

Clinicians' Perspectives on the Validity of Neuropsychological Memory Tests for Use in Aotearoa New Zealand

Grace Johnstone¹, Josh W. Faulkner², Katie M. Douglas³, and Paul M. Skirrow³

¹Department of Psychology, University of Otago

²School of Psychology, Te Herenga Waka Victoria University of Wellington

³Department of Psychological Medicine, University of Otago

Neuropsychological memory testing is vital in the assessment of neurological conditions but evidence suggests that cultural factors can significantly impact upon the validity of the scores obtained. The lack of normative data for different populations may lead to misleading scores/inappropriate diagnoses. The present study aimed to assess clinicians' perspectives around the validity of memory tests for use in NZ. Clinicians (n = 76) across NZ answered an online survey assessing the perceived validity of memory tests within the domains of working memory, visual learning and retrieval, auditory learning and retrieval, and prospective and autobiographical memory. Several of the most commonly used tests with a large language component were judged as lacking validity amongst clinicians. Current findings provide further evidence that common neuropsychological tests may lack validity and that the development of NZ-normative data and/or NZ specific tests are now long overdue.

Keywords: *Keywords: Neuropsychological ; Cognitive; Validity; Culture; Indigenous; Language; New Zealand.*

INTRODUCTION

Neuropsychological testing is a key component of diagnostic and functional assessments for a wide range of neurological and developmental conditions (Lezak et al., 2012). The vast number of neuropsychological tests available to psychologists is continually increasing, with clinicians typically selecting tests based on their own experience, training, or accessibility (Ross-McAlpine et al., 2018).

A key component of neuropsychological tests is their capacity to determine where an individual's cognitive ability sits in comparison to their same-aged peers (Evans, 2010). Appropriate normative data are, therefore, critical to ensure that clinical interpretations are accurate. Unfortunately, normative data tends to be biased towards WEIRD demographics (Western, Educated, Industrialised, Rich, Democratic; Quesque et al., 2022) and can fail to account for the influences of culture, language, education level, and gender (Rivera Mindt et al., 2021; Schaefer et al., 2021). As a result, current normative data may lead to an under- or over-estimation of a person's true ability, which, in turn, may limit the accuracy of their diagnosis.

There is increasing evidence that accounting for cultural factors is critical to ensure the accuracy of neuropsychological tests (Rivera Mindt et al., 2021; Schaefer et al., 2021). For example, research has demonstrated visual perception, spatial reasoning, brain organisation, and brain activation – among other aspects – are affected by culture (Fernández & Abe, 2018). For NZ, in particular, where the population is becoming increasingly diverse, the lack of current culture-specific normative data means that non-European individuals may be more likely to score lower and receive an inappropriate diagnosis (Rabin et al., 2016; Rivera Mindt et al., 2021;

Schaefer et al., 2021; Tan et al., 2021). This is especially concerning given the stigma that can be attached to neurocognitive difficulties and disorders (e.g., Intellectual Developmental Disorders; Jansen-van Vuuren & Aldersey, 2020) and the importance of accurate diagnosis in the development of suitable treatment plans and dispersal of resources (Rabin et al., 2016; Rivera Mindt et al., 2021; Schaefer et al., 2021; Tan et al., 2021).

The current over-dependence on implementing neuropsychological assessment approaches from the UK and North America is particularly problematic as they do not account for Māori and Pasifika communities, the largest and third largest ethnic minority groups in NZ, respectively (Dudley et al., 2016; Statistics New Zealand, 2018). Structural inequities in the healthcare system already place Māori and Pasifika individuals at greater risk of severe physical and mental health outcomes (Steyn et al., 2021; Talamaivao et al., 2020). In order to address these inequities, it is imperative that neuropsychological assessments and treatments are evaluated to ensure their psychometric validity within specific cultures and sub-groups.

To date, few commonly used tests have appropriate normative samples for the NZ context. For example, while the Wechsler Intelligence Scale for Children (WISC-V; Wechsler, 2014a) has NZ samples, other common tests such as the The NEuroPSYchological Assessment (NEPSY-II; Korkman et al., 2007) do not (Wechsler, 2014b). Moreover, while the WISC-V is one of the only tests with normative data for Australian and NZ populations, there were only 19 individuals who identified as Māori included in the sample of 84 New Zealanders, which was not a large enough sample to have included a singular male and female participant who identified as Maori per age group (Wechsler, 2014b). Hence, current

normative data for the WISC-V is unlikely to accurately represent Māori individuals, limiting the ability of clinicians to interpret test data for Māori clients. More recently, Dudley et al. (2022) presented normative data for clinicians to use when interpreting Wechsler Adult Intelligence Scale 4th Edition (WAIS-IV; Wechsler, 2008) scores for Māori clients. While this will help limit unfair interpretations of WAIS-IV test scores for Māori, normative data is needed across all neuropsychological assessments to ensure cultural biases in test score interpretations are reduced.

To address the present lack of cultural standardisation, recent evidence found that nearly 70% of psychologists in NZ made adaptations to neuropsychological tests for children in order to maintain the semantic meaning across cultures (Ross-McAlpine et al., 2018). Perhaps problematically, this is being done in a non-standardised manner, which also impacts the validity of assessment and subsequent interpretation of test data (Howieson, 2019). For these reasons, it is becoming increasingly clear that neuropsychological practises developed in the UK and North America need to be empirically adapted, standardised, and validated within an NZ context.

Previous research examining the cultural validity of specific neuropsychological tests within NZ, such as the Peabody Picture Vocabulary Test (Haitana et al., 2010), WAIS-IV (Dudley et al., 2019), Boston Naming Test (Barker-Collo, 2001), Test of Premorbid Functioning (Dudley et al., 2017), and California Verbal Learning Test (CVLT; Barker-Collo et al., 2002) all demonstrated various cultural biases or factors that needed to be considered in the evaluation of test scores for Māori clients. In general, the majority of tests within the wider domain of neuropsychological memory assessment have not been examined within a NZ context. It is thought that memory tests are potentially more likely to be influenced by language/cultural factors due to the difference in encoding and retrieval practises for familiar and unfamiliar words (Barker-Collo et al., 2002). For example, the cognitive load of remembering an unfamiliar word is much greater than a familiar word, as individuals commonly experience “interference” whereby the brain will relate familiar words with the new word they were exposed to (Barker-Collo et al., 2002; Mashal, 2013).

