A survey of psychologists administering cognitive and neuropsychological assessments with New Zealand children

Kate S. Ross-McAlpine, Janet M. Leathem and Ross A. Flett
Massey University, Wellington, New Zealand

Psychometric measures form an essential component of cognitive and neuropsychological assessments, yet there is a lack of published research about the practices of psychologists administering these tests in New Zealand. An online survey was developed to determine the views and practices of psychologists who administer cognitive and neuropsychological assessments with New Zealand children. Results of the 66 respondents revealed that the WISC-IV and the ABAS-II-Parent were the most frequently used measures and that lack of access and familiarity were most frequently cited reasons for not using a measure. Most respondents had concerns regarding the cultural sensitivity of tests and thought that New Zealand normative data was needed.

Keywords: neuropsychological assessment; cognitive assessment; test selection; child; New Zealand context

Cognitive and neuropsychological assessments with children must be accurate and valid due to the significant influence that results can have on the life of a child and their family/whanau. Patterns of psychometric test use and how appropriate a test is for the children being assessed, are critical considerations to ensure assessments are accurate and valid.

The selection of a psychometric measures involves multiple considerations, such as whether to use a flexible approach by selecting subtests as needed (e.g., Delis Kaplan Executive Function Scale), a fixed test battery approach (e.g., Luria-Nebraska Battery) or an approach based on cognitive abilities (Koziol & Budding, 2011). Other considerations include psychometric properties (validity and reliability/ sensitivity and specificity), availability of parallel forms and time, cost and whether to use computerised or traditional administration (Lezak et al., 2012).

A New Zealand perspective on test use was gained by a survey of members of the New Zealand Council for Educational Research (psychologists, consultants and counsellors) (Dunn & Dugdale, 2002) who considered whether test selection should be based on relevance, validity and credibility and on assessor factors such as training, experience and competence, not on cost-effectiveness and employer policies. The most five commonly used measures, (endorsed as being used once a month or more) were the Beck Depression Inventory (BDI-II; 27%), Wechsler Adult Intelligence Scale (WAIS-III; 22%), Wechsler Intelligence Scale for Children (WISC-III; 16%), Ravens Standard Progressive Matrices (SPM; 15%) and the Myers-Briggs Type Indicator (MBTI; 14%) (Dunn & Dugdale, 2002). Further, recommendations in the second edition of the Professional Practise of Psychology in Aotearoa New Zealand were that test selection should be based on content, psychometric qualities and based on the level the instrument is aimed at (Eatwell & Wilson, 2007). Despite the psychometric emphasis in cognitive and neuropsychological assessments, there is a lack of literature on the patterns of test use in these specific assessments.

The overall assumption is that the measures selected will be appropriate to the individual being assessed. This is achieved by comparing the individual being assessed to a normative sample (Feigin & Barker-Collo, 2007) usually comprised of local people in the geographical area where the tests were developed (Lezak et al., 2012). Typically these areas are North America (Feigin & Barker-Collo, 2007) described more recently as "white US and Canadian populations" (Thames, Karimian, & Steiner, 2016, p. 140) which differs significantly in geographical and cultural context from where the measures were often used. As there can be substantial differences between normative groups and the population being assessed, it is important to be aware of cultural differences when completing neuropsychological assessments (Horton, 2008; Wong, 2006). Literature in the New Zealand context is consistent with this (Dudley, Faleafa, & Yong, 2016) with accurate New Zealand normative data for neuropsychological assessments deemed necessary to provide increased specificity and sensitivity of diagnosis (Dudley et al., 2016). While some normative data for New Zealand has been developed (see the Psychological and Neuropsychological Norms for New Zealand data base https://cdn.auckland.ac.nz/assets/psych/about/our-research/documents/psychological-and-neuropsychological-norms-for-new-zealand.pdf), the only published article providing normative data with New Zealand children appeared 14 years ago (Fernando, Chard, Butcher, & McKay, 2003). While the second edition of the Professional Practice of Psychology in Aotearoa New Zealand (Ogden, 2007), lists measures commonly used for neuropsychological assessment with children the list does not appear to be substantiated by empirical research.

Cross-cultural neuropsychology has identified cultural differences with psychometric tools used in the neuropsychological of cognitive assessment of school-aged children (Sobeh & Spijkers, 2013; Mulenga, Ahonen & Aro, 2001). In New Zealand, cultural bias has been investigated by Haitana, Pitama, & Rucklidge, 2010 who found that the

Peabody Picture Vocabulary Test-III (PRVT-111) was largely appropriate for use with Māori children in mainstream schools but not for children attending Māori-medium schools where results were more indicative of stage of English language development than their overall language ability. These authors suggested adaptations such as more culturally appropriate target words (e.g., changing Porcupine to Hedgehog) and recommended inclusion of te reo Māori.

Despite the literature identifying cultural differences, it is currently unclear if psychologists who regularly administer these cognitive and neuropsychological assessments are aware of cultural influences or if they make adaptations for use with New Zealand children. There is no empirical research investigating psychologists' opinions on cultural influences.

In summary, the practice of cognitive and neuropsychological assessments with children in New Zealand has received little research attention. The purpose of the current study was to gether information about the cognitive testing processes with New Zealand children through a survey of New Zealand psychologists. There were two aims 1) to determine the patterns of test use, specifically the frequency of test and subtest use, as well as the reasons for test selection and 2) to provide perspectives on the impact and influences of cultural in cognitive and neuropsychological assessments with children.

