent self-concept (e.g., "My Māori identity is fundamentally about my relationships with other Māori"), and authenticity beliefs (e.g., "To be truly Māori you need to understand your whakapapa and the history of your people"); which show good internal reliability (Houkamau & Sibley, 2010; Sibley & Houkamau, 2013), and a reliable six factor structure (Houkamau & Sibley, 2010). As an additional indication of validity, self-reported fluency in Te Reo Māori and reported number of Marae visits within the last month, correlated highly with the cultural efficacy and active identity engagement sub-scale. The MMM-ICE is scored by calculating a participant's average score on each sub-scale (Houkamau & Sibley, 2010), with higher scores indicating stronger identification with that dimension of being Māori.

Test of Premorbid Functioning (TOPF)

The TOPF is a word reading test designed to estimate PF in adults which contains a list of 70 phonetically irregular words in order of increasing difficulty (e.g., 'eye', 'ceilidh'). The individual's pronunciation is scored as correct or incorrect based on North American English, with administration discontinued after 5 incorrect responses, and a total score given out of 70. In this study, participants attempted all 70 items generating two scores: a $TOPF_{discontinue\ raw}$ which is the score with discontinue criteria applied, and a $TOPF_{total\ raw}$ where the discontinue criteria is disregarded. Comparison of these two will allow determining if one provides a better PF estimate for this sample. Both raw scores were converted to standard scores ($TOPF_{discontinue\ SS}$, $TOPF_{total\ SS}$) which have a mean of 100 and SD of 15, using the age-corrected normative tables from the test manual. The test takes approximately 5 to 10 minutes to administer and score, and has a very high degree of reliability (.96-.99; Holdnack & Whipple Drozdick, 2009), test-retest reliability (.89-.95; Holdnack & Whipple Drozdick, 2009) and concurrent validity with the WAIS-IV Full Scale IQ (*r*= .70, Holdnack & Whipple Drozdick, 2009).

Wechsler Adult Intelligence Scale (Fourth Edition)

The WAIS-IV is a neuropsychological test battery that measures global intellectual functioning (Wechsler, 2008). Scores from the 10 core subtests contribute to five index scores: Full Scale IQ (FSIQ), Verbal Comprehension IQ (VCI), Perceptual Reasoning (PSI), Working Memory (WMI), and Processing Speed (PSI). The index scores have a mean of 100 and a standard deviation of 15. Index scores are categorised into seven qualitative categories from extremely low (69 and below) to very superior (130 and above). The WAIS-IV takes approximately 1 ½ hours to administer. All subtests were administered in accordance with the standardised instructions. Reliability and validity of the WAIS-IV has been established for the American population on which it was normed with index reliability coefficients averaged across age groups ranging from .90 to .98 (Wechsler, 2008, average test-retest reliability coefficients ranging from .87 to .96 (Wechsler, 2008).

Procedure

This study was approved by the Auckland University of Technology Ethics Committee. Recruitment took place over a period of 18 months. Participants were recruited through flyers in Māori health clinics and universities, radio advertisements, presentations to Māori community groups and through the principal researcher's contacts as an individual of Māori descent. Once potential participants identified themselves to a member if the research team, they were contacted by phone and a time and date for a face to face meeting was arranged if they met eligibility criteria.

All face to face meetings were conducted by the main researcher or a Māori research assistant with a tertiary health background. The majority (76%) took place in a Māori friendly health clinic. Each participant was invited to have a support person accompany them. Sessions began with a review of the Participant Information Sheet which explained the purpose of the research and the participant's role should they take part, and participants were encouraged to ask questions. If the participant was agreeable, fully informed written consent was given. After consenting to participate, each participant then completed the MMM-ICE, the TOPF and WAIS-IV. The time taken to complete the administration of these was between 2 ½ to 4 hours. To minimise fatigue, participants were told to ask for a break whenever they needed to. All tests were administered in accordance with their respective manuals. The participant was given a small gift voucher at the conclusion of the session to thank them for their involvement.

Particular attention was paid to ensuring that the testing took place in a way that respected Māori cultural values and processes. This included observing tikanga (protocols) relevant to whakawhanaungatanga (building connections) such as mihi (introductions), karakia (prayers/incantations) and the offering of kai (food). Te Reo Māori (the Māori language) was also

spoken during the sessions where appropriate.

