UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Exploring how objects used in a Picture Vocabulary Test influence validity

by Ilse de Bruin

Submitted in fulfilment of the
requirements for the degree

Magister Educationis

in

Assessment and Quality Assurance

Department of Science, Mathematics and
Technology in Education

Faculty of Education

University of Pretoria

Supervisor:

Dr Vanessa Scherman,

September 2010

ABSTRACT

Multilingualism in the classroom is one of the many challenges found in the cumbersome bag that the South African education system is carrying over its shoulders at present. Globalisation and migration have added to the burden as factors adding further diversity to the already diverse classroom. In South Africa the spotlight is focused on equality. Equality is expected in the education system, and in the classroom and especially in tests. With 11 official languages excluding the additional languages from foreign learners it has become a daunting task to create tests that are fair across multilingual learners in one classroom. Items in tests that function differently from one group to another can provide biased marks. An investigation was done in order to detect any biased items present in a Picture Vocabulary Test. The study was lead by the main research question being: How do objects used in a Picture Vocabulary Test influence the level of validity? The first sub research question was: How do objects used in a Picture Vocabulary Test influence the level of validity? The next sub question was: To what extent is an undimensional trait measured by a Picture Vocabulary Test? The final subquestion was To what extent do the items in a Picture Vocabulary Test perform the same for the different language groups? This Picture Vocabulary Test was administered to Grade 1 learners in Afrikaans, English or Sepedi speaking schools within Pretoria, Gauteng. The sample totalling 1361 learners. The process involved a statistical procedure known as Rasch analyses. With the help of Rasch a Differential Item Functioning (DIF) analysis was done to investigate whether biased items were present in the test. The aim of this study it is to create greater awareness as to how biased items in tests can be detected and resolved. The results showed that the items in the Picture Vocabulary Test all tested vocabulary. Although items were detected that did indeed perform differently across the three language groups participating in the study.

Keywords: test, bias, Differential Item Functioning (DIF), fairness, multilingual, items, equality, culture, language, visual literacy.

ACKNOWLEDGEMENTS

First and most important I would like to thank my Father who was at my side, holding my hand on this arduous three year journey. Thank you for giving me health, a fit mind and the will to carry on when things were going really tough and it would have been easy to give up. Thank you for your Sabbath days when I know I can rest and recover because You gave the Sabbath especially for resting. Thank You for loving me.

To Christo, my husband, we have gone through deep waters these past three years. Thank you for helping out with the boys when I had to sit in front of the computer. Thanks for the hugs when Word decided to hang for the umpteenth time. Thanks for insisting I have a break and taking me for coffee when you see the statistics are driving me crazy. Lief vir jou.

To my two sons, Pierre and Milan, dankie vir julle liefde. Pierre who always says: “Mamma, jy moet jou Meesters geniet”, when he goes off to school. And Milan who always asks: “Mamma hoekom sit jy net voor die ‘nekenaar’ die heeltyd?” Ek is baie, baie lief vir my twee seuntjies!!!

Thank you to my parents for teaching me perseverance – to finish what I started and having faith in me. Dankie dat Mamma my altyd motiveer en rigting gee waneer ek nie weet watter kant toe met my studies nie.

Thank you to Vanessa. We have also walked a long road together. I feel that as we walked the winding pathway upwards you became just as much a friend as a supervisor. Thanks for the chats, advice and ‘light bulb moments’ and the help in times of need - it still means a lot to me.

I am extremely grateful to the University of Pretoria and the CEA for their financial assistance and the opportunity to complete my M. Thank you so much to my friends at the CEA for their support, upstairs Ronald and Gabriel and downstairs Zaneletjie.

Then to my brand new friends: Yvette and Fires and their daughters Danet and Carah. Baie dankie vir die heerlike aandetes wat julle onverwags ons vir oornooi. Dankie dat julle my ‘n breek gee van die rekenaar. Ons geniet julle wonderlike geselskap en die lekker lag. Maar meeste van als dankie dat jul omgee mense is.

TABLE OF CONTENTS

CHAPTER 6

CHAPTER 7

LIST OF FIGURES

TITLE **PAGE**

LIST OF TABLES

LIST OF ABBREVIATIONS

CEA	Centre for Evaluation & Assessment
CEM	Curriculum, Evaluation and Management Centre
DoE	Department of Education
FAL	First Additional Language
GET	General Education and Training
LEA	Local Education Authorities
LO	Learning Outcome
MidYIS	Middle Years Information System
NQF	National Qualifications Framework
NRF	National Research Foundation
PIPS	Performance Indicators in Primary Schools
PIPSSA	Progress in Primary Schools South Africa
PIRLS	Progress in International Reading Literacy
RSA	Republic of South Africa
SAMP	South African Monitoring in Primary Schools
SASSIS	South African Secondary School Information Systems
SAQA	South African Qualifications Authority
UK	United Kingdom
UNESCO	United Nations Educational, Scientific and Cultural Organization
UP	University of Pretoria
VL	Visual Literacy
YELLIS	Year 11 Information System

GLOSSARY OF TERMS

There are various concepts and terms that will be used and referred to in this paper. These are clarified here, and in greater detail in Chapter 2:

Validity: An assessment is high in validity when the results thereof are accurate and trustworthy. Validity is seen as evaluative judgements that are made from the inferences of assessment results. An assessment has a high level of validity when correct interpretations are made and actions are taken about the results (Bond, 2003; Gregory, 2000; Linn, 1998; Mahoney, 2008; Messick, 1989; Popham, 2003; Frederiksen & Collins, 1989; Gay & Airasian, 2003; Sullivan, Karlsson & Ware, 1995).).

Visual Literacy (VL): For the purposes of this study (there are a multitude of definitions) VL is defined as the ability to accurately identify objects and pictures seen in the past when they reoccur in the present in a similar or different manner (Arbuckle, 2004; Bamford, 2003; McDougall, 2004; Avgerinou & Ericson, 1997).

Picture Literacy (PL): Is the ability to understand and interpret a picture (Arbuckle, 2004; DeLoache, 1991; Rowntree, 1990; Carney & Levin, 2002; Hawthorne & Tomlinson, 1997).

Rasch Analysis: Is a statistical procedure that makes use of scientific analysis to evaluate each item that is used in an assessment. By means of a Rasch analysis each item can be analysed to determine its difficulty, value and relevance to a specific test (Boone & Rogan, 2005; de Beer, 2004; Henson, 1999; Pallant & Tennant, 2007).

CHAPTER 1

INTRODUCTION

Educational Researchers constitute a community of enquirers. Doing the best they can and (at their best) ever alert to improving their efforts, they seek enlightenment or understanding on issues and problems that are of great social significance
– Phillips and Burbules (2000, p.2).

1 INTRODUCTION

Validity is a crucial component of in educational and psychological measurement and addresses the issue of whether a test is a good measure of the trait it is interpreted to assess (Messick, 1981, p. 5). The purpose of this research study was to explore how the construct validity for the Picture Vocabulary Test used to assess Grade 1 learners could be increased, thus resulting in the inferences being accepted as a true reflection of the test used in this research study. (Validity is discussed in detail in Chapter 3). To achieve the purpose the study made use of empirical investigations to answer the research questions. This was done by exploring how each item functioned, not the learners or any other related aspects. Since only the items were explored, the study follows a Positivist viewpoint as discussed in Chapter 4.

In the Republic of South Africa (RSA), a multilingual population is estimated at approximately 47.4 million, comprising African, Coloured, Indian/Asian and Whites (Statistics SA, 2006). There are diverse linguistic groups, with English, Afrikaans, isiNdebele, isiXhosa, isiZulu, Sesotho, Sepedi, Setswana, SiSwati, Tshivenda and Xitsonga being the eleven official languages. Owing to various factors, including globalisation and migration, there may also be additional diversity in many classroom environments, leaving the educator facing an array of unfamiliar languages. Although the teacher will not be expected to understand all eleven languages, an equal

standard of teaching and assessment across various language groups is a prerequisite in any sphere of education. To incorporate these factors often proves to be particularly challenging when different languages have to be accommodated within one school, particularly where language is associated with differing referencing frameworks, beliefs, ideologies, knowledge systems and socio-economic determinants.

Certain groups may have more exposure to certain objects and pictures than others, and it is on this basis that this research preceded. The aim was to explore how objects used in a Picture Vocabulary Test influence the level of construct validity as well as the inferences made from results. The Picture Vocabulary Test, which forms part of a larger assessment instrument of the South African Monitoring in Primary Schools (SAMP) project, is administered to Grade 1 learners as part of the South African Monitoring in Primary Schools (SAMP) project (see Section 1.2.4). The study examines the way objects used in a Picture Vocabulary Test influence the level of validity. Three different pictures are used in the assessment instrument, with each having increasingly difficult objects to identify as the test progresses. The number of objects the learner is expected to identify varies for each picture. This test, which originated in Durham, England, was designed for learners living in that country, and as a result the objects they were asked to identify related to their surroundings and environment. In South Africa, learners from the various language groups were not necessarily as familiar with these objects, so the pictures were redrawn to fit a more localised context after extensive content validation (Archer, Scherman, Coe & Howie, 2010, p. 79).

South African language groups have their own distinctive referencing framework, culture, and historical background (Cohen, 2003; Cincotta, Wisnewski & Engelman, 2000) which influence the way people think, act and behave, including within the educational sphere. Language is influenced by culture and culture by language. Cultural knowledge is transferred from generation to generation by language as found in Schieffelin and Ochs (1986). The relationship between culture and language is discussed in the next chapter (Section 2.4). In a multi-lingual and multi-cultural society, teaching and assessing of learners from diverse backgrounds therefore presents certain challenges.

1.1 VALIDITY OF ASSESSMENTS ACROSS LANGUAGE GROUPS

Complications exist in ensuring that the inferences made from the results of various forms of assessment used to assess a class of diverse learners are true and of a high standard. Since each language group has a specific background from which to build a referencing framework, it is exposed to situations and objects which may be unfamiliar to another. If they are to be valid, inferences based on the results of the Picture Vocabulary Test therefore need to take into account this diversity.

Each language group has certain differences in Visual Literacy (VL) (see Section 2.2), based on experiences. If learners from a specific language group performed badly in a Picture Vocabulary Test it is possible that it is a case of the test being inadequate rather than the learners being weak. For instance, the objects presented in the test itself might be objects to which that specific group has not been exposed, or that the objects used in it are not common to that particular group. Some objects are much more familiar to one group than the next. Therefore, the inference that learners in a particular language group have greater vocabulary knowledge and visual literacy due to a high score on the Picture Vocabulary Test used in this study, or vice versa, needs to be examined critically.

There has to be common ground for all language groups exposed to assessments which include Picture Vocabulary Tests. Pictures should be employed that are recognisable by all groups and are not biased towards any particular one. The inferences relating to the results of a Picture Vocabulary Test have to be sound. This is especially important for young learners who have had very little exposure to the world surrounding them, except their home environment and this may also have limited resources (Howie, Venter, Van Staden, Zimmerman, Long, Scherman & Archer, 2007). Therefore, the inferences made relating to the assessment results of a Picture Vocabulary Test must be grounded in sound empirical reasoning. If the Picture Vocabulary Test results are poor for a certain learner then the deductions or inferences made about the learner's ability must be accurate, because the test has been constructed with a level of validity. This is achieved when the objects used in the Picture Vocabulary Test are familiar to each group.

Taking the above-mentioned into consideration, the aim of this study is to create a Picture Vocabulary Test that has a high level of construct validity for every language group participating in the test. The aim of this study, therefore, is to explore how the level of construct validity of the Picture Vocabulary Test could be increased so that the inferences made are a true reflection. In the next sections, a background will be provided to the research study.

With this chapter an introduction is given to the study. The background, context and reason for the study are explained in Section 1.2. The research questions and objectives of the study are described in Section 1.5, and a brief description of the rest of the content of the thesis in Section 1.6.

1.2 BACKGROUND TO THE STUDY

In the following paragraphs a detailed description is given of how the assessment and later this study came into existence.

1.2.1 Performance indicators in primary schools (PIPS)

There was a need, in England, to measure the progress that Grade 1 learners were making and, as a result, the PIPS assessment was started in 1993. PIPS was used to track the academic performance of learners in their reception year of school. Originally, 7 schools, that were randomly selected, were involved in the project. In 1994, the number of schools grew to 32, and over a thousand learners. In 1997, there were over two thousand schools participating in the PIPS assessment, but as a result of the responses and comments of educators and data analyses, this grew to four thousand schools, that being the number currently making use of the PIPS assessment in England. The PIPS assessment is also used in Scotland, New Zealand, Australia, the Netherlands and Germany. The main purpose of the PIPS projects was to provide feedback to the schools and teachers so that schools could follow self-evaluation. This monitoring system also allows the educator to detect learners who are either poor achievers' academically or those who are excelling academically, and to monitor learners as a whole. (Tymms & Wylde, 2003; Tymms & Coe, 2003; Merrell & Tymms, 2002; Tymms, 1999, 2001, 2004).

