Cultural biases in the Peabody Picture Vocabulary Test-III: Testing Tamariki in a New Zealand Sample

Tracy Haitana, *Department of Corrections, Christchurch*

Suzanne Pitama, *University of Otago*

Julia J. Rucklidge, *University of Canterbury*

The Peabody Picture Vocabulary Test (PPVT-III) is a test which is currently used in New Zealand both clinically and in research to measure receptive vocabulary skills (Phillips, McNaughton, & MacDonald, 2004; Reese & Read, 2000; Stockman, 2000). This research investigated issues of cultural bias by comparing PPVT-III scores obtained by 46 Māori children from three different age groups (5-11 years) with scores from the standardisation sample. Results revealed that the PPVT-III appeared to be suitable for use with Māori as a receptive vocabulary measure, although a number of suggestions were made as to ways in which the administration and interpretation of PPVT-III test scores could be adjusted when working with Māori in order to minimise the impact of cultural bias. Additional research is required to establish whether changes to potentially culturally biased items may improve the validity of the PPVT-III for use with Māori.

Issues of cultural bias in psychometric assessment have been widely debated in the research literature, with a range of views arising as to the ways in which cultural differences might compromise the validity of psychometric test results (Ogden, 2003; Ogden & McFarlane-Nathan, 1997; Shepherd & Leathem, 1999). Despite the continued debate about the effects of cultural bias, most professionals acknowledge that test content and administration procedures are invariably culturally bound, with some tests having a higher cultural loading and potential for bias than others. Test developers acknowledge the need to consider the impact of test content, test materials and test conditions on the reliability and validity of a test in an attempt to minimise the effects of cultural bias. In an unbiased measure, high or low scores reflect attributes of the person being assessed and cannot be seen as artefacts arising from culturally different understandings of the test material (Stockman, 2000). Within a New Zealand context, there is very little research that has investigated the appropriateness of using psychometric tools with Māori and more specifically, Māori tamariki.

The Peabody Picture Vocabulary Test (PPVT) is one psychometric tool that was designed and standardised for use with children in the United States, but is used in New Zealand with Māori children both clinically and in research (Phillips et al., 2004; Reese & Read, 2000). A previous version, the PPVT-R was found to produce different results for White American and African American kindergarten aged children (Rock, 2005). The results showed that African American first grade children tested with the PPVT-R obtained scores which suggested that they had approximately half the vocabulary of White American first grade children. The discrepancy between African American and White American children's scores on the PPVT-R remained close to one standard deviation apart, even after controlling for factors such as socio-economic status (SES), parental education, and low birth weight (Rock, 2005).

The PPVT-III, which according to test developers was standardised on a representative U.S. sample, is purported to be a culturally valid test due to the inclusion of a large representative group of ethnic minority children within the wider norm group (Dunn & Dunn, 1997; Stockman, 2000). In gauging the suitability of the PPVT-III for use with an at-risk sample of African American children, Washington and Craig (1999) found that their scores on the PPVT-III did not significantly differ from the norm group. As a result of this finding, they concluded that the PPVT-III represents a valid and culturally fair test suitable for use with African American children. However, on review of the available literature, it was apparent that limited research has been conducted in order to systematically investigate the use of the PPVT-III with people from diverse cultural backgrounds. Despite this seemingly well-intentioned move to reduce any previously existing cultural bias, some researchers argue that the mere inclusion of an ethnic minority norm group does not render a test unbiased. In order for a test to be considered unbiased, it must also measure culturally appropriate knowledge and utilise methods of testing suitable for people from different cultural backgrounds (Palmer, 2004).

In New Zealand a number of researchers have made use of the PPVT-III as a measure of verbal knowledge and receptive vocabulary development in children (Phillips et al., 2004; Reese & Cox, 1999; Reese & Read, 2000).

Although not standardised for use in New Zealand, the PPVT-III was chosen as a research tool because of its widespread usage in similar studies and its position as a well known test of emergent language (Phillips et al., 2004). Reese and Read (2004) measured the reliability of the PPVT-III for use amongst a mixed ethnicity sample of New Zealand children and found that the test-retest reliability of the PPVT-III closely matched that reported by the American norm reference group. This finding led the researchers to conclude that the PPVT-III was a useful instrument for measuring language ability in children from multiple ethnic and economic backgrounds (Reese & Read, 2000).

In another similar study, the PPVT-III was used as a baseline and post-intervention measure of receptive language skills with a sample that included a large proportion of Māori children (Phillips et al., 2004). Again this study assumed that the PPVT-III provided a valid measure of oral vocabulary skills in Māori children, attributing the difference found between pre and post intervention scores on the PPVT-III to the success of the intervention (Phillips et al., 2004). Given the accepted use of the PPVT-III in New Zealand, it seems important to assess whether it is a suitable test to be used with Māori children. Of note, the PPVT-III has now been superseded by a later edition, the PPVT-4, which may have further addressed any issues related to bias. The current paper was produced from research findings obtained during the preparation of a Masters thesis at which time the PPVT-III was the latest available edition.

The purpose of this study was to investigate the appropriateness of using the PPVT-III by administering the test with three age groups of Māori children (5-11 years). The key objectives included: 1) ascertaining the level of comfort and engagement experienced by Māori children when tested with the PPVT-III, 2) identifying which particular PPVT-III test items Māori children found easy or difficult and providing some explanation as to why, and 3) identifying whether Māori children were able to respond correctly to items presented after the final item administered during the standardised delivery on the PPVT-III.