Results

Overall Performance

Means and standard deviations obtained by the sample across measures are presented in Table 2. The samples' mean test scores fell in the average range for all WAIS-IV index scores, as did all WAIS-IV subtest scores. The highest mean scores were for Symbol Search, Block Design and Visual Puzzles. The sample's mean TOPF standard scores also fell in the average range regardless of whether the discontinue rule was applied. In terms of the MMM-ICE, the sample's average scores were higher on items measuring group membership evaluation and lowest on items measuring authenticity beliefs.

Impact of Demographic Variables

Correlations were generated to examine the relationship between demographic characteristics (i.e., age, education) and performance on WAIS-IV index scores, TOPF, and MMM-ICE scores (see Table 3). Age was significantly related to the WAIS-IV PRI, and TOPF raw scores (with and without the discontinue rule applied). Years of education was significantly related to all WAIS-IV index scores, and all TOPF scores. There were some significant correlations between index scores and scores on the MMM-ICE. Scores on the socio-political consciousness subscale, for example, were significantly positively correlated with all TOPF scores.

Table 2 Means and SDs produced by the sample (*n*=284) across measures

Measures	Mean	SD
WAIS-IV		
Indices		
Verbal Comprehension Index	97.73	12.58
Perceptual Reasoning Index	102.30	13.15
Working Memory Index.	100.88	12.89
Processing Speed Index	100.57	15.36
Full Scacle Intelligence Quotient	99.95	12.51
Subtests (scaled scores)		
Similarities	9.80	2.55
Vocabulary	9.63	2.81
Information	9.52	2.89
Comprehension	9.98	3.26
Block Design	10.78	2.52
Matrix Reasoning	9.75	2.96
Visual Puzzles	10.81	2.96
Picture Completion	10.59	2.74
Figure Weights	10.19	2.56
Digit Span	10.69	2.57
Arithmetic	9.68	2.82
Letter-Number Sequencing	10.26	2.56
Symbol Search	10.90	3.27
Coding	9.26	2.88
Cancellation	9.16	3.16
TOPF		
$TOPF_{discontinue\ raw}$	42.88	13.63
$TOPF_{total\ raw}$	45.08	12.23
$TOPF_{discontinue\ SS}$	102.11	13.19
$TOPF_{total\ SS}$	104.29	11.82
MMM-ICE		
Group membership evaluation	6.19	1.42
Socio-political consciousness	5.57	1.59
Cultural efficacy & active identity engagement	5.65	1.43
Spirituality	5.77	1.61
Interdependent self-concept	4.63	1.33
Authenticity beliefs	3.98	1.30
Total Score	246.54	33.05

Relationships between the TOPF and WAIS

Bivariate correlations were conducted to examine the relationships between scores on the TOPF and scores on the WAIS-IV (Table 4). Bivariate correlations can be used as measures of degree and direction of relationship between two variables and their function as regression coefficients (i.e., the squared value of the correlation) can be used to estimate proportion of variance in one measure for which another measure accounts. There were significant positive correlations between all TOPF scores and all WAIS-IV index scores. The strongest correlations were between TOPF scores and the VCI, with squared correlations indicating that the TOPF accounting for 40-45% of the variance. The weakest correlations were between all TOPF scores and the PSI. Not applying the discontinue rule generated stronger correlations between the TOPF and FSIQ. With no discontinue rule applied, both TOPF raw score and TOPF standard score each accounted for 36% of the variance in FSIQ. This is as compared to 32% and 34%, respectively, when the discontinue rule was applied"

Table 3 Correlations between age, education, and MMM-ICE, WAIS-IV and TOPF scores

	WAIS-IV Indices					TOPF			
	VCI	PRI	WMI	PSI	FSIQ	$TOPF_{discontinue\ raw}$	$TOPF_{total\ raw}$	$TOPF_{discontinue\ SS}$	$TOPF_{total\ SS}$
Age	0.07	-0.17**	-0.09	-0.09	-0.09	0.20**	0.20**	0.10	0.07
Education	0.32**	0.19**	0.24**	0.15*	0.28**	0.33**	0.33**	0.30**	0.31**
MMM-ICE									
Total	-0.03	-0.10	-0.12*	-0.03	-0.07	0.05	0.06	0.04	0.04
Subscale1	-0.01	0.01	0.02	-0.05	0.001	0.06	0.06	0.04	0.04
Subscale2	0.15*	0.04	0.02	-0.02	0.07	0.16**	0.17**	0.13*	0.12*
Subscale3	-0.14*	-0.11	-0.12*	-0.11	-0.15*	-0.05	-0.05	-0.05	-0.06
Subscale4	-0.05	-0.02	-0.001	-0.11	-0.06	0.04	0.04	0.01	0.00
Subscale5	-0.05	-0.16**	-0.16**	-0.12*	-0.12*	-0.01	-0.02	-0.02	-0.04
Subscale6	-0.23**	-0.08	-0.06	-0.12*	-0.15*	-0.06	-0.08	-0.04	-0.05