1.2.2 The PIPS Assessment

To construct a reliable assessment of which the inferences are valid and which can be administered in 20 minutes by any capable person with minimal training is an admirable task. When the CEM Centre first developed PIPS, several challenges were encountered, one of which was how to assess the learners as they had no previous schooling and only some had attended pre-school. The second challenge was how to test learners younger than seven, in a manner that was valid and reliable. However, a good assessment is usually a long assessment, and the more items involved the more reliable it is likely to be. Therefore, the longer an assessment the better the reliability, at least to a point before the learner loses interest and becomes distracted. The challenge lies in preventing young learners from becoming bored, since they often have a short concentration span and have to be assessed individually to get the most benefits from the assessments. The younger the child the longer he or she takes to respond to a question. As Tymms (2001, p. 23) writes: “On occasions one has to wait an inordinate amount of time to get a simple response”, but the assessment of young children is very important because it helps to identify any problems that may arise, as well as to make an early diagnosis of academic ability.

The PIPS instrument set out to design an assessment that would be reliable and have results not dependent on who had administered it. An assessment had to be created that would be valid and be able to predict future successes or difficulties of young learners. Focus was placed on the internal and external reliability of the PIPS instrument; as examined in more detail in the methodology chapter.

The content of the PIPS assessment was designed after literature on longitudinal quantitative studies about young learners was read. Various sources of literature could provide relevant information about how, through the measure of vocabulary, later reading could be predicted. The progress of children between the ages of two and four was also monitored from primary school until beyond their schooling career. The PIPS assessment is divided into different sections but each is independent of the other and progresses to a more difficult question, known as ‘adaptive assessment.’ In adaptive assessment the learners are presented with easy questions at first and as

they progress the questions become increasingly difficult. The assessment of that specific section is then terminated when it becomes too difficult for the learner and they can then move on to the next section (Merrell & Tymms, 2002; Tymms & Wylde, 2003; Tymms & Coe, 2003; Tymms, 1999, 2001, 2004).

The PIPS assessment sets out to develop a good monitoring system with which to predict future general academic achievement of the learner. This assessment is complated either on a computer or on paper. In this research study, the paper-based version was used due to a lack of adequate resources at some of the schools. PIPS measures the learners' vocabulary, early reading and early mathematics, and the paper-based PIPS assessment consists of an A4 book with instructions. Questions relating to the various sections are printed on the left for the person making the assessment. On the right-hand side are pictures that form part of the assessment. The learner answers the question by either pointing to the correct answer from various options or by saying the answer. The assessor fills in an answer sheet accordingly. A baseline assessment is made at the beginning of the year and an outcome measure at the end of the year, known as a 'follow-up assessment'. The learners are not asked to repeat the sections they answer correctly in the baseline assessment. The gains are then measured to determine if any progress has been made. The assessment is administered individually to the learner and takes approximately 20 minutes to complete. A table depicting the format and content of the test (Merrell & Tymms, 2005b; Tymms, Merrell & Jones, 2004; Tymms, Merrell & Henderson, 1997, 2000; Tymms, 1999; Tymms & Gallacher, 1995).

Table 3.1: PIPS Assessment Content

FORMAT	CONTENT
Handwriting	The learner is asked to write his or her own name and is allocated marks accordingly.
Vocabulary	The learner is asked to identify objects in a picture; there are three different pictures with progressively more difficult objects to be identified by the learner.
Ideas about Reading	The learner is asked questions about the concepts of print, e.g. Can you show me someone who is reading?
Phonological awareness	The learner is asked to repeat words, e.g. *riotous*, and also asked to identify word that rhyme e.g. *cat* rhymes with *mat*.
Letter identification	The learner is asked to identify a series of upper and lower case letters.
Word recognition and reading	The learner is asked to identify certain words that are accompanied by pictures, but the pictures act as distracters
Ideas about Mathematics	The learner's understanding of mathematical concepts is assessed.
Counting and numbers	The learner is asked to count certain objects and numbers.
Sums	Addition and subtraction are assessed with no symbols.
Shape identification	The learner is asked to identify objects that are, for example, taller, shorter, most, least and so on.
Digit identification	The learners are asked to identify the numbers they see.
Mathematical problems	These are sums with symbols that the learner has to work out, including sentence sums.

Due to the focus of this research, more detail will be given to the vocabulary section of the PIPS assessment. Literature read and researched by the CEM centre has shown that the number of letters a learner knows, as well as their phonological skills

at the beginning of the reception year, plays an important role in the prediction of later reading. Based on this useful information, the vocabulary section was designed.

The Picture Vocabulary Test is a sub test that forms part of a larger instrument of the South African Monitoring in Primary Schools (SAMP) project. The subtest evaluates the receptive vocabulary of the Grade 1 learners (Archer et al., 2010). The first picture used in the assessment is a kitchen scene, with the first object the learners are asked to identify being a carrot. From previous analysis, Tymms, Merrell and Jones (2004, p. 676) assert that "... practically every child starting school at the age of four in England, whose first language is English, can point to the carrots on the picture." The authors further indicate that most learners find it difficult to identify the 'yacht' and 'microscope' at the beginning of the reception year. These objects are found in the last picture where the scene is of a room filled with various objects. The assessment progressively moves to a more difficult object to be identified by the learner and continues until it becomes too difficult. The pictures and more discussions of the PIPS vocabulary assessment and the adapted SAMP vocabulary assessment can be found in Chapter 4.

As mentioned above, the PIPS assessment was adapted for use in countries such as Australia, New Zealand and Scotland, but it did come to the CEM Centre's attention that certain items in the vocabulary section of the assessment behaved a little differently in different countries, for example in Australia. It was more difficult for the Australian speaking children to identify with the word 'wasp' and 'pigeon' than it was for children from England. These items that were seen as unusual and biased were dropped from the next stage of analysis (Merrell & Tymms, 2005a). This was significant because if there were difficulties with Australian children with certain objects, the same might also be said for other countries. This study sets out to explore this point further in relation to the South African context, to which the PIPS assessment came by means of the Centre for Evaluation and Assessment (CEA).

1.2.3 The National Research Foundation Value-Added Project

The research for this master's study was born out of a much larger project, namely the National Research Foundation (NRF) Value-Added Project, which is one of many research projects being conducted by the CEA. It came into being in 2003, as the CEA in collaboration with the CEM Centre started a research project funded by the NRF, a national funding body. The reason for the study was to investigate the adaptation of existing monitoring systems designed in the UK to fit a more South African context. Various assessments were developed by the CEM Centre with the aid of teachers and Local Education Authorities (LEAs) in the UK. These assessments provide the teachers with valuable information about the learner's academic abilities (Tymms & Albone, 2002).

The NRF Value-Added project consists of two initial research projects that focus on two specific points in schooling:

1. **The beginning of the first academic year of the primary schools learner's career.** The original Durham instrument for this phase is known as PIPS and has been adapted to become known as South African Monitoring in Primary Schools (SAMP). (See Section 1.2.4 for a full discussion).

2. **The beginning of secondary schooling.** The original Durham instrument for this phase is known as the Middle Year Information System (MidYIS), which has been adapted to become the South African School Information System (SASSIS).

The reason for choosing the first years of Primary and Secondary schooling was because the educators had very little information on what the learners' academic abilities were at this specific stage of their academic careers (Scherman, Archer, Howie & Lopez, 2006). Through the implementation of these monitoring systems, the educators are given information about the learners' future academic performance and are able to identify any areas that need specific attention. This study focuses on

the primary school monitoring system used to assess Grade 1 learners at the beginning of the year and again at the end of the year (SAMP).

1.2.4 South African Monitoring in Primary Schools (SAMP) project

SAMP assesses Grade 1 learners in the following areas: early phonics, early reading, writing and mathematics (Archer, 2006a & 2006b), but it went through many transformations before it became the instrument it is today.

The original PIPS instrument was used to assess Grade 1 learners (Tymms, Merrell & Jones, 2004) but was designed for learners from England (Archer, 2006a & 2006b). The aim of the instrument was not only to determine the learner's current academic abilities but also to predict the learner's future academic performance (Tymms, Merrell & Henderson, 2000). PIPS was adapted to become Performance Indicators for Primary Schools in South Africa (PIPSSA) to fit a more South African context. The PIPSSA instrument was available in English, Afrikaans, IsiZulu and Sepedi. The PIPSSA instrument was a computer-based assessment and was later adapted to become a paper-based assessment. One of the reasons for this shift to paper-based testing was that a very limited number of schools had access to proper functioning computers (DoE, Draft White Paper on e-education, 2003). Other reasons were that a paper-based assessment allows for a lowering of costs, less administrative work and an easier process of adaptation of the instrument to become more culturally fair. The PIPSSA Picture Vocabulary Test made use of objects that were relevant to England but had to be explored for the South African context. After a number of meetings with professionals from educational and psychological backgrounds the objects were identified that needed to be adapted to fit a more South African background. The SAMP instrument was designed to be much more relevant to the South African context, especially the Picture Vocabulary Test.

1.3 PROBLEM STATEMENT

An important factor influencing a child's emergent literacy is the surrounding natural environment in which the child lives and grows up. There are for instance objects, insects, animals and plants that may not be present in another environment. To give a few examples: a kangaroo or koala bear is typical of Australia, a calabash from which mostly African men drink beer is typically African. Such objects, linked to a specific environment, can influence the construct validity of test designed in the UK when being used in other countries. An object that might be common to one country could be foreign to another. In addition, an object that is easily recognised by learners in one country may be difficult for learners in another country. Considering this point, the way the objects are listed in order of difficulty in the PIPS instrument may or may not be applicable to learners in South Africa. This study sets out to explore whether an empirically sound construct is being measured and how the objects should be arranged in order of difficulty to fit a more South African context, thereby increasing the level of construct validity and ensuring that the inferences reflect the construct it was designed to assess.

When consideration is given to the above aspects, the construct validity of the Picture Vocabulary Test needs to be carefully explored as "validation is the empirical evaluation of the meaning and consequence of measurement" (Messick, 1995, p. 747). If learners participating in a Picture Vocabulary Test have not been exposed to the pictures and objects used in the test, they stand the chance of misinterpreting the pictures and the objects. This leads to questioning the inferences made about the results of the Picture Vocabulary Test. Poor results may not necessarily reflect that the learner has poorly developed vocabulary, but rather depend on whether or not the learner has been exposed to the objects, in order for him or her to identify them.

As there are very large numbers of diverse learners in the South African Grade 1 population, it is of utmost importance that a Picture Vocabulary Test accommodates all the learners. Each learner has a different referencing framework and perception of the world and its surroundings, but to argue that all Grade 1 learners have an equal level of visual literacy as evidenced by the Picture Vocabulary Test, a solid

referencing framework and a proficient perception of the surrounding world, would be naïve. These variations influence the way learners react and answer when asked to identify objects in a Picture Vocabulary Test. These factors have to be taken into consideration since they play an important role in how learners achieve in a Picture Vocabulary Test.

This leads to the question of the importance validity plays in the Picture Vocabulary Test of the SAMP instrument (Gay & Airasian, 2003; McMillan & Schumacher, 2006; Popham, 1999). Inferences that are made from the results of an assessment must be valid, and inferences about the results of the number of correct or incorrect objects identified in a Picture Vocabulary Test must also be true. For example, if a learner were to perform exceptionally weakly in a Picture Vocabulary Test, would the inferences that the learner has a poorly developed vocabulary reflect the truth in the SAMP Picture Vocabulary Test? The inferences in this case are that the learner has a poor level of vocabulary relating to the objects he/she was asked to identify in the Picture Vocabulary Test. However, it could be that the items were not constructed in the correct manner from easy to difficult. If this were the case then the inferences made about the specific learner's vocabulary knowledge based on the results of the Picture Vocabulary Test for this specific learner would be incorrect. The South African Qualifications Authority (SAQA) states specifically that assessment should be valid, reliable, fair and practical (SAQA, 2009). This research study sets out to explore the construct validity and inferences made about the results of the Picture Vocabulary Tests and ascertain whether they are in place.

1.4 RATIONALE FOR THE STUDY

Being a developing country, South Africa is faced with challenges such as multilingualism, poor schooling conditions and a limited amount of resources that are presented in all eleven languages (Scherman, Archer & Howie, 2006). As indicated above, although language groups have a few commonalities, clear distinctions are also evident, which leads learners to develop their own unique perceptions and levels of visual literacy and resulting vocabulary. These factors introduce a challenge for assessment that makes use of various objects that need to be identified in a Picture

Vocabulary Test. The research study sets out to determine whether the inferences made about the results of the Picture Vocabulary Test are true, as is the level of construct validity. Suggestions will be made as to how these factors can be addressed to assure correct inferences and a high level of construct validity.

The researcher became interested in this research study after working in an initial research study called PIPSSA, which evolved into SAMP. The first research study, PIPSSA, was developed overseas and used pictures in the picture vocabulary section which were developed according to the culture and traditions of England, with little or no overlap with South African context, as previous discussed. The broader project deals with the feasibility of adapting the English monitoring systems to the South African context. A key concept in the feasibility of the project is validity, which is multi-layered and is seen as a unitary concept consisting of various components but with the primary focus on content validation. When it was decided, after a panel discussion with various professionals, to adapt the study to suit the South African context, the construct validity had to be investigated to see if it was still of a high level. The items that were difficult for the learners in England were easy for the South African children, and visa versa. An example of this is the word ‘padlock’ that was used in the second Picture Vocabulary Test (discussed in the methodology chapter). For the learners from England the padlock was seen as a relatively difficult object to identify, whereas for the learners from South Africa this seemed to be an easy object to identify, perhaps because of a greater level of crime. A reverse situation also occurred when the learners were asked to identify a yacht in the third Picture Vocabulary Test. The yacht was first on the list of objects to be identified, meaning it was easier for the learners to identify than the rest of the objects that followed. For the learners from South Africa the yacht appeared to be one the most difficult items to identify, perhaps because learners are more familiar with the term ‘boat’.