Method

Participants

A total of 46 Māori children (24 boys and 22 girls) aged between 5 years 0 months and 10 years 9 months participated in this study. Participants comprised three different age groups, with the youngest group of children aged between 5–6 years 9 months (Mean = 5.98 years, *SD* = 0.65), the middle group aged between 7–8 years 9 months (Mean = 7.88 years, *SD* = 0.49), and the oldest group of children aged between 9–10 years 9 months (Mean = 10.11 years, *SD* = 0.59). Participants lived and attended schools in the Canterbury region. Of the 46 participants who took part in the study, 33 attended mainstream schools (72%), and 13 (28%) attended either Bilingual units within mainstream schools or Kura Kaupapa Māori.

In accordance with Ministry of Health guidelines, suitable participants were those who were identified by their parents/caregivers (using the 2001 census question) as Māori, were aged between 5 and 10 years 9 months, had no uncorrected vision or hearing problems, were described by their parents/caregivers as being competent in the English language, and were meeting key developmental milestones as reported by their parents/caregivers.

Participants were recruited over an eight month period, from 7/03/2006 to 7/11/2006. Given the specific characteristics required of the research sample, two main recruitment strategies were used to identify potential participants for this study. Participants were identified using community networks (through consultation hui and using existing Māori community networks), and snowball recruitment methods where additional participants were identified by current participants.

In order to address the generalisability of the data collected, a quota sampling approach was used, whereby participants were matched to the wider population of Māori children in New Zealand by the strata of gender, school type, and school decile rating (Clay, Ellis, Amodeo, Fassler, & Griffin, 2003). While all strata provided an indication of the similarities between the research sample and the wider population from which the sample was drawn, the school decile rating strata also provided a measure of socio-economic deprivation.

Materials and Apparatus

Peabody Picture Vocabulary Test – III

Participants were assessed using the third edition of the Peabody Picture Vocabulary Test (PPVT-III) Form IIIA, a psychometric test that was developed and standardised on a representative U.S. sample in 1997. The test was designed to measure receptive vocabulary in the English language, and is commonly used to assess a person's verbal ability, or as part of a wider battery of tests measuring cognitive functioning . A particular strength of the PPVT-III, which utilises pictures and spoken words rather than print, is the wide age-range (2½ to 90+ years) with which it can be applied. In order to complete the test, examinees are asked to choose which one of four pictures best represents a given word that is read to them by the examiner. The PPVT-III is reported to be a reliable and valid measure that shows consistency between forms and administrations of the test, and a strong relationship with associated measures of language. Test-retest reliabilities range from .91 to .94, convergent validity ranges from .80 to .92 when assessing the extent to which PPVT-III scores correlated with other measures of verbal ability, and discriminant validity estimates range from .62 to .85 when comparing PPVT-III scores with measures of different traits.

Procedure

Kaupapa Māori Research Approach

This study utilised a Kaupapa Māori Research (KMR) approach, which is a methodological framework that has been developed in response to a research history where Māori were researched 'on' rather than 'with'. KMR is not in itself a method, rather it is used to guide the research process in a way that prioritises, values, and protects Māori knowledge and ways of being. KMR dictates critiquing all tools utilised within the study and justifying their ability to validate Māori experiences. It also encourages clear accountability

between researcher, participant and community stakeholders. The KMR principles adhered to in this study were in fitting with the approach adopted by the Māori/Indigenous Health Institute (MIHI) in various research projects, and included consultation and input from key Māori stakeholders in the community at each stage of the research process.

This study was reviewed and approved by the Human Ethics Committee at the University of Canterbury. Potential participants identified by community networks and existing participants were provided with an information sheet. The researcher then contacted each potential participant's whanau by phone or face to face to discuss their willingness to take part in the project. At the first meeting, the study information was reviewed and a consent form was signed by willing parents/caregivers. A screening measure was then administered by the researcher to ensure that participants met the inclusion criteria required to take part in the study. Of those participants approached to take part in the research, one child did not meet the inclusion criteria, four were unable to be contacted by the researcher, and one chose not to participate. Due to the wide range of community networks used to distribute information pertaining to the research, it is impossible to establish how many suitable participants chose to passively decline.

Parents/caregivers were then asked to complete a short developmental survey which covered questions about pregnancy, birth and early developmental milestones so as to provide an indicator of possible childhood adversity that might have impacted negatively on developmental outcomes of participants. With the parent/caregiver present, the researcher met with the child, and explained the purpose and process of research participation. Children were provided with an age appropriate information sheet, and after reading/listening to the form, were asked to sign an assent form if they agreed to take part in the research.

Testing was conducted in a quiet place; 41 participants chose to undertake testing in their own homes, while five elected to be tested in a work environment. Most children preferred to carry out testing whilst sitting on the floor with the examiner, often with their parents or a caregiver nearby. Efforts were made to ensure that each child experienced similar test conditions, with emphasis placed on reducing test anxiety while providing a quiet test environment that was largely free from distractions.

Testing was carried out in a standardised manner in accordance with the PPVT-III examiners manual. Following the completion of standardised administration children were given a short break where they were invited to look through the remaining picture items to see if there were any pictures that they were able to recognise. During the break, the examiner took note of all incorrect responses on a blank response sheet, recording the item numbers, and correct picture numbers onto the sheet. The examiner then revisited incorrect items that had been administered, and asked participants to name the word that best described the picture presented to them in an attempt to identify the reasons for participant errors and allow investigation of possible item bias.