* p<.05
** p<.01

Predictive accuracy of the TOPF in relation to WAIS-IV FSIQ categorisation was then examined (Table 5). As seen in Table 5, using TOPF standard scores with the discontinue rule accurately predicted FSIQ categorisation of 53% of the

Table 4 Bivariate correlations between the TOPF and the WAIS-IV Indices

	WAIS-IV Indices				
	VCI	PRI	WMI	PSI	FSIQ
TOPF					
Raw $Score_{discontinue}$	0.64**	0.42**	0.48**	0.30**	0.57**
Raw $Score_{total}$	0.67**	0.42**	0.49**	0.34**	0.60**
$TOPF_{discontinue}$	0.63**	0.43**	0.49**	0.31**	0.58**
$TOPF_{total}$	0.65**	0.43**	0.49**	0.36**	0.60**

* p<.05
** p<.01

sample. Similarly, using TOPF standard scores without the discontinue rule applied predicted the FSIQ categorisation of 53% of the sample.

Regression towards the mean was evident regardless of whether the discontinue rule was used or not, with none of the categorizations in the extremely low or very superior range being accurately predicted.

Table 5 WAIS-IV FSIQ categories accurately predicted by the TOPF IQ category (*n*=284)

		WAIS-IV FSIQ category						
		Extremely Low n= 1	Borderline n=13	Low Average n=45	Average n=159	High Average n=53	Superior n=8	Very Superior n=5
TOPF discontinue category	Extremely Low							
	Borderline	1	**6 (46%)**	1	6	1		
	Low Average		1	**18 (40%)**	20	1		
	Average		5	23	**93 (58%)**	14	2	2
	High Average		1	2	33	**29 (55%)**	2	1
	Superior			1	7	6	**4 (50%)**	2
	Very Superior					2	8	
TOPF total category	Extremely Low							
	Borderline	1	**3 (23%)**	1	4	1		
	Low Average		4	**9 (20%)**	10			
	Average		5	32	**102 (64%)**	13	1	2
	High Average		1	2	36	**30 (57%)**	3	1
	Superior			1	7	7	**4 (50%)**	2
	Very Superior					2		

Note: Bold print represents number and proportion of FSIQ categorisations accurately predicted.

Discussion

The results indicate that the TOPF is not appropriate for predicting current WAIS-IV performance amongst Māori. The percentage of variance in current IQ scores explained by the TOPF was low regardless of whether the discontinue rule was applied or not, accounting for only 32-36% of the variance in FSIQ scores. Similarly, Lichtwark (2011) reported that the TOPF accounted for 26% of the variance in FSIQ scores in a NZ sample. Previous studies have found that other word reading tests only account for a low percentage of the variance in FSIQ scores (Barker-Collo et al., 2008; Lichtwark et al., 2013; Starkey & Halliday, 2011).

Overseas studies of word reading tests typically report that at least 50-60% of the variance in FSIQ scores is explained (Crawford, Deary, Starr & Whalley, 2001; Crawford, Stewart, Cochrane, Parker, & Besson, 1989), and as much as 70% when applying Japanese and Spanish versions of the NART (Matsuoka, Masatake, Kasia, Koyama, & Kim, 2006; Schrauf, Weintraub, & Navarro, 2006).

Not applying the discontinue rule generated slightly stronger correlations between the TOPF and FSIQ, accounting for slightly more variance. This is likely due to differences in how familiar the TOPF words are in a NZ context. The TOPF words are intended to be listed in order of increasing difficulty, however what is considered difficult in the United States is likely to differ to what is considered difficult in NZ. One hundred and two participants in the current study, for example, mispronounced the word 'porpoise', though it is only number 28 of the 70 words. Porpoises are rare in NZ waters with only one known species found in the area (McKay, 2014). Fewer pronunciation mistakes were made on words considered more difficult (e.g., #42- 'plethora' and #37- 'umbrage'). There was little to no difference between the correlations produced by the TOPF raw scores and standard scores with the WAIS indices.