The researcher became curious as to whether the objects in the Picture Vocabulary Test from the adapted SAMP assessment were in the right order of difficulty. The researcher noticed that the learners were identifying certain objects more readily than others that were supposed to be easy. As a result, a need was identified to explore the construct validity of the new SAMP Picture Vocabulary Test, as well as the level

of construct validity with regard to the inferences made about the results of the Picture Vocabulary Test.

This research will be valuable to:

- Researchers taking part in similar studies wanting more accurate results in assessment conducted with various cultural groups within the South African context.
- Instrument development specialists who will test learners from various language groups in South Africa
- Teachers of learners from various language groups who have to be assessed within the South African context.
- Policy makers in developing multilingual assessment policies for the South African context.

1.5 AIMS AND OBJECTIVES OF THE STUDY

The main reason for conducting this study is to explore how the level of construct validity of a Picture Vocabulary Test could be increased so that the inferences made are a true reflection. Furthermore, the study aims to investigate what would be the most effective manner to present the items. The study also intends to recommend possible ways to present the items so that a high level of validity will be maintained across all three language groups participating in the study.

1.6 RESEARCH QUESTIONS

Against the above background, the main research question that guides the study is:

How do objects used in a Picture Vocabulary Test influence the level of validity?

The main research question was broken down into more detailed questions displayed.

1.6.1 Sub Research Questions

The main research question has been broken down into more detailed questions that can lead the research study to explore objective answers.

How do objects used in a Picture Vocabulary Test influence the level of validity?

This question explores literature to identify barriers to the validity level of the Picture Vocabulary Test. Areas that were most applicable to this study were explored. The areas where barriers could be identified were language, culture and Visual Literacy. These are discussed in detail in Chapter 2.

To what extent is an undimensional trait measured by a Picture Vocabulary Test?

This question explores whether objects used in pictures in the Picture Vocabulary Test are measuring a single trait or ability of the learners. The assumption is that all the objects included in the Picture Vocabulary Test measures the trait; in this case the trait is vocabulary. The ability of the learner to identify objects presented in a Picture Vocabulary Test.

To what extent do the items in a Picture Vocabulary Test perform the same for the different language groups?

Once it has been established that an undimensional trait is measured by the Picture Vocabulary Test the items will be explored even further by means of examining whether the items are functioning the same for different language groups.

How can the identified barriers that decrease the level of validity be minimized?

This question aims to provide suggestions as to how the objects that are barriers to the construct validity and the inferences made can be effectively addressed. The suggestions will try to provide insight into the means to increase the construct validity and decrease barriers that are detrimental to validity.

Considering the abovementioned questions this study follows a Positivist paradigm in order that the research questions could be investigated empirically. A Positivists viewpoint allows the focus to be solely on the items used in the Picture Vocabulary Test and not the learners, schools, educators or anything else. This is based on the fact that Positivism sees science as not needing to have a prior sense of the whole to which different parts belong in order for the different parts to be studied (Fischer, 1991).

1.7 CONCLUSION

This chapter gave an outline of the study that takes place, its origin, the research questions that guided it and the methodology used. This study is undertaken from a Positivist viewpoint to provide advice on how to increase the level of validity for the Picture Vocabulary Test. The various items used in the assessment are investigated to determine their fit and difficulty level in the Picture Vocabulary Test. With the improved understanding of how the items perform suggestions and recommendations can be made as to how the level of validity can be increased. This study is guided by the main research question that leads to the information found in subsequent chapters.

This study is divided into seven chapters, each of which has a distinct purpose of leading to the answer of the main research question. A definite line linking the

chapters to the main goal can be seen. **Chapter Two** is a literature review examining various factors that play a role in influencing the learners' ability to identify objects presented in pictures. In this chapter the eclectic definitions of Visual Literacy are narrowed down to what is most appropriate for this study. The close union of language and culture are discussed and the conceptual framework is given of this study. Validity and reliability and their interrelating role are discussed in **Chapter Three.** In **Chapter Four** the methodology of the study is discussed together with the theoretical framework (Positivism) the study followed. The sample of Grade 1 learners speaking Afrikaans, English and Sepedi participated in a Picture Vocabulary Test that has 22 objects ranging from easy to difficult. The study makes use of the Positivist paradigm using statistical procedures with the help of Rasch analyses to explore the data. **Chapter Five** discusses the findings of the data analysed. The entire group of learners' performances are discussed as well as each individual language group. The learners' abilities are matched to the items' difficulties. Items that are not performing as expected are also identified. Reflections on the study are described in **Chapter Six.** Finally in **Chapter Seven** the findings are discussed and recommendations are made regarding improvements that can be made on assessments used across languages and culture and a conclusion is drawn to the study.

CHAPTER 2

LITERATURE REVIEW

2 LITERATURE REVIEW

2.1 INTRODUCTION

South Africa's education system has undergone a number of changes since the first democratic election in 1994, in particular the language curriculum. Significant to this curriculum is visual literacy, in which the identification of pictures plays a crucial role in the development of reading for young learners.

In this chapter a review of the literature relating to the study is provided. It begins with background information to Foundation Phase education in South Africa (Section 2.1), since that is a basis for the discussion of teaching and assessment of literacy, in line with the aim of the study to explore how objects used in a Picture Vocabulary Test influence the level of construct validity, as well as the inferences made. This Picture Vocabulary Test is part the PIPS and adapted SAMP monitoring system, introduced and outlined in the previous chapter. The assessment is administered in the Foundation Phase at the beginning of the year and then again at the end of the year in the home language of English, Afrikaans or Sepedi, and is considered a highly effective tool for the prediction of the future academic performance of learners. However, there are a number of factors that influence the level of validity of monitoring systems, or at least parts thereof, as well as a learner's ability to identify objects presented in the test. A number of topics, such as Visual Literacy (Section 2.2), Pictures (Section 2.3), Language and Culture (Section 2.4), are discussed as they play a vital role in influencing the level of construct validity of the test. The chapter concludes with a discussion of the conceptual framework (Section 2.5) that emerges from the literature reviewed in this chapter. The relationship between each of these aspects of the framework and how they influence the level of construct validity is explored.

2.2 BACKGROUND TO FOUNDATION PHASE EDUCATION IN SOUTH AFRICA

South African education has undergone a dramatic change with the introduction of a curriculum that follows an Outcomes Based approach to learning. As background information, the description of the education is provided, including the three bands of education, making reference to relevant policies such as the Revised National Curriculum Statement (RNCS, 2002c) and the policy on Assessment and Qualifications for Schools in the General Education and Training Band (2001).

There are three bands of education in South Africa, recognised by the National Qualifications Framework (NQF). The band relevant to this study is General Education and Training (GET), made up of three phases beginning at Grade 0 (also known as Grade R) through to Grade 9, and comprising a total of nine years of schooling before the learner is allowed to legally exit the school system. This study is situated in the first phase, the Foundation Phase, which runs from Grade 0 to Grade 3. The South African Schools Act of 1996 makes schooling compulsory from Grade 1 or the age of seven, but it is not compulsory for learners to attend Grade 0 (Education in SA, 2009).

According to the DoE report of 2006 there are 15,676 primary schools in South Africa, with 6 289 530 learners and 190 389 educators, a ratio of 33:1. In the Foundation Phase alone there are a total of 3 807 756 learners, with 52% male and 48.5% female. More than half of the total number of primary school learners is found in the Foundation Phase (DoE, 2006, p.). Most of the Foundation Phase classrooms are filled with culturally diverse learners from various ethnic backgrounds, placing an immense load on educators and persons involved in education to fulfil the educational needs and requirements set out by the DoE and its policies (DoE, 2006).

2.2.1 The Teaching of Literacy at the Foundation Phase

The RNCS gives a broad overview of what is expected of the educator and what the learner is to be taught in the Foundation Phase, as well as understanding the need to develop the " …full potential of each learner as a citizen of a democratic South Africa" (DoE, 2002a). In the development of literacy, emphasis is put on all learners learning their home language or mother tongue for a minimum of 3 years until the end of the GET band (Grade 3), and at least one additional language such as English which may become the Language of Learning and Teaching (LoLT). This means that learners must become competent in an additional language while maintaining their home language.

Since this research study focuses more on vocabulary and literacy skills, a discussion of the development of literacy will follow, with attention being given to the six language outcomes (DoE, 2002a):

Learning Outcome 1: Listening - the learner will be able to listen for information and enjoyment, and respond appropriately and critically to a wide range of situations.

Learning Outcome 2: Speaking - the learner will be able to communicate confidently and effectively in a spoken language in a wide range of situations.

Learning Outcome 3: Reading and Viewing - the learner will be able to read and view for information and enjoyment, and respond critically to the aesthetic, cultural and emotional values in texts. This outcome plays a vital role in a learner's Visual Literacy development (discussed in detail in Section 2.2).

Learning Outcome 4: Writing - the learner will be able to write different kinds of factual and imaginative texts for a wide range of purposes.

Learning Outcome 5: Thinking and Reasoning - the learner will be able to use language to think and reason, as well as to access, process and use information for learning.

Learning Outcome 6: Language Structure and Use - the learner will be able to use the sounds, words and grammar of the language to create and interpret texts.

These outcomes guide the teacher in the teaching and learning of literacy, focusing on speaking, viewing, reading, writing, reasoning and thinking, as well as increasing exposure to pictures and objects (DoE, 2002c). With this type of exposure the level of visual literacy of the learners can be improved upon. The more trained the learners become at achieving these six outcomes, the greater their competency in participating in any form of literacy test.

The early years of schooling play a vital role in developing emergent literacy through exposing young learners to stories (DoE 2002a). In listening to the stories and understanding how they are constructed, language develops naturally and assists learners when they begin to read and write. Emergent literacy is developed in the following ways:

- Seeing signs in the environment and understanding that they signify something;
- Using rhymes that play with language and develop awareness of the separate sounds of the new language (phonemic awareness);
- Trying to read and write in their language, even though their writing may look like scribbles on a page (DoE, 2002a, p. 9).

Teaching Reading in the Early Grades (DoE, 2008) is a handbook that has been designed by the DoE to provide the educator with guidelines to ensure that all children learn to read. It highlights the core elements needed for teaching and reading in the early grades, including the essential knowledge and skills required to help learners read. The time that should be spent on learning and teaching literacy is 1 hour and 50 minutes per day, or a total of 9 hours and 10 minutes per week. The core elements for the teaching of literacy are as follows:

Reading and Writing Focus Time

During this time, basic literacy skills are taught and it is suggested that an hour a day is spent on reading and writing. The activities are shared writing, shared reading, word-level and sentence-level work, guided reading and writing and independent reading and writing activities. With the word-level and sentence-level, special focus is placed on phonics, spelling, vocabulary development, grammar, sentence work and punctuation.

Listening and Speaking

Listening, speaking and writing form part of literacy development, as well as helping a learner to develop thinking and reading skills.

Writing

With the writing activities, learners learn how to form letters, words and numbers. More time is spent on writing in the Foundation Phase than any other phase.

2.2.2 Assessment in the Foundation Phase

The purpose of assessment is to gain information about a learner's strengths and weaknesses in a particular area, leading to decisions based on valid inferences and that should be both challenging and reflect the knowledge and skills of the learners (Vandeyar & Killen, 2003; Killen, 2002).

The implications of assessment are far-reaching in educational settings, with final decisions based on the results affecting a learner's life and academic path (Maree & Fraser, 2004). For this reason it is crucial that valid inferences are made. Assessment is inextricably linked with teaching and learning and Foundation Phase educators are guided by the Policy on Assessment and Qualifications for Schools in the General Education and Training (GET) Band. A few important points highlighted in this policy (DoE, 2001) reveal that assessment should:

- be authentic, continuous, multi-dimensional, varied and balanced;

- take into consideration the diverse needs of learners and the context. Various assessment strategies should therefore be used;
- be used as an ongoing integral part of the learning and teaching process. This means that assessment should be used to inform and evaluate teaching and learning;
- be accurate, objective, valid, fair, manageable and time-efficient;
- take many forms, gather information from several contexts, and include a range of competencies and uses;
- be free from bias and sensitive to gender, race, cultural background and abilities;
- in the main, be criterion-referenced;
- be transparent so that learners and teachers have a clear understanding of the expectations for any assessment task, and what knowledge, skills, values and attitudes are being assessed.

Teachers are involved in assessing learners through an array of strategies both summatively and formatively. Thus, assessment forms the crux of teaching and learning, as stated in the National Curriculum Statement (NCS) (DoE, 2002b), and should be included in all areas of the learning environment and the planning of lessons. The policy document states that with the help of assessment, educators are able to track whether the desired outcomes have been reached, including the minimum achievement level which learners are supposed to reach in that specific grade, and the achievement levels of the learners in accordance with a specific grade. However, the main purpose of assessment is to enhance learner growth and development and to monitor progress.