The present study aimed to understand NZ clinicians' perspectives on the validity of neuropsychological memory tests for use with adult and child clientele. In the current study, clinicians working in NZ were approached to answer an online survey asking about their perceptions on the validity of neuropsychological memory tests for use with NZ clients.

METHOD

Participants and procedure

To recruit registered psychologists, the study was advertised through professional psychology associations in NZ: the NZ Special Interest Group in Neuropsychology (NZSIGN), the NZ College of Clinical Psychologists (NZCCP), and the NZ Psychological Society (NZPsS). Clinicians completed an online survey between November 2021 and January 2022. One-hundred-and-three psychologists consented to partaking in the survey, however, ten participants did not enter any survey data, five only entered demographic information, and a further 12 were excluded for insufficiently answering questions related to cultural validity (i.e., only providing validity ratings for a handful of tests). As a result, the present study had a sample of 76 psychologists from around NZ.

Approval for this study was granted by the University of Otago Human Ethics Committee (Approval number: 21/120). The Ngāi Tahu Research Committee also reviewed and approved this study for relevance to Māori.

Measures

Descriptive Characteristics

Participants self-reported their scope of psychological practise, number of years practising, current context of neuropsychological assessments, and typical clientele.

Perceived Validity Measures

The authors compiled the list of memory tests for clinicians to evaluate. The second, third, and fourth authors identified neuropsychological memory tests from previous clinical experience (8-, 8-, and 18-years experience, respectively) and a review of the background literature. A group from the professional neuropsychology body within NZ (NZSIGN) also reviewed the collated list of memory tests. Test selection focussed on tests that are commonly utilised within formal neuropsychological assessment in clinical practise, rather than tests intended for ‘cognitive screening’ (e.g., Montreal Cognitive Assessment; Nasreddine et al., 2005) or primarily used in research trials (e.g., National Institute of Health Cognition Battery; Weintraub et al., 2013).

Table 1. Descriptive characteristics of the clinicians.

	Adults	Children	Both Adults and Children
Total n	51	7	18
Mean Number of Years Experience (SD)	14.45 (8.2)	10.14 (7.6)	17.22 (10.0)
Context			
District Health Board	12	4	2
ACC Contracts	28	1	10
Corrections	4	0	0
Private Clients	2	2	1
University/Academia, PHO/Primary Care	2	0	1
NZPB Scope(s) of Practice			
General Scope and Other Scopes	2	0	2
Clinical Psychologist	13	3	6
Clinical Psychologist and Neuropsychologist	32	2	7
Neuropsychologist	4	1	2

Memory tests were grouped into broad domains based on the publishers' descriptions of the primary function of each test. The first domain, working memory, contained 11 tests for adults and 10 for children. The second domain, auditory or verbal learning and retrieval, contained 13 tests for adults and 9 for children. The third domain, visual learning and retrieval, contained 17 tests for adults and 6 for children. The fourth domain, prospective and autobiographical memory, contained 4 tests for adults and no current tests were identified to measure prospective or autobiographical memory for children. The Rivermead Behavioural Memory Test for Children (RBMT-C; Aldrich & Wilson, 1991) formerly measured prospective memory in children, however, this test has been discontinued.

Clinicians were presented with each domain of memory containing the list of respective tests and were asked to rate how valid they perceived each test to be for use within NZ (e.g., language, cultural appropriateness). Participants selected their answer from four categories ranging from "significant validity issues", "some validity issues", "reasonably valid", to "completely valid". The survey also allowed clinicians to indicate they did not have sufficient experience to comment on the validity of a test through a fifth possible answer category: "I don't know this test well enough to comment". A textbox was provided under each test for participants to explain particular validity issues or suggestions for improvements. The survey also included an "other" textbox for participants to identify and rate the validity of additional tests not included in the survey.

Data Analysis

The data were analysed and frequency tables were created for descriptive characteristics and the perceived validity of each test. Text comments around the perceived issues with neuropsychological memory tests were sorted into thematic categories curated by the first author and validated by the last author. Only a relatively small proportion of respondents provided qualitative feedback (see below). Furthermore, responses were typically brief, allowing for only surface- or semantic-level coding and

RESULTS

Descriptive Characteristics

Table 1 presents descriptive information for the sample. The majority of clinicians worked predominately with adult clients and the main context in which memory assessments were conducted was with the Accident Compensation Corporation (ACC). Two clinicians who worked with adults did so within other contexts: criminal procedure, the Intellectual Disability sector, and a 50/50 split between the local District Health Board (DHB; since the completion of this survey, DHBs have been combined into a single health entity, Te Whatu Ora) and ACC. A further four clinicians working with both adults and children identified family court, the Intellectual Disability sector and Oranga Tamariki (child protection services) as other contexts for cognitive assessments. The majority of clinicians were registered clinical psychologists and/or neuropsychologists. In keeping with previously reported data (Skirrow & Ward, 2022), the majority of psychologists registered within the neuropsychology scope were also registered within the clinical scope. One

participant working with children was a registered educational psychologist and one participant working with both adults and children was registered under the 'general' (psychologist) scope of practice as well as counselling psychology and educational psychology scopes.

Perceived Validity for Use with NZ Adults

Figure 1 illustrates the perceived validity of working memory tests for use with adult clients. Clinicians perceived that the most valid tests were the WAIS-IV Digit Span and Letter-Number Sequencing tests, the Wechsler Memory Scale (WMS-IV; Wechsler, 2009) Spatial Addition and Symbol Span tests, and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998) Digit Span test. The Stanford Binet 5 (SB-5; Roid, 2003) working memory tests were rarely utilised, as very few clinicians provided validity ratings. In addition, the Kaplan Baycrest Neurocognitive Assessment (KBNA; Leach et al., 2000) Numbers test was rated "reasonably valid" by four clinicians.

