

# Culturally Valid Assessment Tools: Northern Sotho Translation of the Peabody Picture Vocabulary Test - Revised

Corrieta Pakendorf and Erna Alant

Department of Communication Pathology  
University of Pretoria

## ABSTRACT

*There is currently a great demand for service provision for the African language speakers in South Africa. The difficulties associated with assessing speakers in the absence of assessment tools in the indigenous languages is, therefore, also a pertinent concern. Within the current socio-economic climate in South Africa where test translation and adaptation is often cited as a more viable option than that of developing new tests, very few guidelines exist for the development or adaptation of valid assessment tools for culturally and linguistically diverse population groups. This article is aimed at describing the process which took place when existing English test material, in this instance, The Peabody Picture Vocabulary Test-Revised (PPVT-R) (Dunn & Dunn, 1981) was translated and culturally adapted for the Northern Sotho population in Pretoria and surrounding areas. The findings of the research include practical examples of methodological considerations which should be taken into account while translating and undertaking cultural adaptation of test material. The newly adapted test material was also applied to a sample of 152 North-Sotho speaking pupils in the study area and the test results are discussed.*

## OPSOMMING

*Hierdie studie het voortgespruit vanuit die huidige behoefte aan dienslewering binne die Afrikatale-konteks en die gebrek aan evalueringshulpmiddels in die inheemse tale. Alhoewel die Suid-Afrikaanse sosio-ekonomiese klimaat eerder toetsvertaling en -aanpassing aanmoedig, as die ontwikkeling van nuwe toetsmateriaal, bestaan daar tans min riglyne vir die beplanning en uitvoering van so 'n prosedure. Hierdie artikel verskaf 'n omskrywing van die prosedure wat gevolg is tydens die vertaling en kulturele aanpassing van die PPVT-R vir die Noord-Sotho-populasie in Pretoria en omliggende gebiede. Die navorsingsresultate sluit praktiese voorbeelde in van metodologiese stappe wat geneem kan word tydens die vertaling en aanpassing van standaard-toetsmateriaal. Die nuut aangepaste toetsmateriaal is ook toegepas op 152 Noord-Sotho-sprekende leerlinge in die toets-area en die resultate word bespreek.*

**KEY WORDS:** African languages, cross-cultural language testing, cultural validity, diverse populations, mother-tongue, non-mainstream, PPVT-R, receptive vocabulary, test translation, test adaptation.

## INTRODUCTION

Speech-language pathologists in South Africa are now, more than ever, faced with the evaluation of the African Language speaking child. Especially in cases where speech-language and/or learning problems are suspected, the challenge for the speech-language pathologist as professional team member is to evaluate the child in a non-biased manner and then to determine if the child demonstrates a communication difference or disorder (Roseberry-McKibbin & Eicholtz, 1994). Unfortunately, there are realities that present major obstacles to providing non-biased and appropriate services for these children. For example, a general need exists amongst professionals who are not knowledgeable about these cultures or languages for original tests in the African languages. This indicates the importance of the development of such evaluation instruments. It must, however, be kept in mind that the study of normal and pathological communication is couched in cultural terms. Knowledge of each community's own definition of "normalcy" in language and communication is,

therefore, a prerequisite for the development of original and culturally fair tests. This necessitates ethnographically-based research on the subject of African languages and culture (Taylor, 1986). Such a research undertaking is both resource- and time-consuming (Jordaan, 1989) and the development of original tests is thus not always a practical option within the current South African situation.

One of the alternatives to the application of inappropriate language tests with the non-mainstream language user is that of modifying or revising existing tests in ways that will make them appropriate for the intended test population (Bradley in Taylor, 1986). The present study's main aim was to explore methods for the adaptation of an existing translated (Northern Sotho) language test into a culturally valid version for a specific population. The importance of a cultural orientation, as suggested in Taylor's conceptual framework, becomes evident especially when applying the same test to different populations. In a study aimed at the investigation of the validity and cross-cultural generalisability of English language tests, Zeidner (1986) found that test scores show significant bias, tend-

ing to disadvantage the non-standard English user. Various researchers have stressed the need to recognise the limitations of using Standard English tests on different cultural and linguistic groups (Alant & Beukes, 1986, Roseberry-McKibbin & Eicholtz, 1994).

The importance of considering the cultural and sociolinguistic background of the intended test population when adapting existing tests or when creation of a new test is undertaken, has been strongly emphasised in the international literature (Miller, 1984; Taylor, 1986). Furthermore, it is important that, not only the translation of a test, but also the applicability of test material such as stimulus pictures must be taken into account when adapting an evaluation instrument for the intended test population.

The present study included the evaluation and further refinement of the most recent English-to-Northern Sotho translation (Koekemoer, 1994) of the PPVT-R as well as an investigation of the cultural appropriacy of the original picture plates for the intended test population. Northern Sotho is the home language of approximately 10% of all South Africans and two previous studies (Labuschagne 1990, Koekemoer, 1994) had been conducted aiming at translating the original English word lists into Northern Sotho. A third study (Naidoo, 1994) was aimed at the adaptation of the PPVT-R for the Zulu-speaking community in KwaZulu-Natal. These studies provided a starting point and certain guidelines which could be utilised during the present research.