The SAMP project contains a Picture Vocabulary Test, which forms part of a larger assessment instrument of the SAMP project, and is administered to Grade 1 learners. This Picture Vocabulary Test forms the basis of this research study as it explores how objects used in one influence the level of construct validity and whether the inferences made are valid. In this research study, a baseline assessment is made at the beginning of the year, and a follow-up or summative assessment at the end of the year.

However, within the South African educational context many issues of language arise that have a notable effect on the development of literacy and its subsequent assessment. It is the desire of the government that home language is taught wherever possible, with special relevance to the Foundation Phase. Reading and writing should take place in the learner's home language. It is seen as a barrier to learning if the learner lacks confidence to express him or herself in the language used for teaching, learning and assessment. This barrier is exacerbated if the teaching, learning and assessment are in the learner's second language rather than the home language. If the learner is being assessed in a second language then, according to policy, he or she must be assessed according to the assessment standards of the First Additional Language (FAL) (DoE, 2002b). Sensitivity must be shown to learners with language barriers and necessary steps should be taken to overcome these barriers, as stated in the RNCS (DoE, 2002c). It is also noted in this document that young learners have varying degrees of attention and that the younger the learner the shorter the attention span. This point is highlighted for its relevance to the development of the original assessment used in this research study. One of the challenges of PIPS was to design an assessment that would be short enough to keep the attention of a young learner but also be reliable and have a high level of validity.

The discussion above gives an enlightened perspective of how literacy and assessment, with their challenges, are developed in the early grades. In the paragraphs that follow, factors are identified that affect the way learners perform in the Picture Vocabulary Test used in this study.

2.3 VISUAL LITERACY (VL)

Visual Literacy (VL) is an exceptionally broad and mystifying concept, because unlike the word 'vocabulary', VL is used across numerous disciplines. Each of these disciplines has its own relevant definition of the term VL and in addition, each discipline prescribes different attributes and expectations of it.

VL is referred to variously in the RNCS (DoE, 2002c), whereby the learner is expected to make meaning of and interpret visual texts. The learner has to be able to communicate effectively by making use of different visual modes and must also be able to create, design, discriminate between and interpret visual materials. Furthermore, it is expected that the learner correctly interprets visual images that are combined with text. These aspects are central to the development of VL. In the passages to follow, the attributes of VL most pertinent to this study will be investigated.

The various disciplines that have explored and make use of VL are, inter alia: Psychology, Perceptual physiology, Media studies, Biochemistry, Art History, Sociology, Cultural studies and Educational Technology. Research into the importance of VL in the education of Biochemists, carried out by Schonborn and Anderson (2008; 2006) and Schonborn, Anderson and Grayson (2002), found that it had been ignored for too long when considering that external representations of physical and molecular structures can often be confusing. They further found that VL is not automatically acquired but had to be explicitly taught to students. They concluded that VL is seen as being interdisciplinary and forms part of the modern world. In the Arts discipline, VL is explored to learn how to enable children to become more aware of and how to interpret art (Yenawine, 2003).

Although Biochemistry and Art can be said to be at opposite poles, both these fields see the relevance of the role VL plays in the way images are understood and interpreted. The images used in Biochemistry are highly technical and complicated, made up of molecular and cellular structures. The images need to be correctly identified in order to make meaning of what is seen by the Biochemists, otherwise incorrect diagnoses can be made. On the other hand, it is equally important for children to learn how to interpret art and articulate this in words, so in turn be able to interpret other images and improve their knowledge and language (Yenawine, 2003).

There are as many definitions of VL as there are disciplines using the term. To find a single definition for VL has therefore been problematic and elusive, as noted by Williams (2007), McDougall (2004), Sims, O'Leary, Cook and Butland (2002) and Cassidy and Knowlton (1983). For Raney (1999, p. 1), the term is "like words or like

holy relics?", because a proper definition across disciplines is elusive and the term could be seen as only belonging to the past with no single definition found today. Table 2.1 lists the numerous definitions found in the literature explored. The definitions most applicable to this study are highlighted in green in the table:

Table 3.2: Definitions of Visual Literacy

Definitions of Visual Literacy	Author
A group of vision competencies a person develops by seeing and simultaneously incorporating other sensory experiences	Debes
VL can be defined as a group of skills which enable an individual to understand and use visuals for intentionally communicating with others	Ausburn and Ausburn
VL is the ability to understand (read) and use (write) images and to think and learn in terms of images, i.e. to think visually	Hortin
Visual literacy is what is seen with the eye and what is 'seen' with the mind	Bamford
Visual literacy itself is defined as the active reconstruction of past experiences with incoming visual information to obtain meaning	Sinatra

VL was first identified in 1946 by Dale (as cited in Arbuckle, 2004) as one of the major modes of literacy, with the others being print and audio, but the actual term 'Visual Literacy' was first defined by Debes in the late 1960's, as a group of "visual competencies" a person develops by seeing and simultaneously incorporating other sensory experiences. When these competencies are developed, a visually literate person is able to identify objects and symbols within their environment (McDougall, 2004, p. 56). However, this definition was found to be deficient by Arbuckle (2004),

Bamford (2003) and Avgerinou and Ericson (1997), who felt it was too broad and diffuse for addressing research problems.

Ausburn and Ausburn (cited in Avgerinou & Ericson, 1997, p.281) suggest that: "VL can be defined as a group of skills which enable an individual to understand and use visuals for intentionally communicating with others", and as in this study, it is expected that the learner "will be able to understand and identify visuals (objects) to intentionally communicate effectively the answers required in a Picture Vocabulary Test".

A definition by Hortin (cited in Avgerinou & Ericson, 1997, p.281), slightly more appropriate for this research study, is: "the ability to understand (read) and use (write) images and to think and learn in terms of images, i.e. to think visually". This definition is adapted for the study, namely VL as the ability to understand (identify) and use (name) objects in images and/or pictures, and to think and learn in terms of them. This amounts to thinking visually. The aim of the Picture Vocabulary Test, thus, is for the learner to understand and identify the images or pictures and the objects used in it.

Another definition of VL appropriate to the Picture Vocabulary Test is provided by Bamford (2003, p. 1), as: "what is seen with the eye and what is 'seen' with the mind." Bamford further argues that a visually literate person can "…discriminate and make sense of visual objects and image". Linking this with the expectations of the Picture Vocabulary Test would mean that the learner sees (seen with the eye) objects used in the pictures of the test and correctly identify the vocabulary associated with the picture. The learner then has the ability to recall and identify the objects in these pictures ('seen with the mind'). However, this is based on an assumption that the learner has had previous exposure to the objects used in the Picture Vocabulary Test and can then later recall, identify and name them.

A definition by Sinatra (as cited in Avgerinou & Ericson, 1997, p. 282) of VL as "the active reconstruction of past experiences with incoming visual information to obtain meaning" is the most applicable to this study, as it is expected that the learners participating in the Picture Vocabulary Test are able to combine previous events in

their lives with what they are seeing and make sense of it on a cognitive level The learner should thus be able to find personal significance in an object, in this study identify and make meaning of, for example, a carrot, butterfly or cash, from visual information to which they have previously been exposed. Although Sinatra has not stipulated or explained what is meant by 'visual information', it is here taken to refer to objects, items or situations that have contributed to a learner's VL. The definition of VL most suited for this study therefore would be: the ability to accurately identify objects and pictures seen in the past when they reoccur in the present in a similar or different manner.

The VL competency of a learner can be identified by a number of factors mentioned in the next section.

2.3.1 Identifying Factors of Visual Literacy Competency

The characteristics of VL identified by Johnson (in Arbuckle, 2004, p. 448) are the ability to:

- see the difference between light and dark
- recognise difference in brightness
- distinguish colour from greys
- recognise differences and similarities in colour
- see distance, height and depth
- see movement
- understand simple body language
- recognise a whole shape when parts are covered or hidden
- recognise groups of objects that are commonly seen together (e.g. knife, fork and spoon)
- sequence objects that are not commonly seen together into some kind of meaning
- see similarities and differences in shapes

The above can certainly provide guidance to the competency of a person's VL level but cannot be considered as qualifying factors. They cover a very broad area and no distinction is made between, for example, gender, age groups, physical disabilities or culture. For example, if a person was colour-blind, and depending on the degree of colour blindness, he or she might not be able to distinguish colour from grey nor recognise differences or similarities in colour. It may also be significant in terms of gender that one in twelve males have a degree of colour-blindness, but that it is rare in females (Ridgen, 1999). This does not necessarily mean that colour-blind males are visually illiterate or incompetent, merely they may not all see the same colours. This example is cited as evidence that gender may be relevant in examining the topic, amongst other factors such as age, visual competence, culture, and socio-economic status, all of which have to be considered when gauging a person's competence in VL. The aspects required for this research study are described in greater detail.

2.3.2 Developing Visual Literacy

Burton (2004, p. 3) asserts that there are three factors that describe the process of VL:

1. **Visual Perception,** the way information and objects from the physical world are seen and taken in for meaning to be derived. It could also be the way information and objects from the environment of a person is internalised to make a mental picture of what was seen or experienced.

2. **Visual Imagery,** the way that information and objects or events are processed internally and then recreated "in the mind's eye". This information consists of objects or events recalled from memory by means of past visual experiences of these objects.

3. **Visual Communication,** the way of conveying and receiving visual information or messages using purely visual means, with no use of text.

To these three factors, can be added another, which although not mentioned in the literature referred to in this study is important to include:

4. **Visual Ability,** the way objects and pictures can be represented in different contexts and ways, and yet still be identified and made meaning of by the person viewing them.

A diagrammatic representation of the four factors involved in the process of developing VL. When visual perception, visual imagery, visual communication and visual ability are all equally and substantially developed, a concrete foundation is laid for VL.

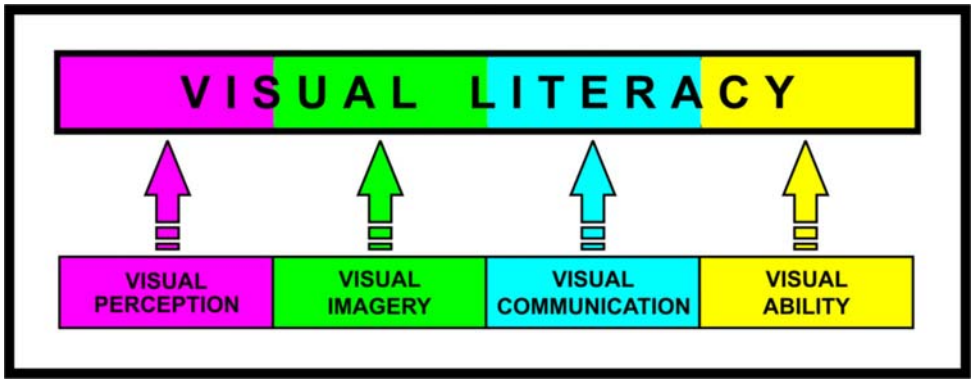


Figure 3.1: Foundation of Visual Literacy

If these four factors are adequately developed, the learners participating in the Picture Vocabulary Test will be able to accurately identify the objects shown to them. If the VL of the learners overlaps with what is presented in the Picture Vocabulary Test the inferences made will be well-founded and the test will have a high level of validity. Objects that are presented in the world around the learners have to be made meaning of, above all to be seen (visual perception). These objects then have to be processed and stored in the learner's mind to be recalled later 'in the mind's eye' (visual imagery). Additionally, objects can be used to communicate without the use of text; an example would be a picture of a hamburger on a map that represents a restaurant (visual communication). Continuing with the example of a map, a number of different objects can be used together to represent a road, a bridge, play parks, and hotels. These various objects are grouped together to make a map, but they can

also be used to depict a story in a storybook. In other words, the same objects are presented in different ways for different reasons, but the learner is still able to understand and make meaning of the objects presented in diverse contexts.

Considering the above factors, an ideal can be reached of what can be expected of learners who participate in a Picture Vocabulary Test. A visually literate learner, thus

- has the ability to see and identify objects in his or her environment
- is able to make meaning of unfamiliar objects seen
- is able to memorise objects seen within the environment. Once meaning has been made of the objects seen, these objects are internalised and stored in memory
- can recall from memory objects that have been seen previously
- can identify the object if it is seen again
- can identify previously seen objects in various contexts
- can name previously seen objects in various contexts
- is able to continually develop his or her VL through exposure to new unfamiliar objects
- has the ability to store and recall new objects that have been seen
- has his or her own unique level of VL, similar to someone having their own unique degree of vocabulary

Some factors tend to be more developed than others, and within time and explicit teaching a learner's VL can be increased.

2.3.3 Furthering the Development of Visual Literacy

In order to overcome the difference in the VL levels of learners, Linney (cited in Arbuckle's study, 2004, p. 453) suggests that three methods be incorporated into the teaching of literacy:

- **Making pictures that people understand.** This is a design-centred response, in which the designer tries to adapt his or her pictures more closely to the type of level of VL in the community.

- **Improving visual literacy skills.** This is a people-centred response. If people get the opportunity to learn pictorial conventions with which they were previously unfamiliar, they will become better able to understand whatever pictures they do get to see.

- **Increasing exposure.** The more pictures people see, the better they become at understanding them, helping them learn how to make their own pictures and using existing visual aids more actively and more often.