For the auditory learning and retrieval domain (Figure 2), CVLT, RBANS Story Memory, and WMS Logical Memory tests were considered by a small number of clinicians ($n = 3, 6,$ and 1 , respectively) to have 'significant' validity issues. The RBANS List Learning and Rey Auditory Verbal Learning Test (RAVLT; Schmidt, 1996) were perceived as the most valid with the greatest percentage of clinicians selecting "completely valid" (30% and 26.6%, respectively). The Rivermead Behavioural Memory Test (RBMT; Wilson et al., 2008); First and Second Names, Stories, and Words) and Doors and People tests (all subtests; Baddeley et al., 1994) had the lowest number of clinicians provide validity ratings, with the rest of the sample determining they did not know the test well enough to comment.

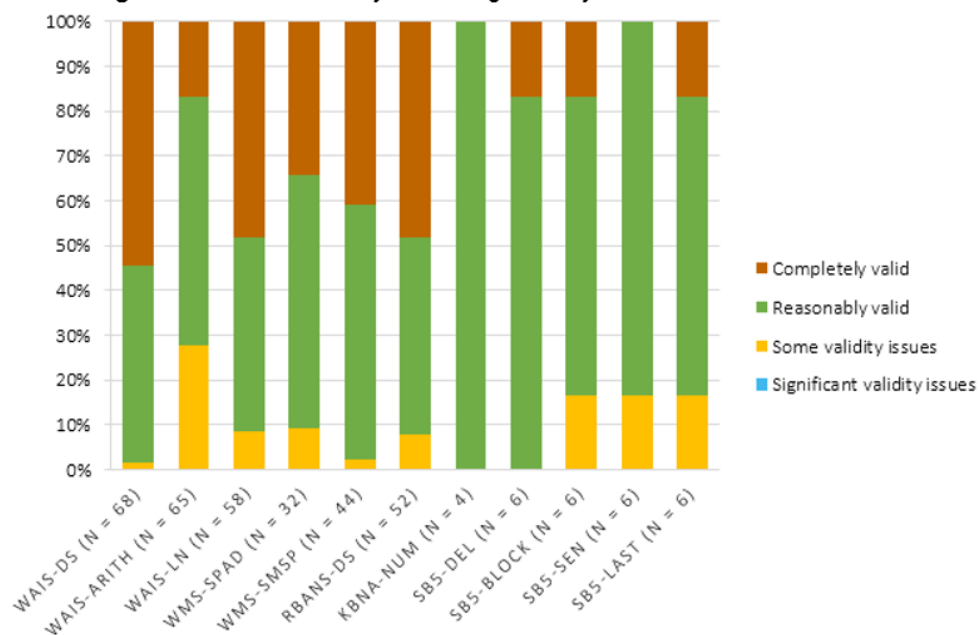
The tests perceived as most valid within the visual learning and retrieval domain were the Rey Complex Figure Test (RCFT; Meyers & Meyers, 1995) and modified Taylor Complex Figure Task (TCFT; Figure 3; Hubley, 1999). The WMS-IV Visual Reproduction 1 & 2, WMS-IV Designs 1 & 2, and RBANS Figure Recall tests were perceived to be 'reasonably' to 'completely' valid.

Within the prospective and autobiographical memory test domain, less than 16% of clinicians provided validity ratings, as the majority of participants did not know the test well enough to comment. From the ratings provided, the RBMT Appointments and Autobiographical Memory Interview (AMI; Kopelman et al., 1990) tests were perceived as 'reasonably' to 'completely' valid by all clinicians, however, a small number of clinicians ($n = 1$ and 2 , respectively) suggested that both the RBMT possessions and Cambridge Prospective Memory Test (CPMT; Wilson et al., 2005) had 'some' validity issues.

Perceived Validity for Use with NZ Children

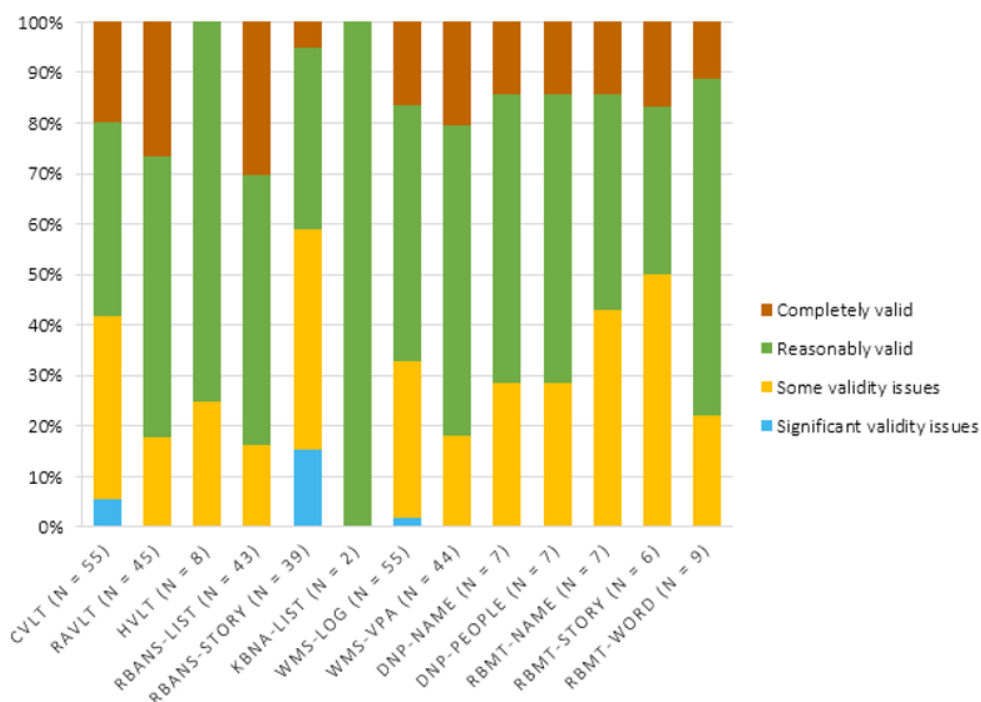
The perceived validity concerns of working memory tests for use with children is shown in Figure 4. Most clinicians perceived commonly used tests including the WISC-V Digit Span, Picture Span, and Letter/Number tests as reasonably valid, with only 4.8% ($n = 1$), 14.3% ($n = 3$) and 11.1% ($n = 2$) of clinicians respectively identifying "some validity issues". The NEPSY Sentence

Figure 1. Perceived validity of working memory tests for use with NZ adults.