The evaluation of a child's receptive vocabulary skills is a practical starting point for any language assessment, especially in the South African context where limited information is available regarding linguistic analysis of the African languages (Jordaan, 1989). Each child's knowledge of vocabulary determines his or her ability to think, speak and understand a language (Miller, 1984). The importance of vocabulary is also reflected in the fact that vocabulary teaching has become an important focus in second language teaching and learning (Morgan & Rinvulcri, 1986). It should, furthermore, be possible to apply a receptive vocabulary test even with young children, seeing that receptive language development can be evaluated without requiring any verbal responses from the subject. Especially in the case of a young child, the PPVT-R, when used in conjunction with a measure of school achievement, could be helpful in identifying underachievers (Dunn & Dunn, 1981). Viewed in the light of current efforts to improve the educational standards within the African language communities in South Africa, the relevance of a Northern Sotho instrument for the evaluation of receptive vocabulary once more becomes clear.

## METHOD

### AIMS

#### Main aim

The main aim of the study was to investigate the applicability of a Northern Sotho translated version of the Peabody Picture Vocabulary Test - Revised (PPVT-R, Dunn & Dunn, 1981) to a group of North-Sotho speaking children.

The main aim was realised by means of 2 sub-aims:

#### Sub-aim 1:

The pre-experimental phase of the study was aimed at the development of a culturally fair Northern Sotho ver-

sion of the PPVT-R. The most recent Northern Sotho version (Koekemoer, 1994) of the test was adapted in order to make it as appropriate and as applicable as possible for the intended population.

#### Sub-aim 2:

The applicability of the modified Northern Sotho translation of the PPVT-R to a group of Northern Sotho children in Grade 1, Grade 2 and Grade 3 was investigated during the experimental phase of the study.

## RESEARCH DESIGN

During the pre-experimental phase of this research there was aimed at further development and refining of the existing Northern Sotho version of the PPVT-R. Given the complexity of the cultural adaptation of a test, it was decided to, during this part of the study, implement the qualitative research design as suggested by Creswell (1994). This design made it possible to adjust and re-adjust according to unforeseen results obtained from completed parts of the project. This continuous developmental process also implied the repetition of certain research procedures in a circular manner in order to refine all the test stimuli as well as possible.

In order to maximise the cultural validity of the adapted test material, the pre-experimental phase of the study was also characterised by a participative methodology. This entailed that Northern Sotho mother-tongue speakers, professionals as well as non-professional local residents living in the target area, were involved in the formulation and development of test-material on a consistent basis. Marshall & Rossman (in Creswell, 1994) sustain this methodology by suggesting that when applying the qualitative research design, it entails the researcher entering the everyday life of the setting chosen for the study through ongoing interaction with informants, and by seeking their perspectives and meanings. This methodology characterised the developmental (pre-experimental) phase, but was also implemented during the experimental phase of the research.

It is important that the methods of research are not only innovative but also soundly-based and scientifically accountable. Therefore, an analytical survey design was implemented during the second (experimental) phase of research. During the experimental phase, where pupils were tested individually, the acquired data was of a quantitative nature and statistical assistance was needed to extract the meaning. Intergroup comparisons were done by which the performance of different age groups and academic year groups were statistically compared. In this way the applicability of the newly adapted test material, within the specific scholastic context involved in the present study, could be investigated.

## PROCEDURE

### PRE-EXPERIMENTAL PHASE

#### Participants in the pre-experimental phase

Two different groups of participants were involved in this part of the study.

Firstly, a total of 21 normal Northern Sotho pupils were tested. These pupils had to comply with the same requirements as the subjects involved in the experimental phase.

The distribution of scholastic year groups was as follows: Grade 1 : n = 5, Grade 2 : n = 5 and Grade 3 : n = 11.

Secondly, the researcher used the help of a panel of Northern Sotho speaking adults, consisting of three academically qualified (tertiary education) as well as two non-academic (primary school education) residents from the test area. These selected North-Sotho speakers informed the researcher on the cultural and linguistic characteristics of the intended test population on a continuous basis. In this article the panel is referred to as “the panel of experts”. Two lecturers from African Languages Departments, both native speakers of Northern Sotho and proficient in English were included in the panel. It was important to include native language users who are familiar with the language use and educational demands set in the geographical area. A student, majoring in African languages, was also included in the panel. This student lives in the research area and is proficient in both English and Northern Sotho.

The variation of Northern Sotho which is the subject of interest in the present study has been classified by Schuring (1984) as a Koine-language. One of the outstanding characteristics of such a language is that the vocabulary and idiom are constantly expanding and changing. This implies the necessity of the input of research assistants who are up to date with the latest de-

velopments in terms of the spoken language in the area. Therefore, in addition, two non-professional persons residing in the research area were included in the panel which judged the appropriacy of the translated test items of the PPVT-R as well as the appropriacy of the picture plates which represent these words. Both persons are native speakers of Northern Sotho and proficient in English. Seeing that they do not use these languages on an academic level, these persons showed a fine attunement for the everyday language use in the geographical area included in the study.

Description of procedure and results: Pre-experimental phase

During the first part of the study four procedural steps were followed, certain of which were repeated in a circular manner in order to refine the test stimuli as well as possible.

The first step of the pre-experimental phase of the study entailed the evaluation and further refinement of the existing Northern Sotho version of the PPVT-R by the above-mentioned panel of selected North-Sotho speakers. Specifically the correctness of the Existing English-to-Northern Sotho translation was investigated.