Method

Survey

In consultation with three psychologists who administer cognitive measures with children, an online survey was developed that would take approximately 20 minutes to complete. It consisted of three sections a) clinical practice (country of training, scope of practice, current work sector; years and frequency of experience), b) test selection and c) cultural considerations.

The test selection section, b, focused on the frequency of use and reasons explaining use for measures. For clarity, measures were classified as either comprehensive measures or domain specific measures. Comprehensive measures include the WISC-IV[11] , NEPSY-II, Child Memory Scale (CMS), Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV), Delis Kaplan Executive Function System (DKEFS) and Stanford-Binet Intelligence Scale (SB5) and respondents were also asked about subtest use for these measures. Domain specific measures included for example the Stroop task and Conners Continuous Performance Test (CPT-3), as well as informant scales such as Adaptive Behavioural Assessment System (ABAS-II). Frequency of test use when completeing a cognitive or neuropsychologeical assessment was determined on a 5-point Likert scale based on use when completeing a cognitive or neuropsychologeical assessment (always use, almost always use, sometimes use, almost never use and never use). Respondents were able to endorse reasons against using a test from a given list as well as provide additional reasons explaining test use in an open comment box. The survey also asked about use of computerised administration and scoring and measures of effort. Suggested measures had been compiled from resources such as; Professional Practice of Psychology in Aotearoa New Zealand (Eatwell & Wilson, 2016), ACC Neuropsychological Assessment Services (ACC, 2009) and recent literature on neuropsychological assessment (e.g., Lezak et al., 2012).

The third part of the survey, contained open-ended questions on cultural appropriateness and need for New Zealand normative data, as well asking what cultural adaptations respondents made to administration.

Procedure

Invitation to participate in the survey was distributed via a link in the June 2016 New Zealand Psychological Society (NZPsS) Connections magazine and through an email sent to the members of the New Zealand College of Clinical Psychologists (NZCCP), the New Zealand Special Interest Group in Neuropsychology (NZSIGN) and the Massey University Psychology Clinics. Participants were also recruited through word of mouth within the psychological community. The survey was accessible online from June to August 2016. Due to this manner of recruitment there is no way to accurately determine the representation of the sample.

Data Analysis

The survey was analysed using SPSS 24. Although 97 psychologists started the survey only 66 (68%) responses could be analysed due to varying degrees of incompleteness. Quantitative results were analysed using descriptive statistics while the brief open-ended qualitative comments were analysed in terms of frequencies of most common survey responses similar to previous literature (e.g., Barker-Collo, 2015; Brooks et al., 2016).

Results

Sample

As shown in Table 1 almost all respondents were trained in New Zealand, with most working as clinical, education and general psychologists, in the education sector, district health boards or in private practice. Most (72.6%) had qualified within the previous nine years, and 63.6% conducted assessments 1-3 weekly.

Table 1
Demographic information of participants

		n	Percentage of respondents
Country of Training		65	
New Zealand		58	89.2
UK		3	4.6
USA		2	3
Australia		1	1.5
Other		1	1.5
Scope of Practice		66	
Clinical		28	42.4
General		21	31.8
Educational		14	21.2
Clinical Intern		3	4.5
Current Work*		66	
Education Sector		27	40.9
DHB		21	31.8
Private		18	27.3
ACC Practitioner		11	16.7
University		5	7.6
Community/NGO		4	6.1
Other		3	4.5
Experience with child cognitive and neuropsychological assessment		66	
Years	Less than 1 year	16	24.2
	1 to 4	16	24.2
	5 to 9	16	24.2
	10 to 14	9	13.6
	15 to 19	5	7.6
	20+ years	4	6.1
Frequency	Weekly	27	40.9
	2-3 Week	15	22.7
	Monthly	8	12.1
	2-6 monthly	7	10.6
	6-12 month	6	6
	Once in 12 months	3	3

Note. * = Percentage values for current area of work do not total to 100% because many respondents endorsed more than one current area of work.

1 Due to the recent updated of the 5th WISC edition, if respondents had started using the WISC-V they were encouraged to respondent based on their use of WISC-IV.

Test use

Comprehensive measures

As shown in Table 2, the WISC-IV was by far the most frequently used of the six measures with a very small percentage using the SB5.

Table 2
Current use of comprehensive measures

Measure	Total *n*	*n*	Percentage
WISC-IV	66	62	93.9
NEPSY-II	66	20	30.3
WPPSI-IV	64	18	28.1
DKEFS	66	15	22.7
CMS	65	13	20.0
SB5	65	6	9.2

Note. Total *n* = number of respondents who answered the question.
n = number of respondents who endorsed using the measure

Reasons the comprehensive measures were not used are shown in Table 3 below. An 'other' option provided respondents the opportunity to give alternative reasons. Comments listed within the 'other' option, that were endorsed by more than 10%[2] (CMS, WPPSI and SB), are described by notes [a], [b] and [c].