Although Arbuckle's study was based on Adult Basic Education and Training (ABET) in VL, these methods can also be adapted and used with the teaching of literacy to Foundation Phase learners as required by the RNCS (2002c). These adapted methods are discussed:

- **Making pictures that children understand.** This is a design-centred response, in which the designer tries to adapt his or her pictures more closely to the type of level of VL of the learner participating in the Picture Vocabulary Test. This is done by creating pictures that specifically relate to young learners, pictures that will stimulate their curiosity, grab their attention and be age-related.

- **Improving visual literacy skills.** This is a learner-centred response. If the learners get the opportunity to learn about objects and pictures with which they were previously unfamiliar, they will be better able to understand whatever pictures and objects they do get to see. This is achieved by incorporating new objects with familiar objects so that the relationship that exists between them can be better understood and to ease identification and grouping. An example would be drawing an unfamiliar vegetable with

vegetables familiar to the learner. The learner would then be able to understand into which group the object fell.

- **Increasing exposure.** The more pictures and objects learners see, the better they become at understanding pictures and identifying objects, resulting in objects being identified more actively and more often. This can be done by having posters with objects belonging to certain categories, or by providing magazines that the learners can page through that follow a certain theme, e.g., crafts, sport, home decoration.

According to Sims et al., (2002) because the number of captured visual images is increasing in an age of technology, successful educational outcomes should be at the forefront, with VL cultivated and taught. VL and visual resources are fundamental for enhanced learning and retention, as proposed by Gardner (2003), the so-called father of Multiple Intelligences (MI).

VL plays a fundamental role in the learners' performance ability when they participate in a Picture Vocabulary Test, and they must have past visual experiences of the various objects used in it and an ability to identify them. They must be able to differentiate, make sense of and identify the objects displayed, although this can only happen if they have had past exposure to them. The problem then arises that such objects must also be applied in such a manner that they are identifiable by all cultures, where possible (Arbuckle, 2004; Burton, 2004; Sims et al., 2002).

The study now turns to ways in which the above-mentioned variables can be controlled in order to increase the validity of the test.

2.4 THE ROLE OF PICTURES IN LITERACY

Pictures have been created for thousands of years, from early cave paintings to contemporary digital images. Pictures can be used to describe events or a concept, send a message or tell a story. A definition by the *Oxford Advanced Learner's Dictionary* (Wehmeier, 2010, p. 1094) of a picture includes painting, drawing, portrait,

illustration, sketch or image. Arbuckle (2004, p. 449) defines the word picture as being "...similar to written words, in that words (written or spoken) in any language are symbols for, or descriptions of, other things – objects, events, feelings or concepts – that exist in reality". These two definitions are clearly linked and reinforced by DeLoache (1991, p. 738), that a picture's "…primary function is a representation of something else."

Pictures may represent objects, events, feelings or concepts that take place in the physical world, and may be descriptive or symbolic, realistic or abstract. They may also be regarded as metaphors for life experiences. Rowntree (1990, p. 121) sheds more light by pointing out that pictures can be used to express something words cannot. A Grade One child can draw a picture of his or her home with smoke coming out of the chimney, or a teenager can draw a picture of his or her home and have the figure 'Death' lurking in the background. Both of these pictures could reflect something that seems to worry the child or teenager, as stated by Di Leo (1983). Different abilities are required to create pictures which can be used to decorate, amuse, express, persuade, illustrate, describe, explain, simplify and quantify (Rowntree, 1990), and to interpret them. Some people instinctively tend to decorate, amuse, express, persuade, illustrate, describe, explain, simplify and quantify that which is most on their mind, but not every person has the same ability.

Some of this variation in ability may be attributable to the environment, as Hawthorne and Tomlinson (1997, p. 301) wrote: "Pictures are most effective when their contents are familiar, realistic and depict a single activity". They can also be used to greatly enhance learning by adding text, observed by Fang (in Carney & Levin, 2002, p. 6) as offering motivation to readers, promoting creativity, servings as mental scaffolding, fostering aesthetic appreciation and thus promoting children's language and literacy. Text combined with pictures is known as 'multi-literacy' or 'multi modal'.

Carney and Levin (2002, p. 6) explain in detail the various forms of pictures that exist, such as *representational*, that depict the text content, *organisational*, which provide a structural framework, *interpretational*, which are seen as clarifiers of scientifically and technically complicated concepts, and *transformational* (*mnemonic*), that assist memory and learning. The aim of the Picture Vocabulary Test is to use

pictures that have objects in them that are familiar to the learners and to which they can relate. However, the problem remains, for pictures to be familiar to learners and for them to relate to them, they first need to be learnt. This means, for researchers such as Osgood, Suci and Tannenbaum (cited in Cassidy & Knowlton, 1983), that the pictures are culturally and socially mediated. When learning about a new object or picture, sensory learning takes place (Whelan, 2004; Avgerinou & Ericson, 1997), and when pictures are used in an appropriate manner, learning can be enhanced, as argued by Carney and Levin (2002).

Carney and Levin (2002), in reviewing studies on pictures-in-text, found that research in the 1970's and 1980's revealed that carefully designed illustrations enhanced learning. These findings were strongly supported by research in the 1990's (Peeck, 1993). More recent research has revealed that various forms of pictures are found in texts that result in robust memory effects (Marley, Levin & Glenberg, 2010). The four forms of pictures named above affect learning, with learners performing at much higher cognitive orders, leading to the conclusion that if a person is picture literate they are able to understand and interpret information presented in a picture (Stokes, 2001). As a result their academic performance will benefit from pictures used in text.

In arguing this point further, Cassidy and Knowlton (1983) attempted to research a child being kept from any pictures from birth, the child being their own child. Despite encountering difficulties in keeping pictures away from the child, especially when they were travelling on roads fringed by billboards and advertisements. They nevertheless were able to draw some conclusions from their efforts. At 19 months of age, the child saw a horse on a television screen and, full of enthusiasm, called it a dog (they had a dog as a pet). When this incident occurred, the research was ended and two external judges were called in to ensure that the study was trustworthy, to check for discrepancies and to test the child. The child was then shown photographs and drawings of a dog by the judges, and was able to identify the dog in both. However, because the child still confused the horse with a dog, steps had to be taken to teach the child what a horse looked like. The study illustrates that if a person is not taught what a specific object is they may name it incorrectly, in this case according to one that has certain visual characteristics in common. There is no innate knowledge of objects with which people are born. The relevance of this conclusion to the Picture

Vocabulary Test is that in order for learners to correctly identify objects they first have to be taught what those objects are.

Thus, when developing a Picture Vocabulary Test, the factors mentioned above have to be taken into account. The learner must be familiar with the way the objects are represented in the picture. The manner in which the objects are depicted in the picture may be more familiar to some learners than to others, for reasons discussed throughout this chapter. Just as each word and sentence conjures up its own meaning for each person, so the same can be said about pictures. Each person has his or her own unique way of interpreting a picture, seeing it in their own way. The person's way of seeing is influenced by previous experiences, presumptions, assumptions, expectations and beliefs. The same happens when a picture or image is recalled from memory, with all these factors influencing how it is recalled. Thus, a picture can be seen as a story, but each person reading that picture has their own style of reading and interpreting the story represented by the picture (Moore & Dwyer, 1994; Weber & Mitchell, 1996).

Understanding and interpreting a picture correctly is a cognitive ability that needs to be developed, and in certain cases, with particular groups, this may require greater attention. In many instances, it is taken for granted that the learners are familiar with the picture and the objects represented (Arbuckle, 2004, p. 445), however, the way a picture is created plays an important role in the performance of learners taking part in a Picture Vocabulary Test. Even though the manner in which the objects in a picture are presented to the learner may be new, the learner must be able to interpret the picture and identify the objects used in it. If the learner can correctly interpret the picture and is familiar with the objects used, a Picture Vocabulary Test will have a high level of validity. Just as pictures and objects need to be learnt by a person – as mentioned above - so too do language and culture.

2.5 THE ROLE OF LANGUAGE AND CULTURE IN A PICTURE VOCABULARY TEST

Language and culture are often seen as individual entities but both influence each other making them inseparable.

2.5.1 The Role of Language in a Picture Vocabulary Test

The elements found in language are sounds, letters, structures, syntax, vocabulary and the way they are put together (Crawford-Lange & Lange, 1987, p. 264). Within the educational sphere, language has become a multifaceted phenomenon that challenges any educator when knowledge has to be put across to diverse learners. Research has shifted from studying children from one specific language group to those from diverse linguistic societies. As well as studying children learning more than one language at a time, studies have even reported that judgements are passed on children with certain dialects (Garcia, 1993).

From 20th century language has been studied intensely by various scientists around the world. Language is what defines humankind and places it above all other species. New facts have been discovered about language development in infants. In the article of (Kuhl, Tsao, Liu, Zhang & de Boer, 2001, p. 145) research has shown that at birth infants prefer the language spoken by their mother to any other language. The baby learns speech patterns while *in utero.* Taking a look at the early development of language in infants interesting findings was documented. Cultural anthropologists have noted that across many world cultures a certain speaking style has been adopted when infants and young children are addressed. This speaking style has amusingly been given the name ‘motherese’ or ‘parentese’ (Kuhl et al, 2001, p. 154). A distinctive acoustic signature is found in ‘motherese’ being high pitched with a slow tempo. This form of speech is used by mothers, fathers, grandparents and caretakers. Adults are unaware of the changes they make when talking to infants and young children and feel embarrassed when questioned about it. Interesting to note though was that ‘motherese’ helped infants learn the language

because of the fact that the speech was modified at a phonetic level. People also increase their pitch when talking to pets, similar to 'motherese' but they do not extend their vowels. The reason for this being that the person realizes that the pet will not be able to talk back while an infant will be able to later. This information above evidently demonstrates that language and culture are closely related and are worldwide perceived as attributes that going hand in hand where humankind is involved.

From the moment a baby is born she or he is not only exposed to language but culture as well. Both language and culture are learnt as the baby develops. An example of this would be a small girl that sees her mother carrying a baby on her back and singing lullabies. The small girl will listen to the lullaby's words and watch the behaviour; she then carries her doll on her back and sings the lullabies to the best of her ability trying to imitate her mother. A further example would be children born within a Jewish culture who will learn about the customs of that culture. These customs will become second nature to them. Thereby not only learning Hebrew, but also how the Sabbath is kept and going to the synagogue aside from other traditional values (Jacobs & Giarelli, 2001). The acquisition of language and culture occurs simultaneously, resulting in language becoming a proxy to culture. This then leads to the next point.

Educators in the foreign languages have come to believe that language and culture are inseparable (Crawford-Lange & Lange, 1987). Teachers and children construct knowledge together by "… drawing on and mingling their varied language and cultural resources and experiences" (Cochran-Smith, 1995, p. 499). Adding to this point, the Sapir-Whorf linguistic relativity hypothesis is based on the theme that culture by means of language influences the way people think. According to this hypothesis there are:" … certain properties of a given language that affect the way people perceive and remember" (Ji, Zhang & Nisbett, 2004, p. 58). The intertwining of culture and language can be used to enhance learning by their influence on values and perception (Shanahan, 1997). Therefore, language is not seen as being interchangeable but as complimentary in the modern world. The reason for this, according to Hobsbawm (1996), is the existence of multinational societies.

Children learn language by imitation and observation, noticing how sentences are expressed by others and how these expressions change from one situation to another. Language is also developed along the lines of social identity, social roles and conversational activities (Schieffelin & Ochs, 1986), such as the way they interact and play with other children in their neighbourhood. A further important aspect of language development is through reading books written in the language of the child (Pretorius, 2009, p. 56). When children start to learn a second language they make use of rote memorisation and imitation. If a child does not frequently interact with persons speaking the language he or she is trying to learn, the transfer of the language will be less successful (Garcia, 1993, p. 63).

As in this study, learners from three different languages participated in the Picture Vocabulary Test (see Chapter 1). These three language groups also have their own related culture and often their schooling took place in their second or even third language. The reason for this is that in RSA there is a shortage of teachers in all 11 official languages, particularly African languages, which Sibula (2007) argues should be developed because they are being spoken less. Language links people to the core values of their heritage, family and community (Mills, 2001, p. 398), and the relationship between language and culture has been documented in research from the early 1970's. Tseng (2002, p. 11) states that culture has a direct relationship to language and is key to learning a language. Research done by Charteris-Black (2002) showed that certain linguistic terms were conceptualised according to the subjects' culture. Manifesting that culture influences how language is understood and interpreted.

In a study by Mills (2001) on bilingualism, which explored children who were born in Britain but whose family's origin was Pakistan; it was found that although they spoke two different languages, each with its own culture, they followed the culture of the language from which their heritage originated. Language is seen as being filled with cultural contexts, and influences the way words are understood. A young Pakistan boy, commenting on him speaking various languages namely English, Punjabi and Urdu noted that in Punjabi and Urdu one word can have ten meanings, and so to correctly understand what is being said numerous factors have to be taken into account. He said that a person must look at the entire sentence, the circumstances

and the person speaking the sentence to understand what is being said. To him, English was just a language (Mills, 2001, p. 396), though a counter-argument would be that these factors are the basics of communication and understanding for anyone, and that there are words in English that can have several different meanings that depend on the context in which they are spoken. For Carter and McCarthy (2004, p. 81) however, language can become utilitarian and transactional, and that this is the case with English, which "has indeed become a utilitarian object for its world-wide users…"

Conducting research into the way culture influences categorization, Ji, Zhang and Nisbett (2004) pointed out that when cross-cultural assessment takes place, most often the original instrument is developed in English. The English instrument is then translated into the other native languages of the participants. The differences that occur between the different languages are then attributed to their differing cultural backgrounds. They concluded that cultural background does indeed affect reasoning and that language affects thinking. This informs the way pictures are affected by language and culture.