TABLE 1. Description of the Pre-experimental phase of the study

Step 1 of the pre-experimental phase:	
Procedure and Results	Outcomes
The original English word list together with the original picture plates and existing Northern Sotho translation (from Koekemoer 1994) were given to the panel of experts for evaluation and suggested corrections, if necessary. 15 words were deleted from the original list. The panel judged 9% of the translated words to be incorrectly translated. A further 2% were evaluated as culturally unsuitable words for the target population. See Pakendorf (1996) for detailed information regarding the adaptation of test material.	It was important to supply the pictorial stimuli which represent each word in the PPVT-R, together with the word list to the panel while they were evaluating the translation. In this way the panel members were able to identify certain items where previous translations did not correspond with the given pictorial stimuli (e.g., as in the word “tire” which has more than one meaning).
Step 2 of the pre-experimental phase:	
The second step of the pre-experimental phase was specifically aimed at the evaluation of the original picture plates of the PPVT-R in terms of cultural and conceptual appropriacy for the test population.	
Procedure and Results	Outcomes
A file with the original (English) picture plates from the PPVT-R was supplied to the panel of experts to be evaluated. The panel’s suggestions regarding culturally or conceptually acceptable replacements for certain items was analyzed according to number and frequency of suggestions. The only condition was that the lecturers had to be satisfied with the appropriateness of the modified test material. 6% of the picture plates were viewed as inappropriate by the panel.	In previous studies of a similar nature little attention was given to the appropriacy of the stimulus pictures.  It was, however, noted by the field worker that not only the primary stimulus pictures, but also certain of the distracter pictures might have a negative effect on the targeted test population’s performance. This is an important issue which was closely monitored while analyzing the pupils’ responses to the pictorial stimuli during the pre-experimental phase of the research.
Step 3 of the pre-experimental phase:	
Based on the results obtained during the first two steps of the pre-experimental phase, a third step was carried out with the main aim of substituting certain stimulus words with more relevant Northern Sotho words and the adaptation of certain pictorial stimuli. These were carried out in order to upgrade the quality of the translated test and the applicability of the picture stimuli.	

Procedure and Results	Outcomes
The panel suggested alternative translations for 36% of the existing Northern Sotho word list. Furthermore, the researcher modified 6% of the picture plates which were previously identified by the panel as being culturally unsuitable for the intended test population. The members of the panel were required to reach consensus on the applicability of all replaced items. A modified word list and booklet of the pictorial stimuli were compiled. After implementing the adapted test material, re-adapting and a second application thereof, the panel was satisfied with the applicability of the modified test material for the intended test population.	It was relevant to have academic as well as local residents (from the research area) with lesser academic qualifications on the panel. Both these groups gave different, but relevant recommendations during the evaluation process. The inclusion of multiple sources of data collection in the present research project is likely to increase the reliability of the findings (Mouton & Marais, 1990). Previous studies were noted where translations were obtained from only one person (Labuschagne, 1990) or from only academically qualified persons who did not all originate from the intended research area (Koekemoer, 1994).
<p style="text-align: center;"><b>Step 4 of the pre-experimental phase:</b></p> <p>The fourth and final step of the pre-experimental phase was aimed at the application of the refined Northern Sotho version of the PPVT-R to a North-Sotho population in order to evaluate the relevance of the test items (pre-experimental phase of the study).</p>	
Procedure and Results	Outcomes
<p>A group of 15 pupils were tested individually with the help of one of the panel members. The testing procedures were adhered to as suggested by Naidoo (1994) in a study of similar nature in KwaZulu-Natal. All test items where 42% or more of the pupils' responses were incorrect were identified to be analyzed in further detail. Further modifications to these items were made by the panel of experts.</p> <p>A second group of pupils (n=6) was tested with the newly adapted test material. The test procedure was also modified (See next column). The pupils responded favourably to the new test material during this phase of testing. The modifications proved to be satisfactory.</p>	<p>It was found that the use of a pre-recorded version of the verbal test-stimuli was time-consuming and impersonal. The pupils were not at ease with a tape recorder and the presentation of stimuli was often delayed by stopping and rewinding the tape.</p> <p>Therefore, the stimuli (words) were presented verbally and directly by the North-Sotho fieldworker during the second testing phase. This method, which is the suggested method from the original PPVT-R, proved to be more successful than using a tape recording.</p>

#### **DISCUSSION AND CONCLUSION: PRE-EXPERIMENTAL PHASE OF THE RESEARCH**

As part of the further development of the latest existing translated Northern Sotho version (Koekemoer, 1994) of the PPVT-R two main areas were investigated. Firstly, there was aimed at the development of a culturally applicable Northern Sotho word list and secondly, cultural adaptation of the pictorial stimuli on the test was carried out.

##### **Word list**

The following four major factors were regarded as influential in the development of a culturally fair Northern Sotho word list:

- **Koine language:** It has long been established that Pretoria Sotho is a Koine language (Schuring, 1984). This implies that Pretoria Sotho is an autonomous popular language, and that it differs from the related standard language. Both recent studies on the PPVT-R with Northern Sotho populations in South Africa (Labuschagne, 1990; Koekemoer, 1994) were carried out in the same test area as the present study. It was therefore interesting to note that the evaluation carried out as part of the present study indicated 21% of the previous translations to be non-Koine language for the region,

but rather a translation into the standard form of Northern Sotho.

As stated by Tannen (1982), it is imperative to distinguish between "oral" and "literate" language when evaluating a child's way of expression and his approach to knowledge and thought. The African languages contain a vast number of fixed expressions and especially "animal analogues" and different regions use different expressions. Therefore it is especially relevant to adhere to a region's indigenous form of a language when developing a tool for the evaluation thereof.

- **Multiple data sources:** Regional language use was further promoted by working with a research team (panel of experts) consisting of multiple data sources. Not only lecturers and teachers were included in the panel, as was done in previous studies. During the present study local residents with lesser academic qualifications, but a good command of English and Northern Sotho also made valuable suggestions regarding adaptation of the test material based upon their knowledge of the regional language use.
- **Supplying pictorial stimuli:** When working with the PPVT-R, the word is used as the unit of translation. Single words, however, present the translator with a decontextualised situation (Newmark, 1988). This could explain why it was noted during the present study that 9% of the existing items had previously been incorrectly translated from English into Northern Sotho. By sup-

plying the translators with both the word lists and the stimulus pictures for each item, the translation and the intended stimulus picture could be matched. In this way, a more concrete context for translation of the word was established. Consequently, it was possible for the panel of experts to modify the incorrect translations on the existing list (Koekemoer, 1994) of words.