Table 3
Reasons endorsed for not using the specific comprehensive measures

Reason	Percentage* (*n*)					
	WISC (*3*)	NEPSY (*44*)	CMS (*51*)	WPPSI (*44*)	SB (*58*)	DKEFS (*50*)
Lack of familiarity with test		34.8	50.0	16.7	48.5	47.0
Lack of access to test materials		25.8	37.9	19.7	53.0	42.4
Limited/No training on this test		19.7	18.2	4.5	21.2	25.8
Purpose of the assessment	3.0	15.2	15.2	16.7	6.1	7.6
Cost of test		4.5	1.5	1.5	4.5	4.5
Length of administration time	1.5	10.6		1.5	1.5	
Psychometric properties		4.5		1.5	3.0	
Culturally inappropriate				1.5	1.5	
Reputation					3.0	
Other	3.0	9.1	10.6[a]	34.8[b]	10.6[c]	3.0

Note. * = Percentage values do not total to 100% because respondents endorsed more than one reason.
[a] = Within this 10.6%, most (42.8%) stated they preferred the NEPSY-II and CVLT
[b] = Within this 34.8% most (82.6%) stated they do not work with this age group of children.
[c] = Within this 10.6% most (57%) stated they preferred the WISC

As shown on Table 3, lack of familiarity was the most endorsed reason for participants to not use the NEPSY-II, CMS and DKEFS. Lack of access to test materials was the most endorsed reason for not using the SB5 in current practice. Most did not use the WPPSI due to the age of client (82.6%) and there were preferences for other measures over the CMS and SB5.

Respondents who endorsed using the comprehensive measures were invited to comment on the reasons explaining use[3] . Use of the NEPSY-II depended on the referral question (28%) and use of the WPPSI depending on the interest/ compliance of the child (21%) with 14% stating that it was used in conjunction with observation and school performance. The DKEFS was said to be a good/excellent measure by 18% with 12% stating that it was good to use with adolescents. The SB5 was good to use with low functioning children (18%).

Subtest Use

Respondents indicated the frequency of subtest use (detailed in the appendix) and commented on the reasons for use as detailed below. For the WISC subtests four participants noted that Letter-Number Sequencing was confusing or frustrating and two that Comprehension was helpful in providing information on thinking in social situations and that they preferred starting an assessment with Picture Completion.

NEPSY-II comments included "*Narrative Memory has a story that is far too complex*" and "*Memory for Names is too culture specific*", "*I like the Auditory Attention Test for which there is no other equivalent test*" and use of subtests for specific peoples "*I have used the Affect Recognition and Theory of Mind subtests as a way to assess children with autistic spectrum traits*" and "*I have used Theory of Mind, Memory for Names and Faces for students with a diagnosis of autism*".

The only comment specific to subtest use for the CMS was "*I like the stories in the CMS better than the stores in the NEPSY-II*". No comments were made about the DKEFS or WPPSI-IV subtests and of the SB5 "*some parts are too culturally bound e.g., picture absurdities with the map of the Americas*" and "*block span can be distracting to children*".

Domain-specific measures

The frequency of use of domain-specific or rating scales is shown on Table 4. Again respondents made specific comments on the reasons for use of the measures, and as with the comprehensive measures lack of familiarity and lack of access were the main reasons for not using measures. Test use was also influenced by referral reason or specific assessment type (e.g., for a ADHD or ID assessment). Preference for another measure also explained test use, with the ABAS[4] measures preferred over the three Vineland measures and the DKEFS or NEPSY-11 preferred over the Stroop-Child version.

Other measures. Respondents were asked to list other measures that they used in their practice that were not on the list provided. Most frequently mentioned were the Strengths and Difficulties Questionnaire (*n* = 7), Resiliency scales (*n* = 5), Autism Diagnostic Observation Schedule (ADOS[5]) (*n* = 4) and Beck Youth Inventories (*n* = 2). Measures that were only mentioned once are provided in the footnote below[6].

Effort

Most respondents (72.6%) did not administer a measure to assess effort (n = 62).

2 This value is guided by principals of a content analysis

3 As previously mentioned, reasons are reported based on frequency being 10% or greater

4 Second and third edition included

5 Second and third edition included

6 In alphabetical order; Autism Diagnostic Interview (ADI-R), Age and Stage Questionnaire (ASQ), Achenbach System of Empirically Based Assessment (ASEBA), Battelle Developmental Inventory, Childhood Autism rating Scale (CARS), Comprehensive Executive Function Inventory (CEFI), Children's Depression Inventory (CDI), Cognitive Assessment System, Eyberg Child Behaviour Inventory, Gilliam Autism Scale, Integrated Visual and Auditory Continuous Performance (IVA), Multidimensional Anxiety Scale for Children (MASC), Kaufman-Brief Intelligence Test-2, Naglieri Nonverbal ability test, Parenting scales, Piers-Harris self-concept scale, Performance Validity Tests. Raven Matrices, Session rating scale, Spence anxiety scale, Social competency scale, Outcome Rating Scale, Test of problem solving, Test of word reading efficacy, Wechsler Abbreviated Scale of Intelligence (WASI-II), Wechsler Nonverbal Scale of Ability (WNV), Wide Range Assessment of memory and Learning (WRAML) and Word Memory Test (WMT).

Those who did most frequently used the Test of Memory Malingering (TOMM) (26.1%), followed by child observation or interview (21.7%), Word Memory test (17%), Rey 16 item test (13%) and embedded measures (8%).