The final phase of testing involved asking participants some qualitative questions related to their perceptions of the PPVT-III which was recorded using a dictaphone. At the end of this process, children were thanked and provided with a koha by way of a $10 gift voucher and some stickers.

Administration of PPVT-III with Māori Children

Although a standardised approach to test administration was used as a guide throughout the study, a Māori process was used by the researcher in order to engage Māori children in the research. Prior to testing, time was spent connecting with participants and their whanau through establishing whakapapa and community links, and discussing the process and purpose of the research. The researcher was known to be Māori by all research participants, which in some cases facilitated the use of te reo Māori in response to PPVT-III test items. Of note, after engaging in this process all participants and whanau agreed to take part in the research, and all eligible child participants willingly completed standardised administration of the PPVT-III and other research tasks.

Consultation and Feedback to Key Māori Community Stakeholders

In line with a Kaupapa Māori methodological approach, meetings were arranged with key Māori community contacts and professionals currently working with children in the Canterbury region in order to discuss the ensuing research. The proposed research and methodology was presented to each group for comment as part of the consultation process. Subsequently further meetings were held to present the research findings and interpretations of the data were discussed and feedback was sought.

Results and Analysis

Excel and SPSS version 15 were used to manage and analyze the data. Planned ANOVA comparisons were conducted to compare the standard scores of participants from each of the three age groups, and the standard scores of participants attending different decile schools. T-tests were conducted in order to compare standard scores obtained on the PPVT-III by participants, with the mean score of 100. When comparing the standard scores obtained by children from different school types, participants were sorted into two groups: children attending mainstream classes in mainstream schools were defined as the 'Mainstream' group, and children attending either Kura Kaupapa or Bilingual units within mainstream schools were defined as the 'Māori-medium' group. Post-hoc comparisons were made in cases where significant differences were noted. Effect sizes (ES) were calculated according to Cohen's d where 0.2 indicates a small effect, 0.5 indicates a medium effect and 0.8 is indicative of a large effect size.

Analyses were also conducted to compare the percentage, rate, and number of errors made on individual PPVT-III items. Qualitative data was recorded for each participant on a dictaphone, and later transcribed. Initial analysis involved reading the transcripts, and coding each transcript into general themes.

Results

Demographic Information

Information outlining the composition of the sample and sub-samples is included in Table 1. Analysis of the developmental survey and screening measures revealed the diverse backgrounds and experiences of participants and their whanau who completed the research. Responses were obtained from 44 out of 46 participants, with two choosing not to return completed surveys. Participants closely matched the wider population on most developmental measures obtained. Of participants who took part in the research, 36 (78%) provided information regarding their iwi affiliations. A total of 25 iwi groups were represented, with some children identifying with as many as four iwi. Twenty-four participants identified with more than one ethnic group, with some children identifying up to three or four ethnic groups.

Mainstream and Māori-Medium School Sample Characteristics

Twenty-eight percent of the research sample were attending Kura Kaupapa or Bilingual schools at the time of the study. To allow for comparisons between these groups of children, analyses were conducted in order to ensure that mainstream and Māori-medium school participants represented similar groups of children in terms of sample characteristics. While participants were similar in terms of age measures, the distribution of participants by gender and school deciles differed between children of each school type. As research highlights, children attending Māori-medium schools tend to represent a more diverse group than mainstream school children in terms of their level of experience in both te ao Māori and te ao Pakeha. Therefore, any differences observed between children from mainstream and Māori-medium schools should be interpreted with caution as the influence of confounding cultural factors and differing sample characteristics cannot be ruled out.

A t-test was conducted to compare the standard scores obtained by this sample with the standardisation sample mean of 100. Analysis revealed a significant difference between the mean standard score of 96.3 (SD = 12.5) obtained by Māori children and that of the standardisation sample, with Māori children found to score lower on the PPVT-III than the standardisation mean, $t(45) = -2.0$, $p < .05$, ES = 0.27 (two-tailed).

Additional analyses revealed that much of the observed differences could be explained by school type. Results showed no significant difference between the mean standard score of participants attending mainstream schools (M = 100.8, SD = 11.1) and the standardisation mean, $t(32) = 0.7$, p = .76, ES = .06 (two-tailed). However, when comparing the standardisation mean with that obtained by Māori-medium participants (M = 84.8, SD = 7.9), a statistically significant difference was observed, $t(12) = -15.1$, $p < .0001$, ES = 1.27 (two-tailed).

Comparison by Age

An ANOVA revealed that there were no significant differences found between the standard scores obtained by children in each of the three age groups sampled, $F(2,43) = 1.9$, $p = .16$. The mean standard scores obtained by children in each age group were: 99.6 (SD = 11.2) for the 5-6 year group, 91.5 (SD = 12.8) for the 7-8 year group, and 97.9 (SD = 12.9) for the 9-10 year group.

Comparison by School Decile

An ANOVA comparing the average standard scores of participants from low (1-3), mid (4-7), and high (8-10) decile schools revealed no significant differences between the standard scores of each group, $F(2, 43) = 2.0$, $p = .14$.