Stimulus and distracter pictures

No examples in previous South African studies on the PPVT-R could be found where cultural adaptation of the pictorial stimuli was carried out. In the present study the panel of experts rated 6% of the translated items' pictorial stimuli to be culturally inappropriate for the intended test population. It was interesting that, with the exception of one, all these stimulus pictures represented items of a non-human nature, e.g., helicopter, ostrich and root. This tendency is of importance because it is stipulated in the PPVT-R manual (Dunn & Dunn, 1981) in the original development of test plates in the USA the aim was to "...maintain careful sex and ethnic balance on the plates containing humans...". It therefore seems that the plates containing humans were also appropriate for the test population in the present study.

While planning the methodology of the present study care was taken to preserve as much as possible of the original character of the PPVT-R. Therefore the identified unsuitable pictorial stimuli were only altered to a minimal extent and substitutions were limited as far as possible.

Furthermore, the original distracter pictures were meant to be of a neutral nature and usually belonged to the same semantic category than the intended stimulus word. Seeing that these sets of pictures were developed within a different culture than that of the intended test population, the researcher viewed it as acceptable when the panel of experts suggested various subtle changes to the distracter pictures before the testing commenced.

EXPERIMENTAL PHASE

Participants in the experimental phase of the study

A total of 152 normal Northern Sotho primary school pupils from two schools in the Mamelodi area were tested.

The distribution of scholastic year groups involved in the experimental phase of the study was as follows: Grade 1 : n=50, Grade 2 : n=53 and Grade 3 : n=49. The stratified sampling framework (Smit, 1983) was used in planning the pupil selection. The aim was to select 25 students from each of these three academic year groups in each of the two schools (See Table 2).

As shown in Table 2 there was an even distribution in terms of gender and of the two relevant age groups within each academic year group. Ages were documented according to the child's actual age on the day of testing.

Selection criteria for pupils involved in the study

Information obtained from school records and teachers was used to identify pupils who met the following criteria:

- **Mother-tongue:** Northern Sotho was required as the home language used by the pupils. The teachers at each

school selected pupils whose mother-tongue, according to them, was Northern Sotho. Each pupil also had to state his dominant home language to the field worker before commencing with the test.

- **Intelligence:** In the absence of appropriate intelligence tests for Northern Sotho children, pupils who were viewed by their teachers to be of average or above average intelligence were included in the study. This selection criterion was necessary as below average intelligence could influence language performance and test results.
- **Academic record:** Children were required to have an academic record with no previous failures or repetition of a school year.
- **Sensory and emotional development:** Children were required to have no known visual, auditory, social or emotional problems, as these factors could influence speech and language development. Social and emotional problems could also lead to poor co-operation and behaviour during testing which may influence the results.

Fieldworkers

The time taken for individual administration of the test to the pupils necessitated the employment of research assistants in the study.

- **Selection of field workers:** Four North-Sotho speakers from the study area were recruited. The assistants' levels of education varied from Grade 8 to post graduate university training. Three of the selected assistants had been involved with the pre-experimental phase of this research and were therefore familiar with the test and research goals. All four research assistants were female.
- **Training of field workers:** The assistants were orientated in one morning to ensure their familiarity with the test procedure. The training included explanations of the goal of the study, training in terms of the general administration of the PPVT-R and recording of the responses as suggested by Dunn & Dunn (1981). All procedures and concepts on the test were clarified prior to the onset of formal pupil testing. Furthermore, a simulated testing situation was included and each assistant took a copy of the pictorial stimuli and the test form home in order to familiarise themselves further with the test material.

On the days of testing the research assistants were personally supervised by the researcher, especially since each of them had to present the word stimuli to the pupils themselves.

TABLE 2. Exposition of stratified population groups involved in the experimental phase of the study

GENDER	GRADE 1	GRADE 2	GRADE 3
Male	14 X 6 yrs	15 X 7 yrs	14 X 8 yrs
Male	11 X 7 yrs	11 X 8 yrs	11 X 9 yrs
Female	14 X 6 yrs	16 X 7 yrs	13 X 8 yrs
Female	11 X 7 yrs	11 X 8 yrs	11 X 9 yrs

### General procedures during the experimental phase of the study

In order to assure reliability, specific consideration was given to the following aspects:

- The purpose and importance of the planned study was explained to the principals and teachers involved with the selection of the pupils (n=152) involved in the testing. Letters and data forms were handed out to the teachers and collected a week later. The researchers grouped the selected pupils according to the stratified sampling framework (Smit, 1983).
- The researcher was present in the classroom where all four research assistants simultaneously tested the pupils. Each assistant was positioned in a different corner of the room with adequate lighting, a table and two chairs.
- Each pupil was tested individually by a research assistant and his or her responses were recorded on a separate response sheet by writing the number of the picture plate to which the pupil pointed. The average testing time per pupil was ten minutes.
- Lastly, each pupil's responses during the experimental phase was computed according to the original procedure as described by Dunn & Dunn (1981). The procedure was, however, altered by requiring each pupil to respond to all the test items. The results were, therefore, not scored according to a basal or ceiling item during the testing. In this way more data could be obtained for analysis and making deductions regarding the applicability and arrangement of all the test items.