Table 4
Most frequently used domain specific or rating scale measures

Measure		*n*	Percentage of responses				
			Always	*Almost always*	*Sometimes*	*Almost never*	*Never*
Informant Scales							
ABAS II	Parent	**63**	11.1	12.7	57.1	9.5	9.5
	Teacher	**61**	9.8	9.8	55.7	14.8	9.8
CCBRS	Parent	**62**	9.7	11.3	37.1	14.5	27.4
	Teacher	**61**	9.8	11.5	37.7	14.8	26.2
	Self	**60**	5	8.3	36.7	13.3	36.7
CBCL	Parent	**64**	9.4	25	39.1	10.9	15.6
BRIEF	Parent	**62**	4.8	9.7	17.7	6.5	61.3
	Teacher	**61**	4.9	6.6	16.4	6.6	65.6
BASC	Parent	**63**	1.6	12.7	23.8	14.3	47.6
	Teacher	**61**	1.6	13.1	23	11.5	50
	SDH	**62**	1.6	0	6.5	21	71
	Self	**62**	0	12.9	16.1	9.7	61.3
Vineland-II	Teacher	**63**	1.6	3.2	19	12.7	63.5
	Interview	**63**	1.6	3.2	15.9	11.1	68.3
	Parent	**61**	1.6	1.6	21.3	14.8	60.7
Other Domains Specific Measures							
CVLT-C		**60**	5	0	8.3	3.3	83.3
TEA-Ch		**62**	3.2	1.6	3.2	3.2	88.7
PPVT-4		**60**	3.3	5	28.3	13.3	50
WIAT-3		**63**	3.2	4.8	28.6	12.7	50.8
Conners CPS-3		**60**	3.3	3.3	16.7	11.7	65
WRAT-4		**61**	1.6	1.6	6.6	14.8	75.4
RBANS		**62**	0	1.6	6.5	8.1	83.9
C-AVLT-2		**61**	0	3.3	3.3	3.3	90.2
Stroop - Child		**62**	0	1.6	4.8	1.6	91.9
Bayley Scales		**63**	0	1.6	3.2	12.7	82.5
CELF-4		**60**	0	0	5	5	90
Bender Gestalt Test		**62**	0	0	1.6	3.2	95.2
AWMA		**62**	0	0	1.6	1.6	96.8

Note. ABAS- Parent = Adaptive Behavioural Assessment Scale - Parent, CBCL – Parent = Child Behaviour Checklist – Parent, ABAS II - Teacher = Adaptive Behavioural Assessment Scale – Teacher, CCBRS-Parent = Conner's Comprehensive Behaviour Rating Scales - Parent, CCBRS-Teacher = Conner's Comprehensive Behaviour Rating Scales - Teacher, BRIEF – Parent = Behaviour Rating Inventory for Executive Functioning - Parent, CCBRS - Self = Conner's Comprehensive Behaviour Rating Scales - Self, BRIEF – Teacher = Behaviour Rating Inventory for Executive Functioning - Teacher, CVLT-C
= California Verbal Learning Test – Children's Version, PPVT-4 = Peabody Picture Vocabulary Test – 4th, WIAT-3 = Wechsler Individual Achievement Test – 3rd Edition, Conners CPS-3 = Conners Continuous Performance Scale3, TEA-Ch = Test of Everyday Attention – Children, BASC - Parent
= Behaviour Assessment Scale for Children – Parent, BASC - Teacher = Behaviour Assessment Scale for Children – Teacher, Vineland-II – Teacher = Vineland Adaptive Behaviour Scale –II – Teacher, Vineland-II – Survey = Vineland Adaptive Behaviour Scale –Survey Interview Form, Vineland-II - Parent = Vineland Adaptive Behaviour Scale –Parent, WRAT-4 = Wide range achievement test-4, BASC-SDH = Behaviour Assessment Scale for Children – Structured Developmental History, BASC-S = Behaviour Assessment Scale for Children – Self, RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, C-AVLT-2 = Children's Auditory Verbal Learning Test-2, Bayley Scales = Bayley Scales of Infant and Toddler Development, Stroop = Stroop Colour and Word Test: Children's Version,CELF-4 = Clinical Evaluation of Language Fundamentals -4th, AWMA = Automated Working Memory Assessment.

Computerised administration and scoring

Most respondents did not use computerised administration in their assessments (71.7%) and of those that did the majority used the Connor's CPT (*n* = 5; 9.4%). Comparatively, most respondents used computerised scoring (65%), most commonly for the WISC-IV & V (*n* = 36; 60 %), but also for NEPSY-11 (*n* = 7; 11.6%).

Cultural Considerations

Assessment measures

Respondents were asked to comment on the culturally sensitivity of measures, with 24% expressing some concern as shown in Table 5 (in order of increasing concern). Fourteen respondents mentioned specific tests; 10 indicated the benefit of New Zealand norms in WISC-V and WPPSI-IV and four people had concerns, 2 with the ABAS and 2 with the vocabulary, information or comprehension subtests from the WISC-IV.

Table 5
Respondents comments on cultural appropriateness *(n = 62)*

Themes	Representative Comments
Adequate (16%)	"*no problem*"
	"*generally okay*"
	"*passable*"
Could be better (4.5%)	*"they could be a lot more culturally appropriate"*
	"could be improved"
	"I don't think they fit perfectly to NZ culture"
Improving (9%)	*"getting better"*
	"WISC V has improved cultural relevance"
	"WISC V has relevant norms for NZ/Australia so appears appropriate"
Clinical judgement required to acknowledge culture (11%)	"*cultural background must always be considered*"
	"All of them have biases that need to be taken into consideration when clinically interpreted"
Not culturally appropriate (9%)	*"not very"*
	"not entirely culturally appropriate"
Concern for specific persons and peoples (24%)	*"no norms for Māori or Pasifika children which I do not think is appropriate"*
	"*not great with new immigrants or refugees*"
	"*need to question use with specific populations … including English as second language and Māori"*

Note: All of the themes identified are presented in this table (even ones occurring < 10%) in order to demonstrate the richness and range of opinions evident in the data from this question

Administration changes

Most respondents (69.7%) stated that they changed aspects of administration for use with New Zealand children examples of which are shown on Table 6. Other changes included units ("*pounds/miles*") and places ("New York and Chicago to Auckland to Wellington"). One participant commented they do so "*where it makes understanding clearer for client*" and another that they are "*mindful of my own accent*".