Error Analysis

Error analysis was conducted in order to investigate the suitability of individual PPVT-III test items when used with Māori children by identifying which items resulted in higher than expected error rates, and to further investigate issues of possible cultural bias. Quantitative information was gathered in relation to the percentage, rate, and number of errors made on the PPVT-III by all participants. Qualitative information was also obtained by re-administering incorrect items to participants, and by asking them to describe the picture that depicted the correct response to the stimulus word.

Overall Error Percentage and Rate

Total error percentages were calculated for each participant by subtracting the final item administered from the first item administered, and dividing this figure by the number of

Table 1. Characteristics of Research Participants

	5 - 6 years	7 - 9 years	9 - 10 years	Total (N=46)	Population Strata
Low decile (1-3)	6	7	7	20	16.9[a]
Mid decile (4-7)	8	7	3	18	18.8
High decile (8-10)	2	2	4	8	10.1
Males	8	8	8	24	23.1[b]
Females	8	8	6	22	23.0
Mainstream	12	9	12	33	32.2[c]
Kura/ Bilingual	4	7	2	13	13.8

[a] = Quota acquired directly from Ministry of Education July 2006 SAS System figures describing Māori enrolments in the Canterbury Region by School Decile from Years 1 - 6.

[b] = Quota acquired from 2001 census data describing gender distribution of Māori children aged 5 - 19 years who reside in the Canterbury region (Statistics New Zealand, 2002)

[c] = Quota acquired from Ministry of Education figures describing rates of participation in Māori language by primary school students and includes Māori children enrolled in either bilingual or immersion school settings (Te Puni Kokiri, 2000)

Table 2. Most Common Errors Made by all Participants on the PPVT- III

Item #	Word	# of Errors	Error rate
93	Pitcher	39	91%
84	Wrench	34	79%
99	Bouquet	26	79%
100	Rodent	24	73%
89	Canoe	27	63%

errors made on the test to produce a total error percentage. The average total error percentage for participants on the PPVT-III was 35%, and ranged from 22% to 48% errors. Error rates were calculated by considering the number of errors made by participants on a particular PPVT-III item based on the number of participants who actually completed that item. Table 2 outlines which items resulted in the highest number of errors.

Error Percentage and Rate by School Type

Given the differences observed between mainstream and Māori-medium school children on standard score comparisons, additional analyses were conducted in order to compare errors on the PPVT-III. PPVT-III response sheets were reviewed manually so as to identify which particular test items might be contributing to the lower scores obtained on the PPVT-III by participants attending Māori-medium schools. Information was also collected with regards to which of the four pictures available for each item was chosen as the 'correct' response by participants. This provided the basis for identifying patterns of responses to items resulting in the most errors. Qualitative descriptions of target 'correct' pictures were also reviewed. Table 3 presents the results from this analysis. Items resulting in the highest error rate or with the most apparent pattern of errors were considered for participants from both school types. Table 4 provides details of these items, and presents the error rate obtained by mainstream and Māori-medium participants based on the number of children who completed each item.

Participant Responses to Re-administered Correct Stimulus

Qualitative information was obtained by re-administering previously incorrect items, and by asking participants to describe the picture that depicted the correct response to the stimulus word.

Table 3. Error Analysis for Participants Attending Māori-Medium Schools (n = 13)

Item #	Word	Error Rate (%)	Response to Items*				Examples of Responses to Correct Items
			1	2	3	4	
63	Luggage	9/11 (82%)	-	**C**	4	5	Bag, Suitcase
68	Signal	8/11 (73%)	**C**	2	3	3	Traffic Light
69	Squash	6/11 (55%)	2	4	-	**C**	Kumara
70	Globe	7/11 (64%)	1	**C**	5	1	Earth, World
75	Vase	6/12 (50%)	2	3	**C**	1	Pot
84	Wrench	11/12 (92%)	6	3	2	**C**	Tool, Spanner
86	Tambourine	6/11 (55%)	4	1	1	**C**	-
87	Palm	7/11 (64%)	**C**	2	-	5	Beach, Coconut Tree
88	Surprised	7/11 (64%)	2	5	-	**C**	Hat fell off
89	Canoe	7/11 (64%)	-	1	**C**	6	-
91	Clarinet	6/11 (55%)	2	2	2	**C**	Flute
93	Pitcher	11/11 (100%)	2	-	**C**	9	Jug
94	Reptile	7/11 (64%)	2	**C**	2	3	Crocodile
95	Polluting	6/11 (55%)	1	3	**C**	2	Factory
96	Vine	7/11 (64%)	**C**	6	1	-	-
100	Rodent	6/6 (100%)	1	3	**C**	2	Rat

* = For each word from the PPVT-III there are four picture choices. C relates to the correct picture choice for that particular word

Table 4. Analysis of Items with High Error Rates or Patterns Made by Participants who were Presented with the Items Attending Both Mainstream (n=33) and Māori-Medium Schools (n=13)

Item #	Word	Mainstream Group (n=33)	Māori-Medium Group (n=13)
		Error Rate %	Error Rate %
63	Luggage	82%	25%
68	Signal	73%	57%
84	Wrench	92%	74%
87	Palm	64%	13%
88	Surprised	64%	44%
89	Canoe	64%	63%
93	Pitcher	100%	85%

This information was recorded, and later coded to provide an understanding of the reasons for participant errors. The seven codes developed were: No information given, don't know, NZ term given instead, incorrect response, self correction, concrete response and description of picture given.