Overall, the TOPF accurately predicted IQ categorisation for only 52-53% of the sample; being most accurate for participants in the average range with regression to the mean being evident regardless of whether the discontinue rule was used or not. This is consistent with previous research into word reading tests which report under-estimation of IQs above average, and an over-estimation of IQs below average (Veiel & Koopman, 2001). In a previous New Zealand study (Barker-Collo et al 2008) the NART accurately predicted 41% of classifications for Europeans and 7% for Māori; whilst the Spot the Word test predicted 52% of classifications correctly for Europeans and for 93% amongst Māori. Unfortunately this was based on a very small sample of Māori (n = 14) and was also in relation to the WAIS-III rather the WAIS-IV so it is difficult to draw any conclusions.

It might be hypothesized that word reading tests do not correlate highly with IQ in NZ samples (despite doing so overseas; see Crawford et al [2001]; Matsuoka et al [2006]; and Schrauf, et al [2006]) because word familiarity is culturally dependent and these tests were developed and normed on overseas populations. The research to date however shows that word reading tests remain an inaccurate predictor of IQ even when the test is developed specifically for a NZ population (Lichtwark & Starkey, 2013). The NZART for example, was specifically developed in NZ to ensure that the test consisted of familiar and culturally appropriate words yet it too only accounts for a relatively low percentage of the variance in IQ scores (Lichtwark et al., 2013; Starkey & Halliday, 2011). Lichtwark et al. (2013) suggest that maybe the assumptions on which word reading tests are based are not valid for this population. They point to the changing nature of reading in an increasingly technology driven world, and suggest that individuals today read less and are exposed to fewer irregularly spelt words. While this may be true, it is difficult to see how these changes would be peculiar to NZ.

Another explanation is that the research in this area is only in its infancy in NZ and more time and work is needed to develop a word reading test that is valid for use. It is important to note that there have only been two published studies on the NZART to date (Lichtwark et al., 2013; Starkey & Halliday, 2011). In developing the NZART, for example, Starkey and Halliday (2011) noted many limitations: including that the test was developed on a sample unrepresentative of

the NZ population as a whole; being young, highly educated and predominantly Pākeha, impacting generalizability of the findings.

Sample's overall performance

The sample's mean test scores fell in the average range for all WAIS-IV indices (VCI, PRI, WMI, PSI, and FSIQ). Previous research has shown that Māori perform more poorly than Pākehā on tests reliant on Western education and content (Ogden & McFarlane-Nathan, 1997; Ogden et al., 2003). The fact that the highest mean scores were in Symbol Search, Block Design and Visual Puzzles is consistent with previous research which suggested Māori have particular aptitude for visuospatial tasks (Ogden & McFarlane-Nathan, 1997), though discrepancies between mean scores was small with all mean scores in the average range.

In reflecting upon this, in light of the very low levels of variance explained in IQ by the TOPF compared to that in other countries, it could be hypothesized that this is due to differences in the New Zealand lexicon as well as to differences in the underlying relationship between reading ability and overall intelligence; with New Zealanders IQ scores perhaps being more reflecting of performance based abilities.

It should be noted that the education level of the sample was relatively high with 45.8% having a tertiary qualification compared with only 10% of the general Māori population (Statistics New Zealand, 2013a). It is well documented that education impacts test performance (Ardila, 1995; Manly et al., 1998a; Manly et al., 2002), and indeed education was significantly correlated to all WAIS and TOPF scores in the sample, which is also consistent with prior research (Barona et al., 1984; Strauss et al., 2006). It is therefore possible that the data presented here represent the 'best' performance on the TOPF, and that in a more representative sample the variance explained could have been even lower; though alternatively a more varied education level could have resulted in greater variability in performance across the TOPF and the WAIS, with this greater variability allowing for the relationships to be more easily detected.

Alternatively, it is possible that having a Māori assessor conduct the testing in a manner that upheld relevant tikanga (protocols) may have reduced participant anxiety and enabled optimal performance. In previous research, Māori participants have stated that they would prefer to be assessed by a Māori clinician as they would have a better understanding of them and their worldview (Dudley et al., 2014).