Table 6
Examples of words changed by respondents *(n = 62)* assessing NZ children

Original Word	Changed for NZ Children
Mom	Mum
Purse	Wallet
Squash	Pumpkin
President	Prime Minister
Fall	Autumn
Eraser	Rubber
Juan	Tama
Thongs	Jandals
Trash	Rubbish
Family	Whānau
Faucet	Tap
Store	Shop

Normative Data

Of the 62 respondents who offered an opinion regarding the need for normative data for New Zealand, 80.3% said 'yes', 9.1% were 'not sure' and 4.5% of respondents said 'no'. The most frequent reason for the need of normative data was due to the uniqueness of New Zealand (22%), with comments including, "*I think it is important to have norms for Maori children in particular given the effects of colonization*", "*…the cultural make-up of New Zealand is completely different than any overseas nation*", "*gathering normative data is incredibly important to ensure that we are aware of the differences and unique characteristics of our child population*" and "*there's no excuse really, we need to be able to norm to an New Zealand population so we can use these tests with more validity and reliability*".

Of the respondents who were unsure or stated no, 10 made comments, the majority identifying reasons against collecting norms including "*cost*", "*labour intensive and frankly a waste of time*" and "*likely to reflect Auckland and not the rest of New Zealand*". Additionally, two participants thought New Zealand norms would be ideal but not necessary "*not 100% necessary… but optimal would be to have New Zealand norms*". One respondent commented that collecting New Zealand norms would be an "*incessant luxury*".

Final opportunity for comments included the following insights "*Clinical judgement and experience need to be used with all information gathered from tests and rating scales*" and "*…culturally appropriate assessment needs to focus less on the specific instruments and more on the assessors using them*" and "I*n my experience there is an over reliance on formal tests and not enough information gathered from all the environments a child spends time in*". Two respondents provided more argument for the need for cultural appropriateness measures "*we need more culturally appropriate tests for our indigenous population*" and "*it would be great to have some more locally produced and affordable products that pertain specifically to a New Zealand population*".

Post Hoc Findings

Examination of the relationship between place of work and scope of practice (see Table 7) revealed that most respondents in the educational sector were equally divided between the education and general scopes of practice, while those in the clinical scope had a greater spread across work settings, working privately and in District Health Boards.

Chi-squared tests of independence were used to determine whether test use for the comprehensive measures was equal across scope of psychological practice, years of experience and frequency of assessments.

Table 7
Scope of practice and current work comparison

Scope	Current work (*n**)						
	Education	DHB	Private	ACC	University	Community/ NGO	Other
Clinical	0	15	14	9	4	2	2
Education	14	0	1	0	1	0	0
General	13	4	2	2	0	1	1
Intern	0	2	1	0	0	1	0

Note. * = *n* does not total the number of respondents in each scope as respondents were able to endorse multiple current work settings

Scope of practice

Statistically significant relationships were found between the use of the NEPSY-II (Statistically significant relationships were found between the use of the NEPSY-II ($\chi^2(3) = 13.297$, $p = .004$), DKEFS ($\chi^2(3) = 17.125$, $p = .001$), CMS (χ^2 (3) = 11.518, $p = .009$) and SB5 ($\chi^2(3) = 8.024$, $p = .046$) and scope of practice with participants in the clinical scope using these measures NEPSY-II, DKEFS, CSM and SB5 more (see Figure 1). No statistically significant association was found between scope of practice and use of the WISC-IV (χ^2 (3) = 4.106, p = .250) or the WPPSI-IV (χ^2 (3) = 1.820, $p = .611$).

Frequency of assessments

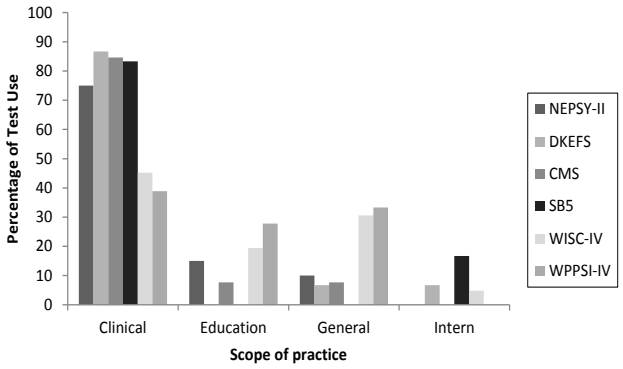

Figure 1. Scope of practice and use of NEPSY-II, DKEFS, CMS, SB5, WISC-IV, WPPSI-IV

A statistically significant relationship was found between the frequency of administering assessments and use of the NEPSY-II ($\chi^2(5) = 13.151$, $p = .022$) and CMS ($\chi^2(5) = 12.709$, $p = .026$), with the NEPSY-II and CMS used more by respondents who administer assessments more frequently than those who administer assessments less frequently. No statistically significant association was found between frequency of administering assessments and use of the WISC-IV ($\chi^2(5) = 5.370$, $p = .372$), DKEFS ($\chi^2(5) = 6.610$, p = .251), WPPSI ($\chi^2(5) = 8.315$, $p = .140$) or SB5 ($\chi^2(5) = 1.538$, p = .909).