Note: WAIS-IV Digit Span (WAIS-DS), WAIS-IV Arithmetic (WAIS-ARITH), WAIS-IV Letter-Number Sequencing (WAIS-LN), WMS-IV Spatial Addition (WAIS-SPAD), WMS-IV Symbol Span (WMS-SMSP), RBANS Digit Span (RBANS-DS), KBNA Numbers (KBNA-NUM), SB-5 Delayed Response (SB5-DEL), SB-5 Block Span (SB5-BLOCK), SB-5 Sentences (SB5-SEN), SB-5 Last Word (SB5-LAST).

Figure 2. Perceived validity of auditory learning and retrieval tests for use with NZ adults.



Note: California Verbal Learning Test (CVLT), Rey Auditory Verbal Learning Test (RAVLT), Hopkins Verbal Learning Test (HVLTL), RBANS Lists Learning 1&2 (RBANS-List), RBANS Story 1&2 (RBANS-STORY), KBNA List Learning (KBNA-LIST), WMS-IV Logical Memory (WMS-LOG), WMS-IV Verbal Paired Associates (WMS-VPA), Doors & People – Names (DNP-NAME), Doors & People – People (DNP-PEOPLE), RBMT Names (RBMT-NAMES), RBMT Stories (RBMT-STORY), RBMT Words (RBMT-WORD).

Repetition test was perceived by 8.1% (n = 1) of clinicians as having “significant validity issues”. Only one clinician provided a validity rating (“reasonably valid”) for the SB-5 Delayed, Block Span, Sentences and Last Word tests.

For the verbal learning and retrieval domain, seven out of nine tests had at least 10% of participants identify

cultural validity issues. The NEPSY Narratives and Children’s Memory Scale (CMS; Cohen, 1997) Stories tests had the greatest percentage of clinicians identify validity concerns (50% and 27.3%, respectively).

Within the visual learning and prospective memory domain, around one to four clinicians provided validity

ratings except for the RCFT test (n = 13). The NEPSY Designs test was perceived to be the most valid followed by the NEPSY Faces and CMS Dot Location tests.

Validity Themes Relating to Adult Memory Tests

Seventeen clinicians (25%) provided open-text responses around the validity of neuropsychological memory tests. Eight tests had no specific validity concerns raised by clinicians; the SB-5 tests, Hopkins Verbal Learning Test – Revised (HVLT-R; Brandt & Benedict, 1991), Doors and People, Recognition Memory Test (RMT; Warrington, 1984), Brief Visuospatial Memory Test – Revised (BVM-T-R; Benedict, 1996), Benton Visual Retention Test, CPMT, and AMI. The RBMT validity issues varied for each subtest. The first and second names test was viewed as Anglocentric and clinicians raised particular concerns with the ‘possession’ test – particularly relating to whether borrowing an item from the client was viewed as appropriate in all cultures. Specific concerns were raised about the appropriateness of several of the commonly used verbal and visual memory tests and are described in detail below.

Language Barriers

A recurring issue highlighted by clinicians for clients whose first language was not English (including Māori, Pasifika clients) was the presence of a language barrier. While using the Māori language for the alphabet and numbers was seen as a simple solution, there were concerns around the validity of substituting English and Māori terminology due to a lack of norms associated with

the use of Māori terms. Additionally, there were concerns around the variability in pronunciation of word lists by clinicians which was thought to potentially influence test scores.

Wechsler Adult Intelligence Scale 4th Edition:

“Can be challenging for people who have English as a second language as it assumes familiarity with English numerical system”

“...different (number of syllables) in English and other languages...”

California Verbal Learning Test 2nd Edition:

“...pronunciation in NZ makes it hard (i.e., cow sounds like car here)...”

Self-Made Adaptations

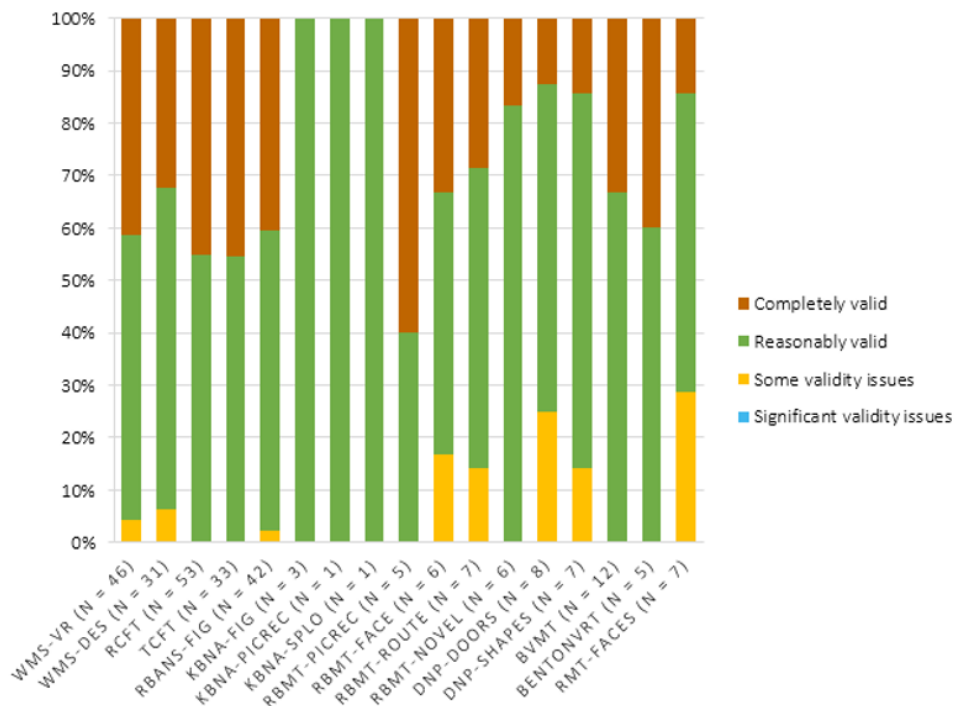
Clinicians frequently described contextual differences that required New Zealand adaptations. While some adaptations were commonly used, concerns were raised about whether the adaptations were beneficial or confused clients more. The use of self-made adaptations were perceived as necessary in some cases to ensure semantic meaning was retained across cultures, however, it was seen as difficult to achieve as some American words do not have a New Zealand equivalent.