### Analysis of the data from the experimental phase of the study

The data obtained during the experimental phase was computerised and calculations, descriptive statistics and analysis were carried out by means of the SAS (DATA procedure + PROC MEANS) and BMDP (PROC NPAR 1WAY + PROC BMDP3S) statistical packages. All the results were compared on a 5% level of difference. The statistical analysis of the data was greatly aimed at comparing the performance of different age groups and pupils on different academic levels.

Three different scores were calculated for each pupil involved in the experimental phase of the study:

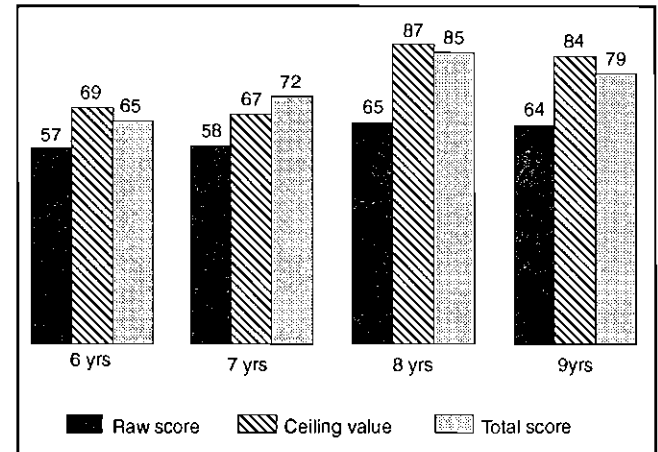
- **Ceiling value:** The last test item where a pupil answers six out of eight consecutive questions incorrectly.
- **Raw score:** The total of all the correct answers below the ceiling item.
- **Total score:** Seeing that the target population differs vastly from the original population as described by Dunn & Dunn (1981) it was decided to further experiment with the research data from the experimental phase of the study by not using the ceiling item (6/8 consecutive answers correct) to determine where to stop measuring the pupils' performance. This implies that each pupil had the opportunity to answer all the items (n=88) included in the present adapted Northern Sotho version of the PPVT-R. The scores derived by means of this procedure is referred to as the "total score" of the pupil. By adhering to the above-mentioned experimental procedure, the possibility of an inappropriate order

of arrangement of test items could be ruled out and more comprehensive statistical data could be gained from the study.

## DISCUSSION AND INTEGRATION OF RESULTS FROM THE EXPERIMENTAL PHASE OF THE STUDY

### PERFORMANCE ACCORDING TO AGE GROUPS

The mean raw scores, ceiling values and total scores obtained by each age group involved in the experimental phase is depicted in Figure 1 below.



**FIGURE 1: Mean raw scores, ceiling values and total scores obtained by each age group**

Statistical analysis of the data, which is visually represented in the above bar chart, indicates a significant increase in raw scores as well as in ceiling values, in conjunction with an increase in age, between the seven and eight year old pupils only. The differences between the six and seven year old groups, and the eight and nine year old groups were not of a statistically significant nature. Although an increase in total scores was measured between the six, seven and eight year old groups, this increase was not of a statistically significant nature. (Refer to Pakendorf, 1996 for detailed statistical data).

According to the above results the important observation was made that an increase in age did not directly imply significant growth in terms of receptive vocabulary skills in the present test population. Therefore, it was relevant to explore the possible effect of factors such as academic exposure on the level of receptive vocabulary development in the test population.

### PUPILS FROM DIFFERENT SCHOLASTIC YEAR GROUPS:

In addition to the procedure of the original PPVT-R (Dunn & Dunn, 1981) the academic exposure of each pupil was also taken into account, seeing that the researcher hypothesised that the development of receptive vocabulary within the selected test population might be more directly influenced by academic exposure than by age.

Statistical analysis of the data depicted in Figure 2 indicated that the performance (mean raw score, mean ceiling value and mean total score) of the pupils in Grade 1 was significantly lower than that of the rest of the pupils

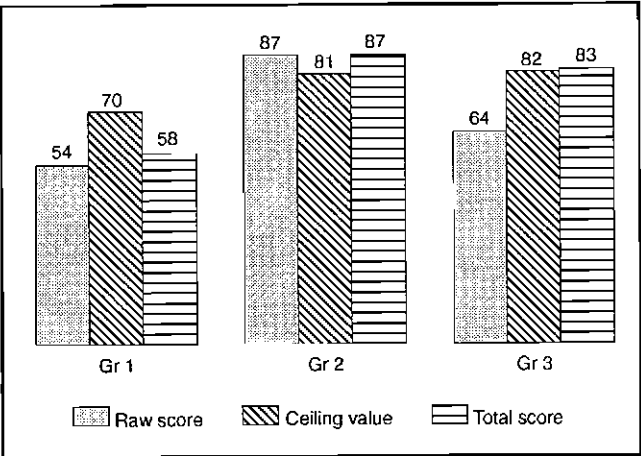


FIGURE 2: Mean raw scores, ceiling values and total scores obtained by each scholastic year group

involved in the study. It is also important to note that the Grade 3 pupils did not obtain statistically significant higher scores than the Grade 2 pupils involved in the study. This indicates that the academic exposure which takes place during the first school year plays a vital role in the development of the present test population's receptive vocabulary.