Responses to re-administered items were collated for all participants, with information obtained for 1196 previously incorrect responses. Once re-administered items had been coded into each of the seven response codes, totals were collated for each of the response types. This allowed a percentage to be calculated for each response code, by dividing the individual code total with the overall total of 1196 responses.

A similar pattern of responding to re-administered items was observed by children across all of the age groups tested. Based on the frequencies observed, it appeared that the most common type of response was a 'Don't Know' response, with a 'Description of Picture Given' being the second most common.

Qualitative Error Descriptions

In addition to coding participant responses to correct stimuli, qualitative information was collected in an attempt to identify themes among participant descriptions of PPVT-III items. This qualitative information was collated, and is presented in Table 5, along with the item number and target word. Qualitative comments are presented for 42 test items. This analysis and dynamic assessment approach may be useful when administering psychometric tests developed and normed with a non-Māori standardisation sample in order to consider the influence of cultural bias on test scores.

Discussion

The purpose of this study was to investigate the appropriateness of using the PPVT-III as a measure of oral language by administering the test with three different age groups of Māori children. The results showed that the PPVT-III was largely an appropriate measure to use with Māori children – particularly for children attending mainstream schools. Amongst children attending Māori-medium schools, the results showed that the current form of the PPVT-III may not be appropriate to use when measuring overall language ability, as it provides a measure of English vocabulary development alone without considering levels of te reo Māori. As a measure of English language development alone, the results suggested a number of possible adjustments that could be made when using the PPVT-III with Māori children in order to increase the suitability of the test. Given the diverse language experiences of children attending Māori-medium schools, the PPVT-III may be more appropriate when used as a measure of a child's stage of English language development, rather than an indicator of age equivalence.

While research has suggested that Māori children's reading acquisition and school based test performance may differ according to age (Westerveld & Gillon, 2001), there were no significant differences found across the three age groups on oral language abilities in the current study. Given the relatively small number of children in each age group, and the different characteristics of children in each sample, it is difficult to generalise from these results. In the event that 'true' differences do exist between the oral language skills of Māori children from different age groups, such differences may have been masked by the nature and size of the sample studied and the variable school types across the age groups.

Consideration should also be given to the possible differences between measures of oral language and reading ability. While Māori children may have little difficulty acquiring oral language skills measured by the PPVT-III, reading is a complex task that requires direct instruction and the acquisition of various component skills prior to achieving reading mastery (Phillips et al., 2004). Therefore, although Māori children across developmental age groups may have obtained similar scores on the PPVT-III, differences may still exist between their performances on reading based tests.

The influence of socio-economic factors on the distribution of economic, social, and political resources has been widely researched, with people from lower socio-economic groups found to experience disproportionate rates of deprivation in relation to these resources (Ministry of Health, 1999). However within this study no significant differences between PPVT-III scores were evident when utilising decile ratings as a possible measure of SES. Further research with a larger sample would be required in order to establish whether differences exist.

Comparisons between Māori children who participated in the research, and the standardisation sample showed that, on average, Māori children obtained a score that was below what would be expected from a normally distributed sample. However, further analyses revealed that these differences could be largely attributed to the influence of school type. While mainstream school children achieved an average that was slightly higher than the normative mean,

Table 5. Qualitative Information Provided by Participants when asked to Identify Previously Incorrect Target Items from the PPVT-III

Target Word	Qualitative Description Provided	Target Word	Qualitative Description Provided
Cactus	Kumara	Vine	Roses, Flowers, Leaves, Garden
Raccoon	Skunk	Dissecting	Frog (operating on/cutting)
Tearing	Ripping, Washing	Bouquet	Roses, Flowers
Camper	Caravan, Ambulance	Rodent	Mouse, Rat
Dripping	Tap	Valley	Farm, Village, Field
Vehicle	Car, Van	Demolishing	Wrecking, Smashing
Luggage	Bag, Suitcase	Hurdling	Jumping, High Jump
Hydrant	Water tank, For fires/firemen	Citrus	Lemon, Orange
Signal	Traffic Lights	Inflated	Balloon
Squash	Kumara, Zucchini, Avocado, Eggplant, Pinecone	Timer	Clock, Thermometer
Globe	Earth, World, Planet	Links	Chain
Nostril	Nose	Garment	Dress
Vase	Pot, For flowers, Spitoon	Fragile	Wineglass, Cup
Horrified	Frightened, Surprised, Yelling, Screaming	Adapter	Plug, Electricity
Trunk	Tree	Feline	Cat
Selecting	Shopping, Food, Lunch orders,	Wailing	Crying
Camcorder	Camera, Video camera	Coast	Beach
Heart	Puku, Tummy, Stomach	Appliance	Iron
Wrench	Tool, Screwdriver, Spanner	Hatchet	Axe
Palm	Beach, Tree, Coconut Tree	Upholstery	Couch, Sofa
Surprised	Frightened, Shocked, Scared	Exterior	Door, Front door
Canoe	Rowboat, Boat, Waka, Kayak	Consuming	Eating, Pizza
Exhausted	Tired, Sleeping	Pastry	Cake, Pie
Pitcher	Vase, Jug, Teapot	Colt	Calf, Pony, Springbok, Horse, Donkey
Reptile	Crocodile	Ladle	Soup Spoon
Polluting	Chimney, Factory, Smoke		

Māori-medium school children obtained an average that was more than one standard deviation lower. What remains unclear; however, are the reasons for the significant differences found between Māori-medium and mainstream school children, with the potential for multiple causal factors to be involved. Causal factors that may have contributed to the discrepancy between standard scores include English language curriculum differences between mainstream and Māori-medium schools, diverse family language backgrounds and experiences, and the influence of socio-economic deprivation. As only 13 Māori-medium school children took part in the study, with none of these children attending high decile schools, possibilities for further comparative analyses were limited.