Wechsler Memory Scale 4th Edition:

“Everybody comments on South Christchurch not being a place”

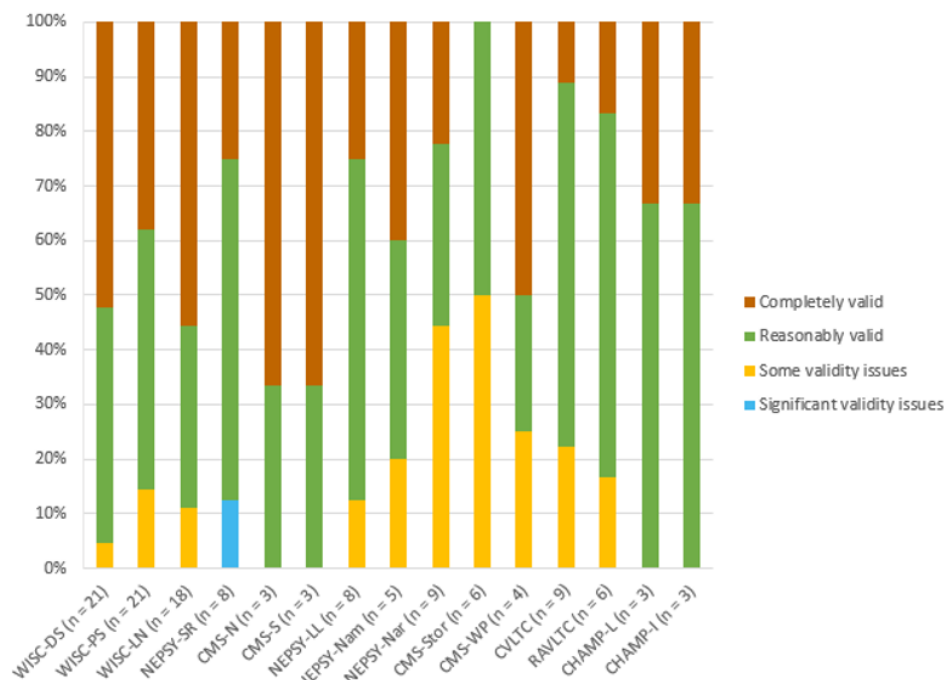
“Stories have cultural loading in terms of language and nature i.e., we don’t have cooks in school cafeterias”

Figure 3. Perceived validity of visual learning and retrieval tests for use with NZ adults.



Note: WMS-IV Visual Reproduction (WMS-VR), WMS-IV Designs 1&2 (WMS-DES), Rey Complex Figure Test (RCFT), Taylor Complex figure Test (TCFT), RBANS Figure 1&2 (RBANS-FIG), KBNA Complex Figure Test (KBNA-FIG), KBNA Picture Recognition (KBNA-PICREC), KBNA Spatial Location (KBNA-SPLO), RBMT Picture Recognition (RBMT-PICREC), RBMT Faces (RBMT-FACES), RBMT Route 1&2 (RBMT-ROUTE), RBMT Novel Task (RBMT-NOVEL), Doors & People – Doors (DNP-DOORS), Doors & People – Shapes (DNP-SHAPES), Brief Visuospatial Memory Test (BVMT), Benton Visual Retention Test Recognition Memory Test (BENTONVRT), RMT Faces (RMT-FACES).

Figure 4. Perceived validity of working memory, verbal learning, and retrieval tests for use with NZ children.



Note: WISC-V Digit Span (WISC-DS), WISC-V Picture Span (WISC-PS), WISC Letter-Number Sequencing (WISC-LN), NEPSY Sentence Repetition (NEPSY-SR), CMS Numbers (CMS-N), CMS Sentences (CMS-S), NEPSY List Learning (NEPSY-LL), NEPSY Names (NEPSY-NAM), NEPSY Narrative (NEPSY-NAR), CMS Stories (CMS-STOR), CMS Word Pairs (CMS-WP), California Verbal Learning Test – Children (CVLT-C), Rey Auditory Verbal Learning Test – Children (RAVLT-C), Child and Adolescent Memory Profile Lists (CHAMP-L), CHAMP Instructions (CHAMP-I).

Repeatable Battery for the Assessment of Neuropsychological Status

“ ‘Cleveland, Ohio’ feels culturally irrelevant to most NZers. I have tended to substitute with ‘Dunedin, Otago’...”

“More problematic is the ‘3 alarm fire’ which is just not a term NZers use and is not easy to change”

California Verbal Learning Test 2nd Edition

“Subway does not work at all for the NZ population. Most NZers equate subway with the fast food chain, which then impacts on the cued recall. Replacing with train increases face validity but also has issues with truck and train being phonetically similar”

Cross-Cultural Biases

Tests involving a base understanding of geometric shapes was seen to reflect a Westernised viewpoint of man-made structures which may not relate to other cultures and their way of living.

Complex Figure Tests

“...it should be noted that cross-culturally, even the concept of ‘normative’ shapes, imagery, etc. may not hold...this can affect, for example, a person’s ability to apply language to the image e.g., triangle, diamond, which assists with memory”

Validity Themes Relating to Child Memory Tests

Three clinicians (12%) provided open-text responses around the validity of neuropsychological memory tests. Clinicians did not highlight any specific validity concerns relating to the SB-5, Child and Adolescent Memory

Profile (CHAMP; Sherman & Brooks, 2015), NEPSY, or Children’s Auditory Verbal Learning Test/Rey Auditory Verbal Learning Test (C-AVLT/RAVLT; Talley, 1994). One clinician commented that the NEPSY Narrative test is “overly complex”. The remaining memory tests for children had specific concerns which are described below through themes and quotation examples.