The overall tendency especially of an increase between the performance of the seven and eight year old pupils and between the performance of the Grade 1 and Grade 2 pupils indicated the importance of making comparisons between pupils of the same age group but from different academic levels and vice versa. Consequently the following intergroup comparisons were carried out:

TABLE 3: Mean raw scores obtained by pupils from the same age group but different scholastic year groups

School Standards Compared	Age group	Mean Raw Scores	P-value: Significant if P < 0.05
Grade 1 vs Grade 2	7 years old	50 vs	P = 0.004*
Grade 2 vs Grade 3	8 years old	64 vs 66	P = 0.57

PUPILS FROM THE SAME AGE GROUP BUT DIFFERENT SCHOLASTIC YEAR GROUPS:

This study investigated the probability that children might show an improvement in performance in relation to their degree of academic exposure and stimulation, rather than as a result of an increase in age. The 7 year old pupils in Grade 2 did better than the 7 year old pupils in Grade 1. However, the raw scores of the 8 year old pupils did not show a significant improvement together with academic progression from Grade 2 to Grade 3. The degree of receptive vocabulary development which takes place during the first two school years in the present target population, therefore, seems to be of special significance.

TABLE 4: Mean raw scores obtained by pupils from the same school year, but different age groups

School Standards Compared	Age group	Mean Raw Scores	P-value: Significant P < 0.05
6 years vs 7 years	Grade 1	67 vs 50	P = 0.17
7 years vs 8 years	Grade 2	65 vs 64	P = 0.07
7 years vs 8 years	Grade 3	66 vs	P = 0.42

\* = Significant difference between the two groups

PUPILS IN THE SAME SCHOOL YEAR, BUT OF DIFFERENT AGE GROUPS:

By means of statistical analysis of the data represented in Table 4, it was found that the pupils' ages in all three academic year groups did not significantly influence their performance in comparison to other (older or younger) children on the same academic level.

These findings are further supported by the mean receptive vocabulary development of pupils belonging to the same age group, but two different scholastic year groups in this study. The 7 year old pupils in Grade 2 achieved significantly higher scores than pupils of the same age group on a lower academic level (Grade 1). The raw scores of the 8 year old pupils did, however, not show a significant increase together with increase in academic level (from Grade 2 to Grade 3). In conjunction with this, it was found that Grade 1 pupils belonging to two different age groups (6 and 7 years old) achieved the same levels of performance. The same was found for Grade 2 pupils (7 and 8 years old) and the Grade 3 pupils (8 and 9 years old) involved in the study. These results, namely that pupils from different age groups but on the same academic level did not show significantly different levels of performance, support the observation that academic exposure plays a major role in the development of receptive language skills in children.

COMPARISON OF PERFORMANCE IN TERMS OF GENDER

Although a significant difference between each age groups' male and female pupils' performance could not be measured, a tendency was noted towards the boys achieving higher raw scores than the girls in each instance (See Figure 3, below). This difference was especially prevalent in the Grade 1 group.

It is known that the African cultures are male-dominated in nature. Therefore, it might be that male children are to a greater extent exposed to language, conversation and learning situations than their female counterparts. This might be a possible explanation for the tendency of boys achieving higher raw scores than girls of the same age.



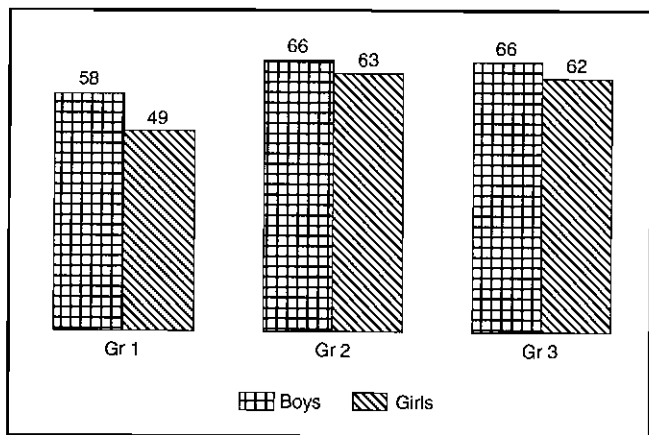


FIGURE 3. Raw scores obtained by each gender in different academic year groups

ERROR ANALYSIS

Keeping in mind that the present research was primarily aimed at the development of a culturally fair version of the PPVT-R, an in-depth error analysis of each response of each respondent was carried out in order to evaluate the validity of the newly adapted test material. From the complete test population's results (experimental phase) 13 items were identified where 50% or more of the pupils' (n=152) responses were incorrect.

Members of the panel of experts were asked to give an opinion regarding possible reasons for the children's poor performance in the identified items, noting that those items might require further refinement or cultural adaptation before this version of the PPVT-R is suitable for implementation in future studies. 70% (n=9) of the 13 identified error-items were not viewed as being problematic or unsuitable for the intended test population. Three (23%) of the 13 error-items' picture stimuli needed further adaptation and only one item (8%) needed further adaptation in terms of translation.

RE-ARRANGEMENT OF TEST ITEMS AND RE-EVALUATION OF PUPILS' PERFORMANCE DURING THE EXPERIMENTAL PHASE OF THE STUDY

During the development of the present Northern Sotho version of the PPVT-R the original order of presentation of the PPVT-R test items (Dunn & Dunn, 1981) were strictly adhered to. The items were originally arranged according to an increasing level of difficulty, but at that time the test was intended for the middle-class American population.

It is, however, possible that the appropriate order of item presentation for Northern Sotho pupils could differ from the original American arrangement. Therefore, as part of the experimental analysis of the results the test items were re-arranged in order of increasing complexity as determined by the number of children responding correctly per item. The items in which the pupils performed best were put first in sequence and the items in which they performed the poorest were put last.