The majority of scores of the MMM-ICE were not significantly correlated with any test scores. This is inconsistent with previous research suggesting a link between acculturation and performance on neuropsychological tests (Arnold, Montgomery, Castaneda, & Longoria, 1994; Manly et al., 1998b). It is possible that the MMM-ICE is not a good measure of acculturation. It was not designed with this purpose in mind, but as a tool to measure the heterogeneous nature of Māori identity (Houkamau & Sibley, 2010). The socio-political consciousness sub-test of the MMM-ICE most closely resembles the definitions given to acculturation in the literature. Houkamau and Sibley (2010) state that individuals who score low on this sub-test are more likely to endorse dominant ideologies and attitudes, and identify as European while being of Māori ancestry, though this is not directly assessed.

The main limitation of the current study is that that the sample was not representative of the general Māori population in terms of education. Future research should endeavour to recruit a sample that more closely matches the general Māori population in this respect. Another more minor limitation of the current study is that it did not establish whether English or Te Reo Māori was the participants' first language. While most Māori speak English as their first language (Statistics New Zealand, 2013a), the findings may not be able to be generalised to Māori whose first language is not English. Thus, any future studies should include languages used as a factor. Future studies should also endeavour to examine the stability of WAIS-IV and estimators of premorbid ability in Māori over time, as well as determining if these scores are impacted by various diseases, as valid and reliable tests of premorbid ability should not only correlate highly with intelligence but also be resilient to the effects of brain damage (Crawford, et al, 1989; Crowell, et al., 2002).

The findings suggest that the TOPF is not a useful tool for neuropsychologists when estimating premorbid abilities of Māori clients. This begs the question of what approach should be adopted by neuropsychologists working with Māori instead. While single measures of premorbid ability are appealing to the neuropsychology profession, it may be unrealistic to expect a single test alone to accurately assess the premorbid functioning of an individual, with emphasis being made that the results from a premorbid measure only form part of the picture (Ogden et al., 2003; Starkey & Halliday, 2011). This fits with the best performance method advocated by Lezak et al. (2012) where the clinician looks not only at test scores, but other data obtained during the clinical interview such as level of education, employment history and previous achievements. All the information collected can then be used to construct a profile of the individual's level of functioning prior to brain injury or disease. An alternative approach which should be considered is the development of regression formulae, which combine performance on test of premorbid ability with demographic factors known to influence these abilities (e.g., age, education). Within the New Zealand context the factors to used require validation before such a formulae could be developed, further it is noted that these formulae typically perform better in research where premorbid ability for a group is considered, rather than that of an individual (Crawford et al., 1989; Veiel, & Koopman, 2001).

Conclusion

Overall the TOPF was not found an accurate means of estimating premorbid intelligence in this sample of 284 neurologically normal Māori. TOPF scores accounted for between 32-36% of the variance in FSIQ scores and accurately predicted IQ categorisation for only 52-53% of the sample. This is consistent with previous NZ research that has begun to question the continued use of word reading tests as a means of premorbid abilities. Future research is therefore needed to

ascertain whether a reliable and valid NZ specific word reading test can be developed. Alternate methods to premorbid estimation may also need to be considered in light of these findings and the research that has preceded it.

References

Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 17*(1), 143-150.

Arnold, B. R., Montgomery, G. T., Castaneda, I., & Longoria, R. (1994). Acculturation and performance of Hispanics on selected Halstead-Reitan neuropsychological tests. *Assessment, 1*, 239-248.

Baddeley, A., Hazel, E., & Nimmo Smith, I. (1992). *The Speed and Capacity of Language-Processing Test*. Suffolk: Thames Valley Test Company.

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*(3), 53-61.

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

Brickman, A. M., Cabo, R., & Manly, J. J. (2006). Ethical issues in cross-cultural neuropsychology. *Applied Neuropsychology, 13*(2), 91-100.

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., 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.

Crowell, T. A., Vanderploeg, R. D., Small, B. J., Graves, A. B., & Mortimer, J. A. (2002). Elderly norms for the Spot-the-Word test. *Archives of Clinical Neuropsychology, 17*, 123-130.

Dudley, M., Wilson, D., & Barker-Collo, S. (2014). Cultural invisibility: Māori people with traumatic brain injury and their experiences of neuropsychological assessments. *New Zealand Journal of Psychology, 43*(3), 14-21.

Feigin, V. L., Theadom, A., Barker-Collo, S., Starkey, N. J., McPherson, K., Kahan, M., … Ameratunga, S. (2013). Incidence of traumatic brain injury in New Zealand: A population study. *The Lancet Neurology, 12*, 53-64.