Anglocentric

Clinicians stated tests were typically unable to account for cultural nuances, which may place Māori clients at a disadvantage compared with Western clients. Basic terminology differences between the US and NZ were identified by clinicians, with concerns these differences caused a reduction in significance and/or relevance each word had for NZ children. Self-made adaptations were identified in an attempt to use more NZ-based terminology that children would have semantic connections with.

Wechsler Intelligence Scale for Children

“Need to use numbers in Te Reo for some Māori children...”

California Verbal Learning Test – Children’s Version

“... ‘Sweater’ is not something most NZ children are particularly familiar with.”

“...I usually ask the adult prior to testing what the child calls a hoodie/sweater/jersey, as I notice many kids stumble over this one more than others.”

Influence of Education

Clinician reports suggested that childhood background was hugely influential and that perceived enjoyment around testing was affected by the role of NZ's education system in the pressure Māori children face when tests are not adapted to fit their cultural needs.

Rey Complex Figure Test

"...I do feel Māori kids can be put off by this when they first see it, it can look like something very 'Pākehā'...some of the kids from artist and carving whānau get it and love it, but others are intimidated"

Wechsler Intelligence Scale for Children

"I do think this test can intimidate Māori children a little more – some get overwhelmed and expect to fail, which I worry is connected for some to the expect(ation) to fail in the NZ education system..."

DISCUSSION

New Zealand has its own, unique culture and language patterns (including Māori and Pasifika influences), however, the majority of neuropsychological tests are imported from overseas. This study sought to evaluate clinicians' perceptions of the validity of many standardised tests of memory for use with a New Zealand population. Generally, clinicians perceived tests with low cultural/language loading (i.e., tests involving numbers or abstract figures), such as the WAIS Digit Span or WISC Picture Span tests, as the most valid across child and adult assessments. Furthermore, the child and adult assessment tests perceived to be the least valid were predominantly language-based involving word lists or stories, such as the NEPSY Narratives, CMS Stories, and RBANS Story Memory tests. Despite the perceived validity issues, most clinicians tended to perceive the issues as relatively minor, rating the neuropsychological memory tests overall as useful and valid for use in NZ.

Notably, the findings from Ogden et al. (2003) align with the clinician perspectives identified in the present study. Ogden et al. (2003) demonstrated that Māori received lower scores on vocabulary and logical memory tests (i.e., tests with a high language component) compared to NZ Europeans. It was also identified that Māori scores aligned with NZ Europeans for tests that contained a large visuospatial component and limited language component (e.g., Rey Complex Figure).

There were several tests included in the study which relatively few clinicians provided validity ratings for (notably, prospective and autobiographical tests). This may reflect a wider lack of use of these tests in an NZ context and demonstrate how psychologists in NZ tend to use tests they have previous experience of/training in and access to (Schaefer et al., 2021; Skirrow et al., 2023).

Aligning with Ross-McAlpine et al. (2018), the present study suggested that NZ clinicians attempt to overcome any perceived language/cultural barriers through self-made adaptations for child assessments (e.g., substituting "Cleveland, Ohio" with "Dunedin, Otago" in the RBANS). Expanding on the findings from Ross-McAlpine et al. (2018), the present study suggested that self-made adaptations are commonly used with both adults and children. While such changes are made by clinicians to increase the cultural relevance to clients, it

raises concerns over the standardisation and validity of the interpretations of test data (Howieson, 2019).

The present study adds to wider concerns that the WISC-V normative data for NZ (Wechsler, 2014b) lacks cultural validity for the diverse clientele in NZ, with clinicians typically describing the tests as Anglocentric. Notably, participants raised the need to use te reo Māori (Māori language) when working with Māori children. This finding complements previous points from Haitana et al. (2010) around the need to examine the impact of te reo Māori in neuropsychological assessments (i.e., whether test performance improves when using te reo Māori). While specific cultural issues were highlighted, clinicians typically perceived the WISC-V tests as useful/beneficial within an NZ context, which aligns with clinician perspectives from the Ross-McAlpine et al. (2018) study.

The presence of specific cultural issues highlighted in the present study implies neuropsychological memory assessment is functional, but not tailored to an NZ context (which has an increasingly diverse population). Therefore, aligning with Rivera Mindt et al. (2021), Schaefer et al. (2021), Rabin et al. (2016), and Tan et al. (2021), cultural factors should be accounted for in order to improve the reliability and validity of neuropsychological memory assessment. Many of the tests we have described have been extensively validated, with large sample norms, in overseas populations. Rather than abandoning the use of these tests, the development of large-sample NZ normative data would allow New Zealanders to receive more accurate assessment, diagnosis and treatment (Feigin & Barker-Collo, 2007).

A limitation of the present study was not gathering data on clinicians' culture and ethnicity. While there are typically few clinicians identifying as Māori – according to the NZ Psychologists Board (personal communication, 2023) only six currently registered neuropsychologists identify as having Māori heritage – it will be important to actively include Māori and Pasifika perspectives in future research when discussing the cultural validity of neuropsychological memory tests. Another limitation was the use of clinician perspectives which may be susceptible to bias (i.e., perceptions may over- or under-estimate validity issues) and the inferences drawn from the data could be limited. For example, Pākehā clinicians may not be as sensitive to cultural validity issues as non-European clinicians which may account for why memory tests were typically perceived as "reasonably valid".. It could be beneficial in the future to assess the validity of neuropsychological assessment from a Kaupapa Māori and Pasifika perspective, while also comparing NZ populations with normative data. Neuropsychology remains a relatively small discipline in New Zealand. Since 76 of 179 (42%) members of NZSIGN (at the time of the study) took part in this study, it is positive that our results may reflect the views of a significant proportion of clinicians practicing in NZ.

The present study has provided insight into some of the validity issues associated with neuropsychological memory tests. Concerningly, some of the tests considered to have the most significant issues for use in NZ appear to represent "standard practice" amongst psychologists undertaking neuropsychological assessment (e.g., WMS,

WAIS-IV, CVLT; Skirrow et al., 2023). In 2007, Feigin and Barker-Collo, described the need for NZ normative data as “urgent”, due to the poor validity of tests validated overseas. More than 15 years later, the need has not reduced, yet New Zealanders still have not seen a significant increase in access to locally-developed or normed neuropsychological tests.