Although it is often taken for granted that pictures are seen as being independent of language and culture, despite the message they are communicating (Hoffman, 2000), cultural backgrounds and languages spoken, among other factors, influence the way pictures and objects are seen and identified by people. The purpose of pictures is to intentionally use signs that have been culturally acquired within culturally established patterns to communicate a message (Debes & Williams, 1974). For this research study, the above quote is adapted to: "…the intentional use of culturally acquired *objects* in culturally established patterns for the purpose of *culturally fair* assessment". The word 'sign' may be understood to be some form of warning or information given, as well as being a key word in the study of semiotics (not the focus of this study), and was consequently replaced by the word 'object' for the purpose of this study.

In the next section, culture and its influence on the level of validity will be discussed.

2.5.2 The Role of Culture in a Picture Vocabulary Test

De Witt and Booysen (1995, p. 36) provide numerous definitions of culture, one of these stating simplistically that culture is: "… a system of meaning shared by a population of people and transmitted to future generations". To expand on this definition, De Witt and Booysen (1995, p. 36) identify two types of culture, namely a material culture which embraces objects, technology and art, and an immaterial culture, which has to do with language, knowledge, skills, values, religion and customs. In this study, the focus will be on the latter.

These above-mentioned definitions of culture oversimplify the concept, and as Webb and Read (2000, p. 1) argue, culture is not genetically predisposed from one generation to the next, but rather is an: "…acquired knowledge, learned patterns of behaviour, attitudes, values, expectations, rituals and rules, a sense of identity and of history…". They further advocate that due to different cultural backgrounds, people have different perceived ideas about: "…work, leisure, time, religion, the role of men, women and children in society, sexual practices, food, dress, and so forth". They also point out that differences in culture are portrayed in various ways through dress, music, art and appearance, and that one of the most important instruments of culture is language.

Culture and language influence every aspect of society at every level, such as home, school, education and work, and are an integral part of each human, being a heritage carried with them, be it consciously or subconsciously. As a result, culture and language play a fundamental role in the educational development of the learner (Webb & Read, 2000). Keeping this in mind, it can be seen that an immense challenge exists in developing a Picture Vocabulary Test to assess learners. Each learner has built up a referencing framework and perception of the world based on a cultural background. Consequently, each learner's acquired knowledge cannot be judged as being insufficient or of a lower standard because of exposure or lack of exposure within the cultural background of his or her language. Thus, if an effective tool for the prediction of the future academic performance of learners is to be

developed, learners should be accommodated in completing a Picture Vocabulary Test that will validate their cultural perceptions and language preferences.

The acknowledgement of different cultures and languages is essential within the educational sphere in order for education as well as educational research to advance. There exists a need to be context-sensitive within education, including assessment, as culture forms an integral part of each human. It is equally important for other countries to be made aware of the cultural differences that exist within South Africa, and how these cultures affect education, assessment and the validity of assessment (Crossley, 2000). South Africa is known as the 'Rainbow Nation', a term first used by Archbishop Desmond Tutu and later by Presidents Nelson Mandela and Thabo Mbeki, to encapsulate both the multi-cultural make-up of South Africa and its coherence across a unified spectrum (the 'rainbow'. It has subsequently been used internationally when referring to the country (Habib, 1996)).

Crossley (2000, p. 319) explains that: "Globalisation has infused the ever-present need to learn about each other with an urgency and emphasis like no other in history". But there are critics who argue that, contrary to belief, globalisation has created its own unique culture and that cultural groups are on the verge of extinction within 'the global village'. Nevertheless, it is with this point in mind that Crossley (2000, p. 322) emphasises the need for consideration towards different cultures, particularly as defenders are becoming more adamant about their being uniquely identified and not subsumed by an analogous global culture.

The cultural background of a learner influences the way he or she perceives a picture and is able to identify objects. When a picture is created with a specific culture in mind the objects could either be familiar or strange to a learner from another culture. This depends on whether the objects are found within the learner's cultural setting. The learner's ability to identify an object presented in a picture correctly is therefore affected, which means that when a Picture Vocabulary Test is designed, pictures and the objects used in them need to be thoroughly thought through when considering their use. Learners exposed to pictures that are not within their cultural sphere may experience a certain level of difficulty in interpreting them and identifying the objects (Barnard, 1988; Cassidy & Knowlton 1983; Debes & Williams, 1974).

An example of this is provided by Cassidy and Knowlton (1983), in which a certain culture in a non-industrialised environment could not identify themselves in a photograph, until they were taught about the method of representation of photography. This specific cultural group had never been exposed to photographs or cameras, and as such may be compared to the Khoisan people of Southern Africa who lived off the land, as either hunter or herders. Their culture, prior to Western colonisation, enabled them to survive severe ecological constraints. Depending on natural sources for food and water, they did not come into contact with mass-produced goods until they were exposed to Western culture (Barnard, 1988).

In order for a Picture Vocabulary Test to have a high level of validity, the person designing the test and the person taking it must be of the same culture or be familiar with the culture of the people taking part in the assessment. If a distinct difference exists between the culture of the person designing the assessment and that of the person taking part in the assessment the validity of the test will be in jeopardy, as depicted in Figure 2.2:

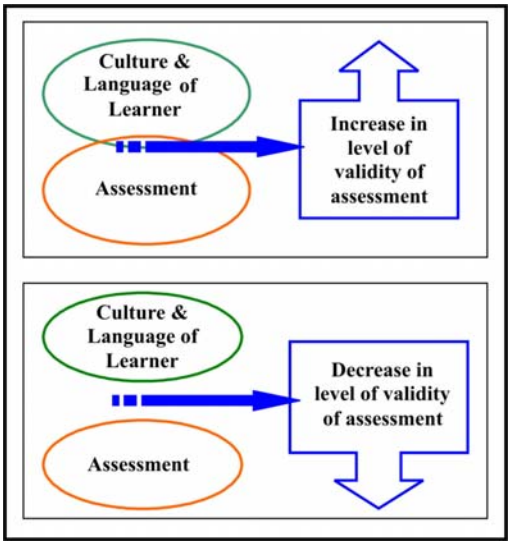


Figure 3.2: Culture & language & assessment influence

A simple hypothetical example intended to illustrate this point would be if a South African teacher of English were to set up a test for Mexican children learning English, in which the question were asked: “What do you wear on your head on a hot sunny

day?" Allowing for the level of language learnt and possible need for translation, the teacher might expect the answer to be 'a hat' or 'a cap', but the Mexican child would likely answer 'a sombrero', a much wider piece of headwear than generally worn by South Africans. Neither answer would be considered incorrect, because a sombrero is, like as hat or cap, generically a piece of headwear. However, because the teacher and learner come from different cultures, with certain objects specific to it, there would be a variation in identifying and naming similar objects.

Taking the above-mentioned into consideration, one can deduce that each culture has its own referencing framework with relation to objects and pictures, which in turn is expressed in language. Each culture has had an exposure to the world seen uniquely through its eyes. What might be common knowledge to one culture could be foreign to another. This places each culture on differing academic pedestals, although it does not imply that some cultures are academically inferior to others, but rather that technologically rich cultures and poor cultures do, to a certain degree, perform differently academically. Technologically rich cultures are surrounded by technology and rely on technology for their daily existence while technologically poor cultures focus more on their natural surroundings. Learners from technologically poor cultures have had a different exposure to the world than their counterparts.

Sternberg (cited in Papalia, Olds & Feldman, 2002, p. 323) asserts that there is a tendency to include questions in assessments that use vocabulary or ask for information and skills which are more familiar or meaningful to certain cultural groups than to others, placing certain learners at a disadvantage. Research reported by Papalia, Olds & Feldman (2002) calculated that Latino and Native American children had lower IQ scores due to language difficulties, performing better in performance tasks than verbal tasks.

Cultural bias also plays a role in the performance of learners when it is presented in assessments, particularly in a Picture Vocabulary Test. Assessments are seen to be culturally biased when they only accommodate a specific culture but are administered to multiple cultures. If the objects used in a Picture Vocabulary Test are selected to suit one specific culture, the test will not have a high level of validity if

administered to other cultures. Comparisons across cultures are becoming increasingly popular in assessments (van der Vijver & Poortinga, 1997).

The SAMP Picture Vocabulary Test has already been adjusted after recommendations made by a panel of experts. This was done in order for the pictures to be more appropriate for a South African context, based on objects found within South Africa. However, the possibility exists that there still could be cultural bias present in the test which could have an effect on learners from various language groups and their ability to identify objects presented in the test. The research done in this research study will indicate whether further adjustments need to be made to make the Picture Vocabulary Test more appropriate and valid for each language group.

Points from the above literature are now considered in terms of their applicability to the development of a conceptual framework appropriate to this study.

2.6 CONCEPTUAL FRAMEWORK

Relevant to this study is Pettersson's (1998) communication model, also cited in Kirsten (2004), which states that in order for effective communication to take place there must be common experiences between the communicating parties: *"…both "sender/encoder" and "receiver/decoder" function within their own "field of experience""* (Kirsten, 2004, p. 19). In the model adapted from Pettersson, the statement changes to say: In order for a Picture Vocabulary Test to have a high level of validity there must be commonalities between the learners' *visual literacy*, *language* and *culture* and the items used in the test. These three factors influence the learners' performance in a Picture Vocabulary Test as well as the validity level of the Picture Vocabulary Test. If any of these three factors do not relate to the items used in a Picture Vocabulary Test then the validity level is in serious jeopardy.

Figure 2.3 is an adapted version of Pettersson's communication model that reflects the factors that influence the validity level of the Picture Vocabulary Test:

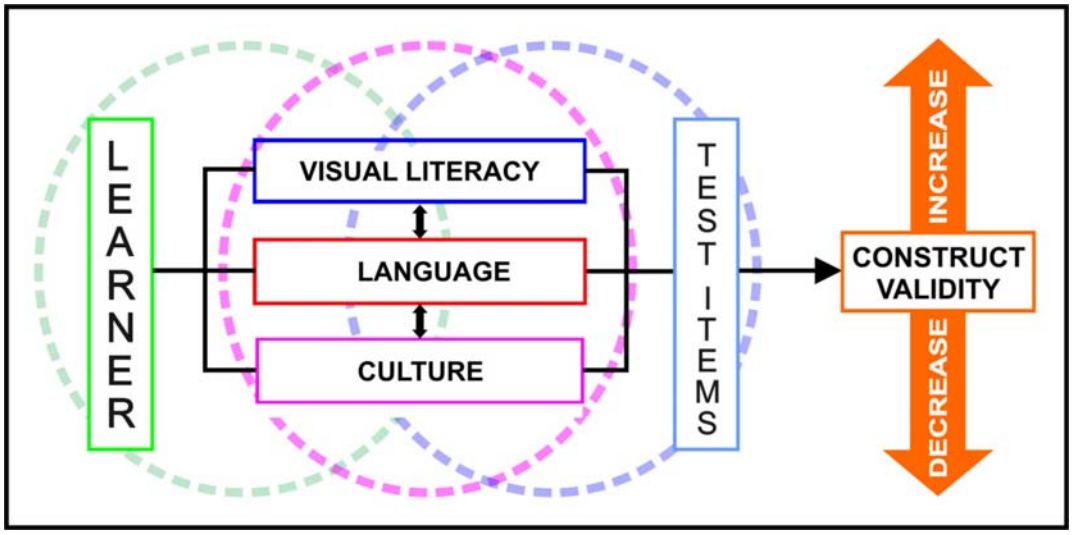


Figure 3.3: Conceptual Framework

Interpretation of Conceptual Framework

The above diagram of the Conceptual Framework can be interpreted as follows: If a clear and definitive overlap can be found between the VL of the learner and the objects in the test, the culture of the learner and the objects in the test, and the language of the learner and the objects in the test, the test will have a high degree of validity.

For a Picture Vocabulary Test to have a high level of validity, it must be designed in such a manner that the learners' 'field of experience' is taken into consideration. In this case, the 'field of experience' of the learners is their level of VL, the culture to which they belong and the language they speak. When considering the design of a Picture Vocabulary Test there must be an overlap of the learner's "field of experience" and the objects used in the test. Pettersson points out various factors that influence the receiver's perception, namely time and stage of development, cultural status and social status (Pettersson, 1998).

The first factor that could influence the level of validity is the learner's VL level. The higher the VL of the learner the greater the chance of the learner identifying the objects presented in the test, as argued above. The learner must have had past experiences with the objects portrayed in the test to be able to successfully recall these objects and identify them. But, if the learners have a low level of VL they will

have difficulty in identifying the objects presented in the test. It is crucial that the VL level of the learner is matched with the objects used in the Picture Vocabulary Test.

The second factor is language. Each language group has its own perceptions and understanding of the surrounding world. Language influences the learners' ability to relate and identify objects because of cultural influences. Since language and culture are intertwined and each language group has its own unique traits, consideration must be given to objects presented in a Picture Vocabulary Test. If these objects do not perform at equal levels of difficulty across the different language groups then the level of validity is threatened.