Years of experience

A statistically significant relationship between years of experience and use of the CMS ($\chi^2(5) = 11.797$, p = .038), was found with respondents with more years experience using this measure more frequently. No statistically significant association between the years of experience and use of the WISC-IV ($\chi^2(5) = 8.037$, $p = .154$), NEPSY-II ($\chi^2(5) = 2.909$, p = .714), DKEFS ($\chi^2(5) = 6.141$, $p = .293$), WPPSI ($\chi^2(5) = 5.071$, p = .407) or SB5 ($\chi^2(5) = 8.458$, $p = .133$).

Discussion

This study sought to explore the current practices of psychologists conducting cognitive and neuropsychological assessment with children in New Zealand, in particular to determine the frequency of test and subtest use, provide perspectives on test selection and cultural considerations.

Consistent with previous research (Dunn & Dugdale, 2002) the WISC-IV was the most commonly used comprehensive measure to assess cognitive and neuropsychological function of New Zealand children. The most commonly used domain specific/rating scales, the ABAS, CBCL and CCBRS, reflect the emphasis on adaptive and difficult behaviour in the context of cognitive and neuropsychological assessments. The focus on adaptive behaviour is not surprising as it is required for a DSM diagnosis of Intellectual Disability.

The focus on behaviour scales in the context of cognitive and neuropsychological assessments reflects the importance placed on considering a child's behaviour. Additionally, the importance the respondents placed on gathering information from parents, and to lesser extent teachers, was also demonstrated. Consideration of a child's behaviour and obtaining perspectives from third parties in the context of neuropsychological assessments is consistent with the literature (New Zealand Psychologists Board 2013; Teeter et al., 2009).

Ideally, psychologists should be familiar with and have access to a range of measures to provide the best assessment for a child. This was explained by Darby & Walsh (2005) who said "*the experienced neuropsychologist will gradually develop quite a large armamentarium from which to choose*" a test measure (p 404). Yet the results of this survey indicate that test selection appears dominated by pragmatic considerations such as lack of familiarity and access. Opportunities for psychologists to familiarise themselves with a range of measures and increased access to psychometric materials would result in test selection being less dominated by pragmatic considerations and aim towards an ideal assessment. Shifting test selection from pragmatism to idealism would allow for the most appropriate test being used to assess a child.

In literature overseas, effort is seen as a necessary component in comprehensive neuropsychological assessment with children to ensure assessments are a reflection of true ability (Deright & Carone, 2015; Perna, 2016). It is crucial that cognitive and neuropsychological assessments are an accurate representation of ability in order to avoid incorrect interpretations or diagnoses. However, majority of our respondents did not use a measure of effort in their assessments with children. This finding might be due to only 16.7% of respondents working for ACC where assessment of effort is expected. Of those who did use a measure of effort, the most common measure (TOMM, 26.1%) was a stand-alone measure of effort, compared with embedded measures (8.7%). This is somewhat similar to research with adults in New Zealand (Barker-Collo & Fernando, 2015) where the TOMM was used slightly more frequently (39.7%) than embedded measures (38%). However it is inconsistent with previous research in North America (Brooks et al., 2016) where embedded measures were used more frequently in neuropsychological assessments with children than stand-alone measures.

Concerns raised regarding cultural sensitivity in the current study align with recent discussions on the need to recognise the diversity within New Zealand (Dudley et al., 2016). Respondents' concerns demonstrate an awareness and sensitivity to culture within their psychological practice e.g., changing '*family*' to '*whānau*' during assessment administration. This finding reinforces comments made by Haitana et al. (2010) suggesting that future research should investigate the use of te reo Māori in cognitive assessments with children in New Zealand.

The focus on the diversity of New Zealand was also reflected with the finding that the majority of the survey respondents considered it important to obtain normative data for New Zealand children (80.3%). These professional opinions are consistent with the literature (Dudley et al., 2016). Normative data for New Zealand has been collected for the WISC-V with a sample of 528 children from New Zealand and Australia. This recently updated measure can therefore be used in New Zealand with greater confidence than other measures. Since the WISC-IV is frequently used during a cognitive or neuropsychological assessment with New Zealand children, providing more accurate cognitive and neuropsychological assessments in New Zealand in the future is likely.

The post hoc findings showed that clinical psychologists were more likely to ustilise a range of psychometrics than other scopes of practice. This may be due to the varying core competencies for different scopes of practice as defined by the New Zealand Psychologists Board. A skill described as 'completion of cognitive intellectual assessments and neuropsychological screening' (pp. 21) is only detailed within the clinical scope of practice (New Zealand Psychologists Board, 2011).

It is recognised that the size and self-selecting nature of the sample may have resulted in sample biases and this is identified as a limitation of the current study. The findings have highlighted test and subtest preferences and provided insights into why particular measures are being selected. The majority of respondents had a sensitivity to and consideration for diversity and culture which was demonstrated through comments made about whether measures are culturally appropriate, concerns for specific peoples and persons, adaptations made for use in New Zealand and the need for normative data. Future research in the area of neuropsychological assessments in New Zealand should develop normative data for measures which are frequently used in order to ensure neuropsychological assessments are the most appropriate for every child in Aotearoa, New Zealand.

References

ACC (2009). Appendix 1: Psychometric tests of Service Agreement Schedule for Neuropsychological Assessment Services (pp. 19)

Barker-Collo, S., & Fernando, K. (2015). A survey of New Zealand psychologists ' practices with respect to the assessment of performance validity, *44*(2), 35–42.

Board, N. Z. P. (2011). *Core Competencies For the Practice of Psychology in New Zealand*.