Kaupapa Māori researchers working in the area of Māori-medium education have made a number of recommendations for practitioners when working with children attending Māori-medium schools. They highlight the need to move from a focus on age based norms to stage based norms when assessing te reo Māori abilities of Māori-medium school students so as to accommodate the more varied experiences and learning needs of students engaged in Māori-medium education. This approach could also be adopted when assessing the stage of English language development of Māori-medium students, by considering children's language backgrounds and levels of exposure to the English language, as well as considering the results of more formal assessment methods. Given the language demands placed on children when attempting all manner of standardised tests, these findings may imply the need for wider consideration of cultural factors when interpreting any psychometric test results for children attending Māori-medium schools.

Administration with Māori Children

An unexpected outcome arising from the qualitative analysis and community consultation hui, was the weight that was placed on the importance of utilising a flexible assessment approach that incorporated Māori processes. It is difficult to estimate the extent to which these results could be replicated if a standard approach to assessment were to be used. However, it is likely that fewer participants would have chosen to take part in the study had a Māori process not been adhered to. A clear strength of the Kaupapa Māori Research approach used throughout this study was the time that was taken when meeting with participants, to establish connections and engage them in the kaupapa of the research. This enabled participants and their whanau to voice any questions or concerns they had at the outset, and to establish a degree of trust.

Given that for some Māori children, the PPVT-III comprised items with which they had little experience, the positioning of this test within a familiar environment assisted with the development of rapport, and facilitated the ease and comfort of participants throughout the assessment process. The use of frequent encouragement by the examiner, as well as correct pronunciation and accurate understanding of Māori words was an additional factor which at times, worked to mediate the impact of participants' lack of familiarity with the test content.

Administration with Māori-Medium Participants

In addition to the administration approach discussed thus far, observations and current literature provide suggestions for ways in which to maximise the performance of bilingual children on standardised assessment measures such as the PPVT-III. Although a number of children reported disliking the error analysis phase of the research, this more dynamic approach to assessment provided valuable information which assisted when interpreting test results. For example, when re-administering incorrectly identified items to participants, the reasons for initial errors became clearer. In some cases, particularly with Māori-medium participants, errors were made due to a lack of exposure to more global English words that were depicted by familiar objects (e.g. the word citrus for the picture of an orange), highlighting the bias of some items towards English first language speakers.

Another useful dynamic adjustment to the standardised administration approach was the practice of testing below the first item administered, and beyond the last item administered during the standardised administration period. While for many participants, administering items below the first item administered was unnecessary in terms of scoring the test, this practice enabled participants to increase their sense of efficacy as they became more familiar with testing requirements. Presenting items above the last item administered posed more difficulties; particularly given the large number of incorrect items required before the ceiling point is reached in a set for a participant (eight out of 12 items incorrect). Had all participants completed one or two sets above the last item they had correctly identified, this would have enabled score adjustments to be made in response to problematic items or self-corrections made during re-administration of incorrect items.

As well as utilising more dynamic assessment techniques, Māori-medium students benefited when attention was given to the manner in which PPVT-III words and instructions were delivered. Encouraging and responding to Māori-medium participants' use of te reo helped to facilitate their comfort and engagement with the assessment experience, despite any lack of familiarity with test material. Ensuring participants had an accurate understanding of each word presented was also essential to maximise test validity, with specific techniques employed including word repetition, and clearly modelled word production by facing participants and enunciating the syllables in each word.

Error Analysis and Adjustments to Culturally Biased Items

In considering the suitability of the PPVT-III for use with Māori children, close attention was paid to the difficulty or ease with which individual items were completed. For all participants, regardless of age, gender, or school type, there were a number of words and pictures that proved to be consistently challenging. On closer examination of these items, it was discovered that they tended to characterise concepts or images more frequently encountered in the U.S.A than in New Zealand. According to PPVT-III test developers, individual item analysis and review was conducted as part of the standardisation process, with biased words and pictures being removed and difficult items re-positioned to a later point in the test as a result. It is possible that when administering the test with Māori children, some items may have posed difficulties due to the inclusion of culturally biased concepts or images that would not have been apparent when administered with an American sample. With further study, adjustments may be made to improve the familiarity and cultural appropriateness of PPVT-III content for use with Māori children.

Although additional research would be required to investigate the impact of introducing new PPVT-III test items, a number of preliminary adjustments could be made to items that represent culturally biased words and images such as replacing some existing words and pictures from the PPVT-III with equivalent concepts that are more familiar to Māori children. At present, several items depict images of animals that are not commonly seen in New Zealand. For example, item 29 (target word Porcupine) and item 44 (target word Raccoon) could be replaced by pictures and names of animals that are encountered more frequently by Māori children. For instance, while the target word Porcupine could be replaced by a word such as Hedgehog, Raccoon could be replaced by a word such as Possum.