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

Harwood, M. (2010). Rehabilitation and indigenous peoples: The Māori experience. *Disability and Rehabilitation, 32*(12), 972-977.

Holdnack, J. A., & Whipple Drozdick, L. (2009). *Advanced clinical solutions for WAIS-IV and WMS-IV: Clinical and interpretive manual*. San Antonio, TX: Pearson.

Houkamau, C. A., & Sibley, C. G. (2010). The Multi-Dimensional Model of Māori Identity and Cultural Engagement. *New Zealand Journal of Psychology, 39*(1), 8-28.

Kaufman, A. S., McLean, J. E., & Reynolds, C. R. (1988). Sex, race, residence, region, and education differences on the 11 WAIS-R subtests. *Journal of Clinical Psychology, 44*, 231-248.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.

Lichtwark, I. T. (2011). *Estimating premorbid IQ in New Zealan*d. (Master's thesis, University of Waikato, Hamilton, New Zealand). Retrieved from http://researchcommons.waikato.ac.nz / handle/10289/6503

Lichtwark, I. T., Starkey, N. J., & Barker-Collo, S. (2013). Further validation of the New Zealand Test of Adult Reading (NZART) as a measure of premorbid IQ in a New Zealand sample. *New Zealand Journal of Psychology, 42*(3), 75-83.

Manly, J. J., Jacobs, D. M., Sano, M., Bell, K., Merchant, C. A., Small, S. A., & Stern, Y. (1998a). Cognitive test performance among nondemented elderly African Americans and whites. *Neurology, 50*, 1238-1245.

Manly, J. J., Jacobs, D. M, Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society, 8*, 341-348.

Manly, J. J., Miller, S. W., Heaton, R. K., Byrd, D., Reilly, J., Velasquez, R. J.,et al The HIV Neurobehavioral Research Center (HNRC) Group. (1998b). The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals. *Journal of International Neuropsychological Society, 4*, 291-302.

Matsuoka, K., Mastake, U., Kasia, K., Koyama, K., & Kim, Y. (2006). Estimation of premorbid IQ in 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.

McKay, W. 2014 'Very rare' porpoise on Dunedin beach. (2014, September 18). *The Press*. Retrieved from http://www.stuff.co.nz/the-press/news/10515523/Very-rare-porpoise-on-Dunedin-beach.

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

Ogden, J. A. (2001). First do no harm. Culturally-appropriate neuropsychological assessment for indigenous people: A position paper. *Brain Impairment, 2*(1), 1-10.

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 neurological assessment of young Māori men. *New Zealand Journal of Psychology, 26*(2), 2-12.

Razani, J., Murcia, G., Tabares, J., & Wong, J. (2007). The effects of culture on WASI test performance in ethnically diverse individuals. *The Clinical Neuropsychologist, 21*(5), 776-788.

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.

Sibley, C. G., & Houkamau, C. A. (2013). The Multi-Dimensional Model of Māori Identity and Cultural Engagement: Item response theory analysis of scale properties. *Cultural Diversity and Ethnic Psychology, 19*(1), 97-110.

Starkey, N. J., & Halliday, T. (2011). Development of the New Zealand Adult Reading Test (NZART): Preliminary findings. *New Zealand Journal of Psychology, 40*(3), 129-141.

Statistics New Zealand. (2013a). 2013 QuickStats about culture and identity. Retrieved from http://www. stats.govt.nz/Census/2013-census/profileand-summary-reports/quickstats-cultureidentity.aspx

Statistics New Zealand. (2013b). 2013 QuickStats about Māori: Work and income. Retrieved from http://www.stats.govt.nz/Census/2013-census/profile-and-summary-reports/quickstats-about-maori-english/work-and-income.aspx

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neurological tests: Administration, norms and commentary* (3rd ed.). New York, NY: Oxford University Press.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811.

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. (1997). *WAIS-III administration and scoring manual*. San Antonio: The Psychological Corporation.

Wechsler, D. (2008). *The Wechsler Adult Intelligence Scale* (4th ed.). *Australian and New Zealand language adapted edition*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2011). *Test of Premorbid Functioning (TOPF)*. San Antonio, TX: Psychological Corporation.

Corresponding Author

Margaret Dudley
Auckland University
Private Bag 92019
Auckland 1142
New Zealand
Email: m.dudley@auckland.ac.nz