Brooks, B. L., Ploetz, D. M., Kirkwood, M. W., Brooks, B. L., Ploetz, D. M., A, M. W. K., … Kirkwood, M. W. (2016). A survey of neuropsychologists ' use of validity tests with children and adolescents. *Child Neuropsychology, 22*(8), 1001–1020. https://doi.org/10.1080/09297049.2015.1075491

Code, T., Board, P., & Zealand, N. (2013). *Guidelines on the use of Psychometric Tests February 2013*.

Deright, J., & Carone, D. A. (2015). Assessment of effort in children : A systematic review. *Child Neuropsychology, 21*(1), 1–24. https://doi.org/10.1080/09297049.2013.864383

Dickinson, M. D., & Hiscock, M. (2011). The Flynn Effect in Neuropsychological Assessment. *Applied Neuropsychology, 18*(2), 136–142. https://doi.org/10.1080/09084282.2010.547785

Dudley, M., Faleafa, M., & Yong, E. (2016). Cross-cultural neuropsychology in Aotearoa New Zealand. In *Professional Practice of Psychology in Aotearoa New Zealand* (pp. 247–264).

Dunn, K., & Dugdale, J. (2002). Patterns of Use of Psychological Tests in New Zealand. In *Inaugural International Test Users' Conference* (pp. 1–10).

Eatwell, J., & Wilson, I. (2007). The effective use of psychometric assessments in decision making. In *Professional Practice of Psychology in Aotearoa New Zealand* (pp. 317–333).

Feigin, V. L., & Barker-Collo, S. (2007). Normative neuropsychological data : Do we need them in New Zealand ? *New Zealand Family Physician, 34*(6), 441–443.

Fernando, K., Chard, L., Butcher, M., & McKay, C. (2003). Standardisation of the Rey Complex Figure Test in New Zealand children and adolescents. *New Zealand Journal of Psychology, 32*(1), 33–38.

Haitana, T., Pitama, S., & Rucklidge, J. (2010). Cultural biases in the Peabody Picture Vocabulary Test-III : Testing Tamariki in a New Zealand Sample, *39*(3).

Horton, A. (2008). Multicultural Neuropsychological Assessment: The future of neuropsychology. In *The Neuropsychological Handbook* (pp. 345–366).

Koziol, L., & Budding, D. (2011). Pediatric neuropsychological testing: Theoretrical models of test selection and interpretation. In A. Davis (Ed.), *Handbook of pediatric neuropsychology* (pp. 443–455). New York: Springer.

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

Mulenga, K., Ahonen, T., & Aro, M. (2001). Performance of Zambian Children on the NEPSY: A pilot study. Developmental *Neuropsychology, 20*(1), 375–383.

Ogden, J. (2007). Practising clinical neuropsychology in Aotearoa New Zealand. In I. Evans, J. Rucklidge, & M. O'Driscoll (Eds.), *Professional Practice of Psychology in Aotearoa New Zealand* (pp. 351–366).

Perna, R. (2016). Effort Testing In Child Cognitive and Neuropsychological Testing Effort Testing In Child Cognitive and Neuropsychological Testing. *Journal of Pediatrics and Neonatal Care, 5*(May), 2–3. https://doi.org/10.15406/jpnc.2016.05.00194

Sobeh, J., & Spijkers, W. (2013). Development of neuropsychological functions of attention in two cultures: A cross-cultural study of attentional performances of Syrian and German children of pre-school and schoola ge. *European Journal of Developmental Psychology, 10*(3), 318–336. https://doi.org/10.1080/17405629.2012.674761

Teeter, P., Eckert, L., Nelson, A., Platten, P., Semrud-Clikeman, M., & Kamphaus, R. (2009). Assessment of behaviour and personality in the neuropsychological diagnosis of children. In C. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of Clinical Child Neuropsychology* (3rd ed., pp. 349–382). Springer.

Thames, A., Karimian, A., & Steiner, A. (2016). Neuropsychological assessment of ethnic minority children. In S. Graves & J. Blake (Eds.), *Psychoeducational assessment and intervention for ethnic minority children: Evident-based approaches* (pp. 133–161). American Psychological Association.

Wong, T. M. (2006). Ethical controversies in neuropsychological test selection, administration, and interpretation. *Applied Neuropsychology, 13*(2), 68–76. https://doi.org/10.1207/s15324826an1302_2

Corresponding Author

Kate Ross-McAlpine
School of Psychology
Massey University
Wellington
New Zealand
Email: kateross.mcalpine@gmail.com