After analysing the qualitative comments offered by participants in response to several incorrectly identified items, a range of alternative New Zealand terms describing the target picture were provided. Item 53 (target word camper) was described as being a Caravan by a number of children, highlighting the need to adjust either the target picture or word. Alternative words that may be more appropriate when presenting the same image to Māori children could include either Caravan or Campervan, with Campervan being the

closer representation of the current target picture. Item 63 (target word Luggage) also posed problems for a number of Māori participants, with Baggage providing a closer representation of New Zealand terminology. When describing the target image depicted by item 68 (target word Signal) most tamariki used the term Traffic Lights. Future amendments to the PPVT-III could use the term Lights in place of the word Signal when working with Māori children. Tamariki also provided alternative words for items 82 (target word Camcorder), 84 (target word Wrench), and 93 (target word Pitcher). While the image depicting the target word Camcorder was described by Māori children as a Video camera, Wrench was more commonly associated with the word Spanner, and Pitcher was described as a Jug.

In addition to challenges brought about by the use of unfamiliar terms, some items resulted in high error rates due to their depiction of images not commonly seen in a New Zealand context. For example, item 65 (target word Hydrant) depicted an image of an American steel fire hydrant, while the target picture for item 69 depicted an image of the vegetable squash not often seen in New Zealand grocery stores. When re-administering item 69 (target word Squash) to Māori participants, the image was described as being a Kumara, Zucchini, Avocado, Eggplant, or Pinecone. Future adaptations of the PPVT-III for use with Māori children could trial the use of more familiar alternative terms and images, such as Post-box and Pumpkin, to replace items 65 and 69.

After reviewing the responses provided by Māori-medium participants to PPVT-III items producing high error rates, it became apparent that not all target and decoy pictures were functioning effectively. According to PPVT-III test developers, decoy pictures were intended to act as neutral distracters, with no single decoy designed to be more or less attractive or misleading than others. Items 88 (target word Surprised) and 89 (target word Canoe) may be improved for use with Māori by developing an alternative set of picture plates. Close investigation of participant responses revealed that many children who responded incorrectly to items 88 and 89, had chosen the same decoy images. This suggests that the decoy images did not allow Māori children to discriminate easily between correct and incorrect choices. As well as improving the precision of target and decoy images, future versions of the PPVT-III may also consider updating the picture plates in order to improve their overall relevance and visual appeal.

After reviewing the reasons for participant errors through the re-administration of incorrectly identified items, it appears that most participant errors occurred in response to items that were unknown to them. What is not clear, however, are the reasons that these items were unknown to participants, with further research required in order to gauge whether PPVT-III items are currently ordered in a difficulty hierarchy that matches the experiences of Māori children.

Limitations

The current paper was produced from results obtained from a Masters thesis, at which time the PPVT-III was the latest available edition. The PPVT-III has now been superseded by the PPVT-4 and the authors acknowledge that this may address some of the issues of cultural bias raised in this research paper.

Detailed analyses of the results revealed the diverse characteristics of the research sample in terms of demographic and developmental backgrounds. High rates of attendance at early childhood services among participants also represented a deviation from what would be expected in the wider population given the lower rates of participation in early childhood education by Māori children overall. Despite these differences, the research sample bore close resemblance to the wider population on all other measures and strata utilised.

Small sample sizes also limited the range of analyses that were possible. In addition to the small number of participants attending Māori-medium schools, there were inadequate numbers of participants overall to allow more comprehensive error analyses to be conducted. For example, Rasch scaling methods employed by PPVT-III test developers require at least 100 participants to complete all test items in order to identify points at which an item may be functioning differently in relation to the entire item pool.

While the use of te reo Māori throughout the administration period was always initiated by participants, it is possible that the combination of both English and te reo Māori may have added increased complexity to the cognitive demands involved in the task and may subsequently have impaired test performance. While standardised administration instructions were always presented in English, some participants chose to respond in te reo (e.g. tahi, rua, toru, wha), and for these participants encouragement was often provided in te reo Māori by the test administrator in order to validate and mirror their use of language. Further research investigating the use of te reo Māori in psychometric assessment would be required in order to clarify the impact on test performance.

Given that the focus of this research was to investigate the suitability of using the PPVT-III with Māori children, all of the participants who took part in the study identified as Māori. While it is possible that difficulties may arise when administering the PPVT-III with non-Māori children, further research is required in order to establish what, if any adjustments could be made to the PPVT-III for use with non-Māori New Zealand children.

Conclusions

Although non-parametric sampling methods such as those employed in this study have been criticised in the past for producing research that cannot be generalised, the research sample who took part in this study were found to closely match the wider population of Māori residing in Canterbury and New Zealand on a number of strata. Despite this criticism, researchers have emphasised the utility of such methods when recruiting highly defined and representative minority samples (Clay et al., 2003; Hicks Patrick, Pruchno, & Rose, 1998; Perkins, Devlin, & Hansen, 2004). These findings have implications regarding the appropriateness of using a standard approach to test administration when working with Māori, particularly when using

psychometric tests developed outside of New Zealand. Although standardised test administration procedures are provided with the intention of minimising the effects of bias introduced by examiner variation, adherence to procedures that do not prioritise Māori beliefs, values, and experiences may disadvantage Māori. Given that the PPVT-III was thought to be a relatively unbiased test, the continued unadjusted use of psychometric tools with Māori that are based on concepts with a known cultural loading is a questionable practice.