Conclusions

While neuropsychological memory tests are generally perceived to be valid, the specific concerns raised around the tests being Anglocentric and lacking semantic meaning provide areas to improve neuropsychological memory assessment within an NZ context. Since

neuropsychological memory testing is important for the differential diagnosis and treatment of neurological conditions, it is vital the assessment measures are reflecting an individual's true neurological functioning; otherwise, individuals may be labelled as impaired/injured/dementing when they are not or clinicians may fail to detect an impairment in individuals who are experiencing cognitive issues. Furthermore, there is a significant need for validated NZ adaptations of the key tests (e.g., WAIS, WISC-V and WMS with NZ terminology) or the development of large-scale NZ norms.

References

- Aldrich, F. K., & Wilson, B. (1991). Rivermead Behavioural Memory Test for Children (RBMT-C): A preliminary evaluation. *British Journal of Clinical Psychology*, 30(2), 161-168. <https://doi.org/https://doi.org/10.1111/j.2044-8260.1991.tb00931.x>
- Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1994). Doors and People Test.
- Barker-Collo, S., Clarkson, A., Cribb, A., & Grogan, M. (2002). The impact of American content on California Verbal Learning Test performance: A New Zealand illustration. *The Clinical Neuropsychologist*, 16(3), 290-299. <https://doi.org/https://doi.org/10.1076/clin.16.3.290.13856>
- Barker-Collo, S. L. (2001). The 60-item Boston Naming Test: Cultural bias and possible adaptations for New Zealand. *Aphasiology*, 15(1), 85-92. <https://doi.org/https://doi.org/10.1080/02687040042000124>
- Benedict, R. H. B. (1996). *Brief Visuospatial Memory Test - Revised*. Psychological Assessment Resources.
- Brandt, J., & Benedict, R. H. B. (1991). *Hopkins Verbal Learning Test - Revised*. Psychological Assessment Resources.
- Braun V & Clarke V (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology* 3:2, 77-101. DOI: 10.1191/1478088706qp0630a
- Cohen, M. (1997). *Children's Memory Scale*. Pearson Australia.
- Dudley, M., Barker-Collo, S., Wilson, D., & Garrett, N. (2022). Age stratified normative data for Māori on the Wechsler Adult Intelligence Scale (4th edition; WAIS-IV). *New Zealand Journal of Psychology*, 51(2), 15-25. https://doi.org/https://www.psychology.org.nz/application/files/7316/6631/4270/Dudley_15-25.pdf
- Dudley, M., Faleafa, M., & Young, E. (2016). Cross-cultural neuropsychology in Aotearoa New Zealand. *Professional Practice of Psychology in Aotearoa New Zealand*, 247-264.
- Dudley, M., Scott, K., & Barker-Collo, S. (2017). Is the test of premorbid functioning a valid measure for Māori in New Zealand? *New Zealand Journal of Psychology*, 46(3), 72-79. <https://doi.org/https://www.psychology.org.nz/journal-archive/TOPF-in-NZ-Maori-private.pdf>
- Dudley, M. D., Barker-Collo, S. L., Wilson, D. L., & Garrett, N. K. (2019). Factors associated with Māori performance on the WAIS-IV. *Archives of Clinical Neuropsychology*, 34(7), 1203-1216. <https://doi.org/https://doi.org/10.1093/arclin/acy110>
- Evans, J. J. (2010). *Handbook of Clinical Neuropsychology*. Oxford Academic. <https://doi.org/https://doi.org/10.1093/acprof:oso/9780199234110.003.02>
- Feigin, V. L., & Barker-Collo, S. (2007). Normative neuropsychological data: Do we need them in New Zealand? *New Zealand Family Physician*, 34(6), 441.
- Fernández, A. L., & Abe, J. (2018). Bias in cross-cultural neuropsychological testing: Problems and possible solutions. *Culture and Brain*, 6, 1-35. <https://doi.org/10.1007/s40167-017-0050-2>
- Haitana, T., Pitama, S., & Rucklidge, J. J. (2010). Cultural biases in the peabody picture vocabulary test-III: Testing in tamariki in a New Zealand sample. *New Zealand Journal of Psychology*, 39(3), 24-34.
- Howieson, D. (2019). Current limitations of neuropsychological tests and assessment procedures. *The Clinical Neuropsychologist*, 33(2), 200-208. <https://doi.org/https://doi.org/10.1080/13854046.2018.1552762>
- Hubley, A. M. (1999). The modified Taylor Complex Figure: A comparable measure to the Rey-Osterrieth Figure. *Archives of Clinical Neuropsychology*, 14(8), 734. <https://doi.org/https://doi.org/10.1093/arclin/14.8.734>
- Jansen-van Vuuren, J., & Aldersey, H. M. (2020). Stigma, acceptance and belonging for people with IDD across cultures. *Current Developmental Disorders Reports*, 7, 163-172. <https://doi.org/https://doi.org/10.1007/s40474-020-00206-w>
- Kopelman, M., Wilson, B. A., & Baddeley, A. (1990). *Autobiographical Memory Interview*. Pearson Australia.
- Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY - Second Edition (NEPSY-II)*. APA PsycTests. <https://doi.org/https://doi.org/10.1037/t15125-000>
- Leach, L., Kaplan, E. F., Rewilak, D., Richards, B., & Proulx, G. (2000). *Kaplan-Baycrest Neurocognitive Assessment*. Melbourne, Pearson Australia.
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). Neuropsychological assessment: A compendium of tests and assessment techniques. In *Neuropsychological assessment*.
- Mashal, N. (2013). The role of working memory in the comprehension of unfamiliar and familiar metaphors. *Language and Cognition*, 5(4), 409-436. <https://doi.org/10.1515/langcog-2013-0024>
- Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial: Professional Manual*. Psychological Assessment Resources, Odessa.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). *Montreal Cognitive Assessment (MoCA)*. APA