The third and final factor is the culture of the learner. Cultural influences lead the learner towards a certain view of the world; this influences how pictures are seen and which objects are learnt. Each culture has its own "field of experience" and therefore common ground needs to be identified between the culture of the learner and the objects used in a Picture Vocabulary Test. If the objects used in a test are designed for one culture but the test is administered to other cultures as well, the validity level of the test will drop dramatically. If the objects used in the test correlate with the learner's culture the validity level will be high.

The Picture Vocabulary Test will be successful in incorporating a high level of construct validity and having sound inferences in place. This study aims to provide suggestions as to how to make this possible.

2.7 CONCLUSION

The aim of this study is to provide suggestions on how to increase the level of construct validity for the Picture Vocabulary Test. Three major role players were identified that could have an impact on the level of construct validity. The VL, culture and socio-economic status of a learner have the ability to influence his or her performance in the test. If there is no overlap between these three factors, the objects used in the test and the learner then the construct validity level will be at

stake. The conceptual framework was designed according to the relationship between these aspects.

In this chapter the background to South African education and related policies were given. The three prominent factors that influence construct validity were identified and discussed. The first factor was Visual Literacy if it is not adequately developed the learner will experience difficulty in identifying and relating to objects presented in a Picture Vocabulary Test. The second factor, language, is present in a child's life from the time she or he is formed in the mother's womb. Language creates an understanding of the surrounding world and how it functions. Additionally language influences how objects are perceived and conceptualised. Language forms an integral part of culture, which is the third factor that was identified. Culture influences what a learner is exposed to and what objects are found in his or her surroundings. Language and culture both influence the learners' ability to identify objects in a Picture Vocabulary Test. Finally a conceptual framework was created to depict how all the aspects fit together and influence each other. With the conceptual framework it can be shown how validity can be influenced and what aspects have to be given attention to, to increase the construct validity level. A definite overlap must be created between the language, culture and Visual Literacy of the learner and the objects used in the assessment.

CHAPTER 3

VALIDITY AND RELIABILITY

3 VALIDITY AND RELIABILITY

3.1 INTRODUCTION

In Chapter 2, the study's aims of exploring how objects can influence the level of construct validity of a Picture Vocabulary Test were discussed, and a review conducted of the literature on the various factors that play a role as to how the validity level can be influenced. In this chapter validity and reliability are discussed. The reason why they are discussed in a chapter of their own is to provide a better understanding of what their role was in this study.

3.2 VALIDITY

Although there is much debating going on about validity and an entire thesis can be written about it only the relevant aspect to this study are mentioned and discussed in the sections to follow.

3.2.1 Validity

Bond (2003, p. 179) comments that:

…validity is foremost on the mind of those developing measures and that genuine scientific measurement is foremost in the minds of those who seek valid outcomes from assessment.

From this above quote, validity can be seen as the core of any form of assessment that is trustworthy and accurate (Bond, 2003, p. 179). Validity, according to Messick (1989, p. 6)

> … always refers to the degree to which empirical evidences and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores.

Borsboom, Mellenbergh and van Heerden (2004, p. 1061) have a different take on validity stating that: "…a test is valid for measuring an attribute if (a) the attribute exists and (b) variations in the attribute causally produce variation in the measurement". They do not agree with Messick's conception of validity. In this study Messick's viewpoint will be followed.

Forming the crux of this research project, not only is validity an essential issue for assessment but for measurement as a whole. In addition, the assessments can be used across countries and cultures, but if this is not the case, assessments can be seen as being biased. What is more, validity influences the way that instruction changes once the results of an assessment have been correctly interpreted (Gregory, 2000; Linn, 1998; Mahoney, 2008; Messick, 1989; Popham, 2003; Frederiksen & Collins, 1989; Gay & Airasian, 2003; Sullivan, Karlsson & Ware, 1995).

For instance, an Intelligent Quotient (IQ) test measures the intelligence of the learner (existing attribute), and not all learners will have the same intelligence (variations in the attribute). This implies that when a specific attribute needs to be investigated, the interpretations or inferences made from the test have to be valid. As Ukrainetz and Blomquist (2002, p. 60) put it: "…how well a test measures what it is purported to measure".

Validity is an evolving complex concept because it relates to the inferences regarding assessment results. Focusing on the consequences of the inferences made implies that they should be appropriate and adequate. Messick (1989, p. 6) points out that inferences are hypotheses, and when these inferences are validated it amounts to hypothesis-testing. As a result, validity is seen as evaluative judgements that are made on the inferences of assessment results or test scores, that is whether correct interpretations are made and actions are taken based on the inferences. These evaluative judgements need to be correct and reflective of the truth. An assessment

or test cannot be said to be valid, only the inferences about the test. In Messick's (1989, p. 5) own words: "…what is to be validated is not the test or observation device but the inferences derived from the test scores or other indicators…" An inference can be seen as the interpretation made by a person about a test or assessments results and for this reason it would be incorrect to say that a test is valid since only inferences about the test can be valid or not. In all research, the phenomena being researched must be accurately described through the findings, but if this does not happen then the level of validity is questioned (Gregory, 2000; Mahoney, 2008; Messick, 1989; Graziano & Raulin, 2000).

On the whole, validity is seen as a unitary concept. An example would be if various researchers had to examine one specific research study and also come up with the same conclusion, then the research study would be internally valid. Conversely, with external validity the results and conclusions can be generalised to other situations or with other subjects. Two different types of validity were portrayed in the above example but many other forms of validity exist, making validity a unitary concept (Howell 2002; Opie, 2005; Cohen, Manion & Morrison, 2000; McMillan & Schumacher; 2006). Validity cannot be adequately summarized by a numerical value but rather as a "matter of degree", as stated by Linn and Gronlund (2000, p. 75). The validity of assessment results can be seen as high, medium or low, or ranging from weak to strong (Gregory, 2000).

To summarise, validity refers to the appropriateness of the inferences made about the results of an assessment. Inferences being "…conclusions derived from empirical evidence bearing on score meaning…" (Messick, 1989, p. 6). Secondly, validity is a matter of degree and not a specific value. Thirdly, validity is applied to a specific purpose or use and therefore is not valid for all purposes. Fourthly, validity is seen as a unitary concept, meaning that there are a number of different types of validity. Lastly, validity is concerned with an evaluative judgment about an assessment (Gregory, 2000, p. 75). Of all the different types of validity that exist, construct validity is seen as the most important form. Construct validity forms the basis for any other type of validity and from a scientific point of view is seen as the whole of validity (Mislevy, 2007).

Content-related validity is also another type of validity. As its name implies it explores how the content of the assessment performs. In order to determine content-related validity the researcher is concerned with determining whether all areas or domains are appropriately covered within the assessment. Furthermore, it deals with how the assessment is designed, for example the size of the font, sufficiency of work space for learners, correct language usage and clarity of instructions (Fraenkel & Wallen, 2003). With the original PIPS instrument the content-related validity was carefully considered and of a high standard.

3.2.2 Construct Validity

In 1955, Cronbach and Meehl elaborated on the various methods to determine construct validity. For a construct to be scientifically acceptable it had to be located in a 'nomological network', made up of laws that are either statistical or deterministic. These laws tie observable properties to one another, in other words the same topics or constructs are grouped together (Garrison, 1994; Moss, 1992; Cronbach & Meehl, 1955).

The main purpose of a researcher by exploring construct validity is to determine whether the inferences made about the results of the assessment are meaningful and serve the purpose of the assessment. Construct validity is empirically explored by means of Rasch and, as mentioned above, is central to any quality assessment. Whenever a certain attribute has to be measured, construct validity is involved, as it is the most applicable form of validity to assess measurements (Andrews, 1984; Creswell, 2005; Mahoney, 2008; Messick, 1981, 1989; Popham, 2003; Embretson & Gorin, 2001; Gay & Airasian, 2003; McMillan & Schumacher, 2006).

In particular, construct validity is concerned with the efficacy of a test to gauge learner knowledge about the relevant topics of concern. The test must be relevant, appropriate and utilised correctly, with the focal point being the integration of evidence that produces inferences about assessment results. These inferences must be meaningful, trustworthy and serve the purpose of the assessment for construct validity to reach its goal (Messick, 1989).

If these three steps are integrated into an assessment then the degree of construct validity will be high. In the same way, this study explores whether these three steps are successfully implemented in the Picture Vocabulary Test, as discussed in Chapter 1. On the other hand, if the construct validity of an assessment is not the central focus, it means that the assessment does not assess what it is supposed to, causing the validity level to lower. If an assessment does not produce the same results across different groups then the level of construct validity comes into question.

Cronbach and Meehl (1955, p. 283) view construct validity as asking to what extent a test is culture-free. There are two major threats that exist for construct validity, the first major one being that the construct is underrepresented because it has limited facets of the construct, or too few relevant items to accurately assess the desired topic. The next major threat is 'construct-irrelevant variance', meaning that the test has too much reliable variance, for instance, making certain items easier or harder for certain learners, in such a way that is irrelevant to the construct being measured (Messick, 1989; Moss, 1992).

Linking the above to this study, the focus of construct for the overall SAMP assessment was to determine the relevant level of the academic knowledge and skills possessed by Grade 1 learners across different fields (see Chapter 1, Section 1.2.4). As for this study, the construct of the Picture Vocabulary Test was under investigation, in particular its implementation across three different language groups. When assessing learners using a Picture Vocabulary Test, the objects used in the test must be familiar to all participating learners, implying that most of the objects used in the three pictures of this study must have been observed by the learners at some time or another in their past. For the Picture Vocabulary Test to have a high level of construct validity, all the objects (items) that need to be identified by all learners have to perform the same across all three groups. Alternatively, if this is not the case this study will provide suggestions as to how the level of construct validity can be increased. Meaningful, accurate and justifiable inferences can only be achieved through a high level of construct validity.

To conclude, the level of construct validity plays a vital role regarding the inferences made about the test scores of the learners participating in a Picture Vocabulary Test.

3.3 RELIABILITY

A test is seen as being reliable when it can be used by a number of different researchers under stable conditions, with consistent results and the results not varying. Reliability reflects consistency and replicability over time. Furthermore, reliability is seen as the degree to which a test is free from measurement errors, since the more measurement errors occur the less reliable the test (Fraenkel & Wallen, 2003; McMillan & Schumacher, 2001, 2006; Moss, 1994; Neuman, 2003). In the same way, Maree and Fraser (2004) ask how far the same test would produce the same results if it was administered to the same children under the same conditions. This helps the researcher and educator to make comparisons that are reliable. The more errors found in an assessment the greater its unreliability, and *visa versa*. Reliability is a very important factor in assessment, and is presented as an aspect contributing to validity and not opposed to validity.

Messick (1989) transformed the traditional definition of validity - with reliability in opposition - to reliability becoming unified with validity. Thereby Messick (1989) has accepted a unified concept of validity which includes reliability as one of the types of validity; thus contributing to the overall construct validity. As Messick (1989, p. 8) states:

Hence, construct validity is a sine qua non in the validation not only of test interpretation but also of test use, in the sense that relevance and utility as well as appropriateness of test use depend, or should depend, on score meaning.

Here Messick (1989) explains that not only is construct validity essential for test interpretation but also for test use. The test must be relevant and be able to be utilised in a reliable manner.

With a Rasch analysis, the '*item reliability index*' was examined, reflecting whether the items could be replicated in the same order if they were given to another sample

group that had similar or equal abilities (Bond & Fox, 2001). If items, or in this case the objects order used in the pictures, are not replicable across the three groups, then the reliability of the test is in jeopardy. In this study the real person and real item separation reliabilities were explored. Similar to internal consistency, separation reliability values of between 0 and 1 had to be obtained, and those that are high are beneficial to an assessment (Scherman, 2007).

The items performance can be assessed through Rasch analyses which can alert a person to the ordering of the items. Each item's difficulty is situated along a logit scale, together with its degree of error. The more information available about the difficulty of the item, the more the estimation error decreases. If items are clumped then the difficulty of the items are not equally dispersed, which influences the level of validity as well as the reliability of a test. With the help of Rasch, items that are clumped or too easy or difficult can be identified and dealt with accordingly. If an assessment is focussed only on reliability, the validity level of the assessment will decrease. As was pointed out earlier Messick's (1989) unified concept of validity also includes reliability and overlaps each other. Consequently, a balance has to be reached between reliability and validity.

3.4 CONCLUSION

With this study empirical investigations took place by means of Rasch analyses, in order to determine the level of validity of the Picture Vocabulary Test. Following Messick's (1989) concept of validity and reliability interrelating within an assessment. These empirical investigations that took place lead the study to integrate a Positivist theoretical framework which aided in investigating the items and the level of validity of the assessment distinct from the learners or other factors. Statistical procedures were used as noted earlier and these are discussed in Chapter 4 to follow.

CHAPTER 4

METHODOLOGY

4 METHODOLOGY

4.1 INTRODUCTION

As mentioned before this research study was quantitative, with the research design illustrating the procedures followed and how the data was collected and analysed in order to answer the research questions. The results and findings of this study will be discussed in Chapter 5.

A detailed description of the Rasch analyses that were conducted to analyze the data of the Picture Vocabulary Test is provided in Section 4.1.6. In Section 4.1, the research methodology pertaining to this study is discussed, how the data was collected and the sample selected. The ethical considerations were taken into account in Section 4.2, before the conclusion in Section 4.3.