Accurate and unbiased assessment is an essential requirement when identifying difficulties, and providing interventions capable of responding to the needs of Māori. While work continues towards the development of culturally unbiased assessment tools that are based on Māori concepts and experiences, the majority of practitioners and tools currently available to work with Māori clients do not operate from a Māori perspective. Findings revealed that although the PPVT-III was not completely free of bias, it appeared to be largely suitable when used with Māori children. However, due to the diverse language and cultural backgrounds of Māori children, it was discovered that flexible adjustments to the standardised administration approach improved the suitability of the test for many participants.

While Māori children attending mainstream schools produced PPVT-III results which mirrored those of the standardisation sample, children attending Māori-medium schools produced scores that were more indicative of their stage of English language development rather than their overall language ability. In fitting with recommendations made by researchers working in the area of Māori-medium education, the PPVT-III could continue to be used, but in a manner that understands and addresses the needs of Māori-medium students. In doing so, the PPVT-III could be utilised to assist with teaching interventions when introducing English as a subject with Māori-medium students, or to gauge the stage of language development of new students with diverse language backgrounds. This research shows that culture plays an important role in the ways in which people respond to standardised assessment. One cannot assume that a test developed based on concepts and ideals valued by one culture will translate equally to another.

Acknowledgements

To the key Māori stakeholders, tamariki and their whanau who supported the kaupapa of this research – kia ora. To practitioners who strive to deliver culturally responsive services to indigenous people throughout the world – kia kaha.

References

Clay, C., Ellis, M. A., Amodeo, M., Fassler, I., & Griffin, M. L. (2003). Recruiting a community sample of African American subjects: the nuts and bolts of a successful effort. *Families in Society, 84*(3), 396.

Hicks Patrick, J., Pruchno, R. A., & Rose, M. S. (1998). Recruiting research participants: a comparison of the costs and effectiveness of five recruitment strategies. *The Gerontologist, 38*(3), 295-302.

Laing, S. P., & Kamhi, A. (2003). Alternative Assessment of Language and Literacy in Culturally and Linguistically Diverse Populations. *Language, Speech & Hearing Services in Schools, 34*, 44-55.

McCloskey, D., & Athanasiou, M. S. (2000). Assessment and Intervention Practices with Second-Language Learners Among School Psychologists. *Psychology in the Schools, 37*(3), 209-225.

Ministry of Health (1999). Our Health, Our Future: The Health of New Zealanders 1999. Wellington: Ministry of Health.

Ogden, J. A. (2003). Adapting neuropsychological assessments for minority groups: a study comparing white and Maori New Zealanders. *Brain Impairment, 4*(2), 122-134.

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

Palmer, S. (2004). Homai te Waiora ki Ahau: A tool for the measurement of wellbeing among Maori - the evidence of construct validity. *New Zealand Journal of Psychology, 33*(2), 50-58.

Perkins, M. R. V., Devlin, N. J., & Hansen, P. (2004). The validity and reliability of EQ-5D health state valuations in a survey of Maori. *Quality of Life Research: an International journal of quality of life aspects of treatment, care and rehabilitation, 13*(1), 271-274.

Phillips, G., McNaughton, S., & MacDonald, S. (2004). Managing the mismatch: enhancing early literacy progress for children with diverse language and cultural identities in mainstream urban schools in New Zealand. *Journal of Educational Psychology, 96*(2), 309-323.

Reese, E., & Cox, A. (1999). Quality of adult book reading affects children's emergent literacy. *Developmental Psychology, 35*(1), 20-28.

Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur Communicative Development Inventory: Words and Sentences. *Journal of Child Language, 27*, 255-266.

Rock, D. A., & Stenner, J. (2005). Assessment issues in the testing of children at school entry. *The Future of Children, 15*(1), 15-34.

Saenz, T. I., & Huer, M. B. (2003). Testing Strategies Involving Least Biased Language Assessment of Bilingual Children. *Communication Disorders Quarterly, 24*(4), 184-193.

Shepherd, I., & Leathem, J. (1999). Factors affecting performance in cross-cultural neuropsychology: From a New Zealand bicultural perspective. *Journal of the International Neuropsychological Society, 5*, 83-84.

Statistics New Zealand. (2002). *2001 Census of Population and Dwellings: Maori* (No. Catalogue 02.217.0001). Wellington: Statistics New Zealand.

Stockman, I. J. (2000). The new Peabody Picture Vocabulary Test-III: an illusion of unbiased assessment? *Language, Speech & Hearing Services in Schools, 31*(4), 340-353.

Te Puni Kokiri. (2000). The Maori Language and Education. *Fact Sheet:8* Retrieved 21 November, 2006, from http://www.tpk.govt.nz/maori/language/fs8language.pdf

Tzuriel, D. (2000). Dynamic Assessment of Young Children: Educational and Intervention Perspectives. *Educational Psychology Review, 12*(4), 385-435.

Washington, J. A., & Craig, H. K. (1999). Performance of at-risk African American preschoolers on the Peabody Picture Vocabulary Test-III. *Language, Speech, & Hearing Services in Schools, 30* (1), 75-82.

Westerveld, M., & Gillon, G. (2001). Oral language sampling in 6-year-old New Zealand children from different cultural backgrounds. *New Zealand Journal of Speech-Language Therapy, 56*, 5-17.

Corresponding Author:
Tracy Haitana
CPPS: Christchurch Psychologists' Office
PO Box 25 146
Christchurch
New Zealand.
tracy.haitana@corrections.govt.nz