An experimental statistical comparison was carried out between the raw scores obtained from the original item-order and the re-arranged item-order. It was interesting to note that the increments in raw score, between the different age groups were of a more even nature (see even 2

point increments in Figure 4) when the newly ordered data was implemented, while the original American item order led to a fluctuating growth in performance as the pupils' ages increased. In conjunction with this, it was interesting to note that the original PPVT-R (Dunn & Dunn, 1981) also shows even increments of 10 points / items per one year gained in age for the original population. Therefore, it might be meaningful to investigate a culturally fair order of presentation of the items, rather than simply adhering to the original population's intended item-order.

With the newly adapted test material, the intervals between the Northern Sotho age groups at present consist of two items each. Theoretically, it should be possible to insert extra test items on the same level of difficulty as the surrounding items in the Northern Sotho PPVT-R, in order to increase the gaps between the expected mean performance of each age group to a more substantial difference, e.g., five items.

COMPARISON OF THE DATA FROM THE EXPERIMENTAL PHASE OF THE STUDY WITH EXISTING SETS OF NORMS AND DATA.

Since the diagnosis of language impairment can only be made on the basis of an investigation of a child's stronger language which is also usually his first language (Juarez, 1983), it is important to develop norms for the Northern Sotho population's receptive vocabulary.

It was found in previous studies where translated versions of the PPVT-R were compared with the original (American) norms that it was not feasible to merely translate and maintain the norms and the validity of a test such as the PPVT-R (Koekemoer, 1994; Naidoo, 1994). However, seeing that the present study's methodology differed from that of previously mentioned studies, it was decided to compare the present results with the results obtained from three specific populations for which norms or provisional measurements of performance on the PPVT-R were available at that time. The groups with which the present results were compared were:

- Middle-class American population (English-speaking), according to the original norms for the PPVT-R (Dunn & Dunn, 1981);
- Zulu-speaking children from KwaZulu-Natal. Norms contained in Naidoo (1994);
- Northern Sotho pupils from the Pretoria / Mamelodi area. Results abstracted from Koekemoer (1994)

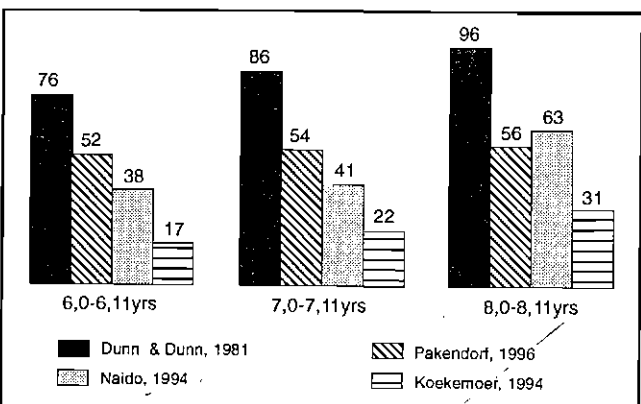


FIGURE 4: Comparison of the mean raw score for each age group involved in this study with the expected mean raw scores of three specific populations



As anticipated, the American (Dunn & Dunn, 1981) raw scores were significantly higher than all three South African groups with which translated versions of the PPVT-R had been applied. It was, however, interesting to note that only the American raw scores and the present version of the PPVT-R showed systematic increase together with increase in the children's age. This is an indication of the value of re-organising the order of test items in a chronological test such as the PPVT-R according to the performance of the population for which the test adaptation is intended.

Most age groups involved in the present study performed on a significantly higher level than that of the Zulu pupils involved in the Naidoo (1994) study as well as the pupils involved in the Koekemoer (1994) study. The present study was carried out within the same target area and with almost completely similar test populations as the Koekemoer study. These results serve as confirmation of the validity of the more stringent refinement process which was followed during the present study. The additional adaptations which were made to the test material, in conjunction with an accurate translation of the word list, therefore had a positive effect on the validity of the test material. It must be mentioned that the PPVT-R test material underwent a more stringent refinement process than in both the previous South African studies on the same subject. Recommendations from these studies were also followed up in most instances in the present study.

Furthermore, the fact that the USA norms were much higher than any of the South African sets of data supports the viewpoint of a number of professional organisations in the USA which called a moratorium on the use of all standardised tests that have not been corrected for test bias when assessing the language of populations which differ from the original intended test population. Specific reference was made to the inappropriacy of the PPVT-R's norms when evaluating a non-mainstream client (Taylor, 1986). The conclusion can be made that the test items (vocabulary) and norms contained in the original PPVT-R do not apply to the levels of cognitive, linguistic, lexical or semantic maturation in South African children whose mother-tongue is an African language.

Another aspect that should be mentioned, and which cannot be measured statistically, is that of the test population's inherent language tradition. As indicated by Holland & Forbes (1986), care should be taken to assist a child from a traditionally orally orientated culture to handle the de-contextual situation which is inherent in a formal language assessment procedure. Therefore the instruction to the research assistants to have a short conversation in Northern Sotho and to explain the test-situation to each child before commencing with the formal evaluation, was especially relevant.

## CONCLUSIONS

It is clear that South Africa with its multi-lingual and diverse population groups literacy in a child's mother tongue is essential (Macdonald, 1990). This poses a unique challenge to research in test adaptation and culturally fair assessment procedures. At present limited examples and guidelines exist which can be followed when undertaking a project of this nature. Therefore, value should be given, not only to the information gained from the experimental phase of this research, but also to the developmental proc-

ess (pre-experimental phase) which preceded the pupil testing phase. The insights which were gained regarding different aspects of translating and modifying an existing Standard American English assessment tool for an African language (and culture) could potentially be utilised over a wide spectrum of similar situations.

Furthermore, the involvement of research assistants with appropriate cultural and linguistic backgrounds can not be over-emphasised in this kind of research. By involving carefully selected research assistants the possibility of sources of conflict in cultural assumptions and communicative norms prior to clinical encounters (Taylor, 1986) can be minimised.