APPENDIX

Frequency of WISC-IV, WPPSI-IV, NEPSY-II, DKEFS and CMS subtest use

Subtests	*n*	Percentage of responses				
		Always	*Almost always*	*Sometimes*	*Almost never*	*Never*
WISC-IV						
10 main subtests	**59**	79.7	15.3	3.4	0	1.7
Full 15 subtests	**51**	2	9.8	31.4	29.4	27.5
Digit Span	**59**	88.1	10.2	1.7	0	0
Coding	**60**	86.7	13.3	0	0	0
Block Design	**59**	86.4	13.6	0	0	0
Similarities	**59**	86.4	11.9	0	0	1.7
Matrix Reasoning	**58**	86.2	12.1	1.7	0	0
Symbol Search	**60**	85	15	0	0	0
Vocabulary	**59**	83.1	13.6	3.4	0	0
Letter-no. sequencing	**59**	57.6	15.3	20.3	6.8	0
Comprehension	**59**	55.9	23.7	11.9	8.5	0
Information	**59**	50.8	13.6	16.9	11.9	6.8
Picture Completion	**54**	25.9	11.1	31.5	20.4	11.1
Arithmetic	**58**	24.1	17.2	22.4	31	5.2
Cancellation	**54**	14.8	5.6	35.2	27.8	16.7
WPPSI-IV						
All	**13**	38.5	30.8	15.4	7.7	7.7
Matrix Reasoning	**15**	66.7	13.3	20	0	0
Block Design	**15**	60	26.7	13.3	0	0
Similarities	**15**	60	20	20	0	0
Information	**13**	53.8	30.8	15.4	0	0
Coding	**15**	53.3	20	20	0	6.7
Picture Concepts	**14**	50	7.1	35.7	0	7.1
Vocabulary	**15**	46.7	26.7	20	0	6.7
Picture Memory	**13**	46.2	15.4	30.8	0	7.7
Comprehension	**15**	40	20	33.3	0	6.7
Receptive Vocabulary	**13**	38.5	23.1	30.8	0	7.7
Zoo Location	**13**	38.5	15.4	23.1	15.4	7.7
Bug Search	**13**	38.5	23.1	23.1	7.7	7.7
Cancellation	**13**	38.5	15.4	15.4	23.1	7.7
Object Assembly	**14**	35.7	28.6	21.4	0	14.3
Picture Completion	**13**	30.8	15.4	30.8	7.7	15.4
Animal Coding	**13**	23.1	15.4	38.5	7.7	15.4
Picture Naming	**14**	21.4	28.6	21.4	14.3	14.3
Word Reasoning	**14**	21.4	7.1	50	7.1	14.3
NEPSY-II						
All	**16**	0	0	12.5	18.8	68.8
Inhibition	**18**	16.7	27.8	38.9	16.7	0

List Memory	**18**	16.7	33.3	38.3	11.1	0
Narrative Memory	**18**	11.1	33.3	44.4	11.1	0
Comprehension of Instructions	**18**	11.1	27.8	50	11.1	0
Animal Sorting	**18**	11.1	16.7	38.9	27.8	5.6
Theory of Mind	**18**	11.1	16.7	61.1	5.6	5.6
Affect recognition	18	11.1	5.6	66.7	5.6	11.1
Auditory Attention and Response set	**18**	5.6	27.8	33.3	22.2	11.1
Design Copying	**18**	5.6	16.7	61.1	5.6	11.1
Memory for Faces	**18**	5.6	11.1	55.6	22.2	5.6
Speeded naming	**18**	5.6	11.1	50	27.8	5.6
Word Generation	**18**	5.6	11.1	50	33.3	0
Geometric Puzzles	**17**	5.9	0	35.3	35.3	23.5
Memory for Designs	**18**	0	22.2	50	11.1	16.7
Clocks	**18**	0	11.1	33.3	33.3	22.2
Memory for Names	**18**	0	11.1	44.4	33.3	11.1
Arrows	**16**	0	18.8	43.8	31.3	6.3
Block construction	**18**	0	11.1	44.4	22.2	22.2
Repetition of Nonsense words	**17**	0	5.9	35.3	41.2	17.6
Sentence Repetition	**18**	0	0	66.7	22.2	11.1
Phonological Processing	**18**	0	0	61.1	16.7	22.2
Picture Puzzles	**18**	0	0	44.4	33.3	22.2
Route Finding	**17**	0	0	41.2	41.2	17.6
Oromotor Sequences	**18**	0	0	33.3	38.9	27.8
Statue	**17**	0	0	29.4	52.9	17.6
Fingertip Tapping	**18**	0	0	27.8	44.4	27.8
Imitating hand positions	**18**	0	0	22.2	55.6	22.2
Manual Motor Sequences	**18**	0	0	16.7	61.1	22.2
DKEFS						
All	**10**	20	10	10	10	50
Trail Making Test	**15**	40	20	40	0	0
Verbal Fluency	**15**	40	20	40	0	0
Colour-word interference	**15**	40	13.3	33.3	13.3	0
Design Fluency	**14**	14.3	14.3	35.7	21.4	14.3
Sorting	**15**	13.3	13.3	33.3	26.7	13.3
Twenty Questions	**15**	6.7	13.3	40	33.3	6.7
Tower	**15**	6.7	13.3	33.3	40	6.7
Proverb	**15**	0	6.7	20	46.7	26.7
Word Context	**14**	0	0	28.6	42.9	28.6
CMS						
All	**7**	0	14.3	57.1	0	28.6
Core CMS subtests	**10**	10	20	60	0	10
Stories Recall	**12**	41.7	16.7	41.7	0	0
Word pairs	**12**	16.7	16.7	50	16.7	0
Word Lists	**12**	16.7	16.7	50	16.7	0
Family Pictures	**12**	16.7	0	41.7	41.7	0
Dot locations	**11**	9.1	18.2	54.5	18.2	0
Picture Locations	**11**	9.1	0	54.5	36.4	0

Faces	**12**	8.3	16.7	66.7	8.3	0
Sequences	**12**	8.3	8.3	58.3	25	0
Numbers	**10**	0	10	50	30	10

Note. SB-V is not presented in this table, as so few (9.2%) respondents used it. Three respondents used all subtests *always*, 1 *almost always*, 1 *sometimes* and 1 *never*. Five subtests were identified as being used *always* by 2 respondents and *sometimes* by 1 respondent, Early reasoning, verbal analogies, procedural knowledge, form board, form patterns and memory for sentences.