- PsycTests.
<https://doi.org/https://psycnet.apa.org/doi/10.1037/t27279-000>
- Ogden, J. A., Cooper, E., & Dudley, M. (2003). Adapting neuropsychological assessments for minority groups: A study comparing white and Maori New Zealanders. *Brain Impairment*, 4(2), 122-134.
- Quesque, F., Coutrot, A., Cox, S., de Souza, L. C., Baez, S., Cardona, J. F., . . . Bertoux, M. (2022). Does culture shape our understanding of others' thoughts and emotions? An investigation across 12 countries. *Neuropsychology*, 37(7), 664-682.
<https://doi.org/https://doi.org/10.1037/neu0000817>
- Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in Test-Usage Practices of Clinical Neuropsychologists in the United States and Canada Over a 10-Year Period: A Follow-Up Survey of INS and NAN Members. *Archives of Clinical Neuropsychology*, 31(3), 206-230.
<https://doi.org/10.1093/arclin/acw007>
- Randolph, C. (1998). *Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)*. Melbourne, Pearson Australia.
- Rivera Mindt, M., Marquine, M. J., Aghvian, M., Paredes, A. M., Kamalyan, L., Suárez, P., . . . Cherner, M. (2021). The Neuropsychological Norms for the U.S.-Mexico Border Region in Spanish (NP-NUMBRS) Project: Overview and considerations for life span research and evidence-based practice [Article]. *Clinical Neuropsychologist*, 35(2), 466-480.
<https://doi.org/10.1080/13854046.2020.1794046>
- Roid, G. H. (2003). *Stanford-Binet Intelligence Scales (5th Edition)*. Riverside Publishing.
- Ross-McAlpine, K. S., Leathem, J. M., & Flett, R. A. (2018). A survey of psychologists administering cognitive and neuropsychological assessments with New Zealand children. *New Zealand Journal of Psychology*, 47(1), 13-19.
- Schaefer, L. A., Thakur, T., & Meager, M. R. (2021). Neuropsychological Assessment. *StatPearls*.
<https://www.ncbi.nlm.nih.gov/books/NBK513310/>
- Schmidt, M. (1996). *Rey Auditory Verbal Learning Test: A handbook*. Western Psychological Services.
- Sherman, E. M. S., & Brooks, B. L. (2015). *Child and Adolescent Memory Profile*. Psychological Assessment Resources.
- Skirrow, P.M., & Ward, N. (2022). What's in a Scope? Where Neuropsychology Meets Clinical Psychology. *Journal of the New Zealand College of Clinical Psychologists*, 32 (1), 3-7.
- Skirrow, P. M., Johnstone, G., Faulkner, J. W., & Douglas, K. M. (2023). Measuring memory: A survey of neuropsychological practice amongst New Zealand psychologists. *Applied Neuropsychology: Adult*, 1-8.
<https://doi.org/10.1080/23279095.2023.2251635>
- Steyn, N., Binny, R. N., Hannah, K., Hendy, S. C., James, A., Lustig, A., . . . Sporle, A. (2021). Māori and Pacific People in New Zealand have higher risk of hospitalisation for COVID-19. *medRxiv*, 2020.2012.2025.20248427.
<https://doi.org/10.1101/2020.12.25.20248427>
- Statistics New Zealand. (2018). *Census Data*.
<https://www.stats.govt.nz/tools/2018-census-place-summaries/new-zealand#ethnicity-culture-and-identity>
- Talamaivao, N., Harris, R., Cormack, D., Paine, S.-J., & King, P. (2020). Racism and health in Aotearoa New Zealand: a systematic review of quantitative studies. *The New Zealand Medical Journal (Online)*, 133(1521), 55.
<https://www.proquest.com/scholarly-journals/racism-health-aotearoa-new-zealand-systematic/docview/2442318063/se-2?accountid=14700>
- Talley, J. L. (1994). *Child Auditory Verbal Learning Test (2nd Edition)*. Psychological Assessment Resources.
- Tan, Y. W., Burgess, G. H., & Green, R. J. (2021). The effects of acculturation on neuropsychological test performance: A systematic literature review [Review]. *Clinical Neuropsychologist*, 35(3), 541-571.
<https://doi.org/10.1080/13854046.2020.1714740>
- Warrington, E. K. (1984). *Recognition Memory Test*. WPS Ltd.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale - Fourth Edition (WAIS-IV)*. APA PsycTests.
<https://doi.org/https://doi.org/10.1037/t15169-000>
- Wechsler, D. (2009). *Wechsler Memory Scale - Fourth Edition (WMS-IV)*. APA PsycTests.
<https://doi.org/https://psycnet.apa.org/doi/10.1037/t15175-000>
- Wechsler, D. (2014a). *The Wechsler Intelligence Scale for Children (5th ed.)*. Bloomington: The Psychological Corporation.
- Wechsler, D. (2014b). *WICS-V: Technical and interpretive manual*. Pearson.
- Weintraub, S., Bauer, P. J., Zelazo, P. D., Wallner-Allen, K., Dikmen, S. S., Heaton, R. K., . . . Gershon, R. C. (2013). I. NIH Toolbox Cognition Battery (CB): Introduction and pediatric data. *Monographs of the Society for Research in Child Development*, 78(4), 1-15.
<https://doi.org/https://doi.org/10.1111/mono.12031>
- Wilson, B. A., Evans, J. J., Emslie, H., Foley, J., Shiel, A., Watson, P., . . . Groot, Y. (2005). *The Cambridge Prospective Memory Test*. Pearson Clinical Assessments UK.
- Wilson, B. A., Greenfield, E., Clare, L., Baddeley, A., Cockburn, J., Watson, P., . . . Nannery, R. (2008). *Rivermead Behavioural Memory Test (3rd Edition)*. Pearson Australia.

Corresponding author:

Paul Skirrow
 Department of Psychological Medicine, University of Otago, Wellington, New Zealand.
 Email: paul.skirrow@otago.ac.nz

Funding Information

Grace Johnstone was funded by a Neurological Foundation Summer Studentship.

Acknowledgements

We would like to thank the Neurological Foundation whose summer research scholarship funded Grace Johnstone. We would also like to thank the clinicians from NZSIGN, NZCCP, and NZPsS who participated in this survey. KMD was supported by a Sir Charles Hercus Health Research Fellowship from the Health Research Council of New Zealand (ref: 19/082) during completion of this paper.