It must, however, also be mentioned that even though it might be possible to derive a culturally applicable version from already existing test material, the fact remains that when (i) translating and omitting items from the word list, (ii) implementing cultural adaptation of the pictorial stimuli, and (iii) re-ordering the sequence of presentation of test items, the original test is being modified to such an extent that it hardly retains any of its original characteristics. In essence, the only remaining characteristic of the test is its main aim - in this case the test is intended to evaluate the child's general receptive vocabulary.

Furthermore, within the current socio-economic climate in South Africa where test adaptation is often cited as a more viable option than that of developing new tests as such, the present study has raised certain issues which have potential for further investigation:

- At present an upsurge is taking place in efforts to uplift the standards of education within the African language communities in South Africa. This implies a great need, not only for assessment procedures, but even more so, for theoretically supported intervention strategies and programmes for the expansion of the Northern Sotho pupil's vocabulary.
- As the study progressed, it became increasingly clear that although the measurement of a child's general level of receptive vocabulary development has a purpose within the academic context, the PPVT-R does not provide adequate direction for clinicians and teachers in the planning of further intervention with the child, and thereby promoting academic development. The test merely determines whether or not a child exhibits a shortfall in receptive vocabulary, when measured in comparison to children of his own age.

This observation is supported in the literature: "The test (PPVT-R) does not provide any theoretically supported suggestions regarding which items should be taught at different stages in a child's intervention program" (Taylor, 1986). Although intensive ethnographically based research around the patterns of language development within a specific cultural group is often inhibited by financial and practical considerations, the knowledge gained during the present study indicated that the development of an original test, which is based upon a developmental model for the intended test population could, after all, be the best option.

In conclusion, all these questions and the process of test adaptation should be seen in the light of creating awareness of not only culturally adapting a test because it is widely used in other cultures and communities, but because it is specifically meaningful for the intended test

population and its circumstances. Only by investing our time and financial resources in a calculated manner can we really make a difference to our children's educational situation.

## REFERENCES

- Alant, E. & Beukes, S.M. (1986). The application of the revised version of the Peabody Picture Vocabulary Test (PPVT-R) to non-mainstream children. *South African Journal of Communication Disorders*, 33:7-12.
- Creswell, J.W. (1994). *Research Design: Qualitative & Quantitative approaches*. California: Sage publications.
- Dunn, L.M. & Dunn, L.M. (1981). *Peabody Picture Vocabulary Test - Revised*. Circle Pines: American Guidance Service.
- Holland, A. & Forbes, M. (1986). Nonstandardized Approaches to Speech and Language Assessment. In Taylor, O.L. (ed.) *Treatment of communication disorders in culturally and linguistically diverse populations*. Boston: College-Hill.
- Jordaan, H. (1989). Management of Language disability in the Bilingual, bicultural child. In *SAALD Conference Proceedings: Facilitating Learning: an ecological perspective*, July:217-234.
- Juarez, M. (1983). Assessment and treatment of minority-language-handicapped children: The role of the monolingual speech-language pathologist. *Topics in Language Disorders*, 3(3):57-66.
- Koekemoer, H. (1994). *Die toepaslikheid van 'n Noord-Sotho vertaling van die PPVT-R by 'n groep skoolgaande kinders*. Unpublished undergraduate research report, Department of Communication pathology, University of Pretoria.
- Labuschagne, D.A. (1990). *The application of the PPVT-R to an urban group of North-Sotho speaking children*. Unpublished undergraduate research report, Department of Communication pathology, University of Pretoria.
- MacDonald, C.A. (1990). *Crossing the threshold into standard three*. The consolidated Main Report of the Threshold Project. Pretoria: Human Sciences Research Council.
- Miller, N. (1984). Language use in bilingual communities. In N. Miller (ed.) *Bilingualism and language disability: Assessment and remediation*. London: Chapman and Hall.
- Morgan, J. & Rinvolucri, M. (1986). *Vocabulary*. Oxford: Oxford University Press.
- Mouton, J. & Marais, H.C. (1990). *Basic concepts in the methodology of the social sciences*. Pretoria: Human Sciences Research Council.
- Naidoo, P. (1994). *Test translation in a South African context using the Peabody Picture Vocabulary Test - Revised*. Unpublished M.A. in Speech and Hearing Therapy, University of Durban-Westville.
- Newmark, P. (1988). *Approaches to translation*. Oxford: Pergamon.
- Pakendorf, C. (1996). *The translation of the Peabody Picture Vocabulary Test-Revised into Northern Sotho: A preliminary investigation*. Unpublished M. in Communication Pathology, Department of Communication Pathology, University of Pretoria.
- Roseberry-McKibbin, C.A. & Eicholtz, G.E. (1994). Serving Children With Limited English Proficiency in the schools: A National Survey. *Language, Speech, and Hearing in Schools*, 25:156-164.
- Schuring, G.K. (1984). *Die Omgangs-Sotho van die Swart Woongebiede van Pretoria*. Unpublished D.Phil dissertation, Rand Afrikaans University.
- Smit, G.J. (1983). *Navorsingsmetodes in die gedragswetenskappe*. Pretoria: Van Schaik.
- Tannen, D. (1982). *Spoken and written language : exploring orality and literacy*. Norwood: Ablex.
- Taylor, O.L. (1986). *Treatment of communication disorders in culturally and linguistically diverse populations*. Boston: College-Hill.
- Zeidner, M. (1986). Are English Language Aptitude Tests Biased towards Culturally Different Minority Groups? *Language Testing*, 3(1):80-98.