The main research question that was asked in this study was:

How do objects used in a Picture Vocabulary Test influence the level of validity?

The main research question has been broken down into more detailed questions that can lead the research study to explore objective answers.

What barriers to validity used in a Picture Vocabulary Test can be identified from literature?

Literature was explored to identify barriers most applicable to this study to the validity level of the Picture Vocabulary Test. These areas were language, culture and Visual Literacy as discussed in Chapter 2.

To what extent is a unidimensional trait measured by a Picture Vocabulary Test?

The objects presented in the pictures in the Picture Vocabulary Test are supposed to measure a single trait or ability of the learners. The trait or ability that the Picture Vocabulary Test is supposed to measure is vocabulary. This was investigated to determine whether this was the case.

To what extent do the items in a Picture Vocabulary Test perform the same for the different language groups?

On establishing that a unidimensional trait was indeed measured by the Picture Vocabulary Test the items were explored even further in order to determine whether the items were performing the same across the three language groups. Statistical procedures were used to get to an answer for this question.

How can the identified barriers that decrease the level of validity be minimized?

This question aims to provide suggestions as to how the objects that are barriers to the construct validity and the inferences made can be effectively addressed. The suggestions will try to provide insight into the means to increase the construct validity and decrease barriers that are detrimental to validity.

4.2 RESEARCH METHODOLOGY

The theoretical position within which the design of this research study followed was Positivist as mentioned earlier, which made use of a quantitative methodological approach to determine the actual reality regarding the above questions being researched, and to draw conclusions. A systematic form of measurement took place so that the conclusions or inferences made were objective (Eloff & Ebersohn, 2004) which is typical of a Positivist approach.

4.2.1 Positivism

'*Cogito, ergo sum*' – 'I think, therefore I am', a pronouncement by René Descartes (cited in Phillips & Burbules, 2000), greatly influenced modern philosophy. The

French philosopher, in his famous *Meditations,* wrote that he had accepted false opinions to be true from his youth, and promised to rid himself of the opinions he had adopted. Locking himself in a small room with a fireplace, during the winter, he examined all his beliefs. The false opinions created a desire in Descartes to establish a rigid superstructure of the sciences, but the only concrete, secure belief he could find was 'I think, therefore I am'. Descartes being a Foundationalist and a member of its rationalist division, identified the foundation of Positivism based on "…what could not possibly be rationally doubted and seemed indubitably true should be accepted *as* true" (Phillips & Burbules, 2000, p. 6). In other words, truth and reality do exist and can be measured and explained if found, leaving no room for doubt (Clark, 1998; Guba, 1990). These premises were the foundation upon which Positivism was built.

The Positivist approach has been a recurring theme since Plato, who believed that nature had certain unalterable ideas (Loving, 1997) that needed to be tested and proven true. These seeds of Positivism began sprouting in the 17th and 18th century during the period of Enlightment (Ponterotto, 2005). Francis Bacon (1561-1626) had a sincere commitment towards the rules of evidence, feeling that modern science's purpose was to investigate a nature that was waiting to be discovered, and defined by man through induction. The two notions of induction proposed by Bacon was one of pure discovery and the other a method of observing and then testing hypotheses which lead to logic or justification (Abraham, 1996; Alexander, 2006; Loving, 1997; Phillips & Burbules, 2000). Later, in the 1920's the philosophy of Logical Positivism was developed by the Vienna Circle, a small group of philosophers, physical scientists, social scientists and mathematicians, who focused on the 'Received View' (Abraham, 1996; Phillips & Burbules, 2000), postulating that what is seen, is believed. It was seen as a pointless task to make statements about happenings that could not be verified by the senses. Positivism was developed further by B.F. Skinner, a behaviourist psychologist who also had a major influence on scholars in the Positivist movement (Abraham, 1996; Buchanan, 1998; Kidd, 2002; Lather, 2006; Phillips & Burbules, 2000; Schulze, 2003; Wardlow, 1989). In Table 3.1 the basic constructs of Positivism and how they link with this research study (highlighted in **light green**) are depicted.

Table 4.1: Positivism linked to the research study

(Adapted and combined from Scherman, 2007 & Guba & Lincoln, 1994)

Positivism linked to this research study	
Inquiry Aim	To explain, predict or control
In this study	To explain how the items function in the Picture Vocabulary Test
Method	Quantitative
In this study	A scientific Rasch analysis will make the study quantitative
Logic	Deductive
In this study	If the items do not function as expected then we can deduce that the Construct validity which has implications for interpretations about validity
Epistemology	Objectivist - objectively true
In this study	The researcher and the subjects are independent of each other
Ontology	Realism
In this study	When the items are analysed they show a true reflection of how they function in reality to the learners

The basic constructs of Positivism are linked to this study and further explained.

Inquiry Aim

The aim of the inquiry is to make use of a Rasch analysis to determine how each item functions in the Picture Vocabulary Test. The data will be analyzed and the order of difficulty of the objects as experienced by the learners will be revealed. This will all be shown in an item pathway, a figure in which the items are represented on a vertical axis from the easiest at the base to the most difficult at the top. If certain items do not follow the logical flow according to the Guttman scale they will be identified and an explanation given. On a Guttman scale items are arranged in a

pattern from the easiest to the most difficult (Bond & Fox, 2001; Cavanagh, Romanoski, Giddings, Harris, & Dellar, 2003).

Method

This research study followed a quantitative approach, with a Rasch analyses making use of various scientific formulae to explain certain phenomena occurring in an assessment and the order of difficulty investigated (Bond & Fox, 2001).

Logic

As shown in Table 3.1 (above), the logic is deductive. By using the knowledge received from the data analyses, the objects are arranged in order of difficulty.

Epistemology

By taking an objective viewpoint and observing the outcome of the Picture Vocabulary Test, knowledge is gained about the level of difficulty of the items and if the items are performing the same for the three language groups.

Ontology

Ontology is seen as the reality of a situation. In this study three different language groups of learners from diverse backgrounds are being assessed by an instrument originating from the UK.

4.2.2 Rationale for working with Positivism

With the help of Positivism, empirical investigations can be made to answer questions. Positivism is seen from the perspective that science does not need to have a prior sense of the whole to which different parts belong in order to study the different parts (Fischer, 1991). In this study the items alone will be explored, not the learner or any other related aspects.

A Positivist research approach to the educational sphere makes use of methods that directly investigate the questions asked. The different methods allow for a chain of reasoning that is lucid and rational, and that can be replicated across various fields of

study. The most beneficial factor of Positivism is that it is open to professional inspection and critique (Fischer, 1991).

Through Positivism, science determines to find the truth about how physical, social and personal worlds are configured. This is done by means of empirical testing and evidence in order for the truth to be claimed. Science is seen as universal, and though different methods are incorporated, the same methodology is used. Science may be considered rational if its truths are seen as similar across cultures (Mathews, 2004, p. 23), but such a claim is highly problematic since cultures vary widely in their understanding of even the basic tenets of reason and science. With the help of a Positivist approach the items performance across three languages were explored to create a better understanding of the roles they play in the assessment.

By making use of statistical procedures and empirical testing to determine how items perform in the Picture Vocabulary Test, a Positivist approach is taken. Simultaneously the level of validity of the Picture Vocabulary Test was also explored. With the help of Rasch analyses – a statistical procedure used in this study (discussed in the following chapter) the validity level was investigated. Rasch analyses can specifically be used to determine the level of construct validity of an assessment (Tennant & Conaghan, 2007).

Reliability and validity form the crux of any measurement since they are important in establishing the credibility and truthfulness of the findings. Both reliability and validity are represented in many types and forms and have multiple meanings (Neuman, 2003).

4.3 RESEARCH METHODOLOGY

Quantitative research was used in this study as described in McMillan and Schumacher (2006). The study was carefully planned and conducted in order to enhance the credibility of the results. An existing general assessment from the UK, that was used to assess Grade 1 learners, was explored so that statistical conclusion could be made regarding the data and level of validity. By employing quantitative

methods, measurements are made of each item's difficulty compared to the learner's ability, thus helping to establish whether the items follow the correct order of difficulty as well as to detect any possible bias that will influence the level of construct validity. Table 3.2 summarises the different instruments used and analyses conducted in order to address each research question.

Table 4.2 Research questions, instrument and analysis

Research Question	Instrument	Analysis
What barriers to validity used in a Picture Vocabulary Test can be identified from literature		Literature review
To what extent is a unidimensional trait measured by the Picture Vocabulary Test?	Picture Vocabulary Test	Rasch analysis Developmental pathway
To what extent do the items in the Picture Vocabulary Test function the same for the different language groups	Picture Vocabulary Test	Differential item functioning
How can the identified barriers that decrease the level of validity be minimized?	Picture Vocabulary Test	Reflections on the analysis

4.3.1 The SAMP Sample

SAMP chose the target population of Grade 1 learners speaking Afrikaans, English and Sepedi within Pretoria, Gauteng, South Africa. These languages were selected because they are the most dominant in the Pretoria area and were also the most accessible population for the SAMP project. Multi-phase sampling was used whereby schools were stratified according to medium of instruction. Eight schools were

selected randomly from each medium of instruction from the DoE databases. A sample of 22 schools was selected, including 2 dual medium schools. The sample was inspected to ensure geographic representation of the Pretoria area and found to be satisfactory.

4.3.2 Instrument

The instrument in this study is a Picture Vocabulary Test that was used to assess the sample group of Grade One learners' ability to identify certain objects. A paper-and-pencil test (Gay & Airasian, 2003) was used, and the learners had to identify various objects presented in the Picture Vocabulary Test as pointed out by the fieldworker, who noted the answers on a sheet with a pencil. A paper-and-pencil test makes use of a standard set of questions presented to the learner, requiring cognitive tasks to be completed (McMillan & Schumacher, 2001, p. 189).

The PIPSSA Picture Vocabulary Test, as described in Chapter 1, was the original instrument developed in Durham specifically for UK learners. The pictures were slightly contextualised to accommodate South African learners but the difficulty order of the objects remained the same as presented in the PIPSSA Picture Vocabulary Test. South African learners were thus asked to identify objects in the modified SAMP Picture Vocabulary Test which followed the order of difficulty used in the original PIPSSA Picture Vocabulary Test.

The objects presented in the pictures acted as stimuli for the learners to answer the questions asked by the fieldworkers. There were three different pictures in the Picture Vocabulary Test, each with progressively difficult objects for the learner to identify. The first picture was a kitchen, in which the learners had to identify 7 objects. The second picture was of a bedroom window overlooking a field where 10 objects had to be identified and the third a child's bedroom where 5 objects that had to be identified (see Chapter 4 as well as appendices A, B and C).

4.3.3 Data collection

The SAMP assessment takes place at the beginning and end of the year. The schools participating in the SAMP assessment were visited on various days. The

Picture Vocabulary Test which is a subtest of the SAMP assessment was administered to the Grade One learners on a one-to-one basis by the fieldworkers, who were trained to ensure that the assessment was administered in a standard way to all the learners. The assessment took place in the area designated to the fieldworkers by the participating school over a two day period. The fieldworker fetched each Grade One learner from his or her classroom then followed the correct protocol by setting the learner at ease before the assessment started. The fieldworker then asked the learner to identify various objects from different pictures used in the Picture Vocabulary Test. The learner received a mark of 1 for each correct answer and 0 for an incorrect answer. Each object was worth one mark, and these make up the total number per picture a learner could achieve. The correct answers given by the learner were compared to the total number of objects the learner was asked to identify per picture. The results for the Picture Vocabulary Test were then worked out per learner, and captured electronically.

The responses were marked on an optical reader in pencil by the fieldworkers administering the test. An optical reader is a form that allows the fieldworker to colour in a circle next to the correct response. If the response is incorrect the circle is not coloured in. Once the participating schools had completed the SAMP assessment, the optical readers were sent to independent data capturers. The optical readers were processed through specialised machines. The circles coloured in pencil allowed for the data to be magnetically screened. The lead in the pencil allows for easy recognition by the machines capturing the data. This data received from the optical markers was then sent via e-mail to the CEA to be further processed.

4.3.4 Data analysis

“Data analysis is the vehicle used to generate and validate interpretations, formulate inferences, and draw conclusions”, as stated by Scherman (2007, p. 147). The data analysis for this research study followed a quantitative approach; using Rasch analyses (see Section 3.2.6). By following statistical measures of enquiry, exact measurements can be made to determine whether there are any significant differences in the performance of the items used in the test. The inferences made

about the results of the learners with regard to the Picture Vocabulary Tests can help to determine the level of construct validity.

By making use of Rasch analyses, the functions of the items of the test can be scientifically investigated. Rasch analyses are quantitative in nature because of the attributes of 'additivity' and 'ordinality' (Acton, 2003, p. 902). It has been used over the past 40 years but is being used more as a research tool by many researchers as the adequacy of the instrument and its level of construct validity can be verified (Callingham & Bond, 2006; Rasch, n.d.; Tennant & Conaghan, 2007). Analyses of the data made use of Descriptive Statistics (Section 3.2.5), Rasch Analyses (Section 3.2.6), and Differential Item functioning (Section 3.2.7). These are discussed as follows.

4.3.5 Descriptive statistics

In order to report on the data analyzed, the mean, mode, median, range of scores and minimum and maximum standard deviation were measured. This was done for each object as well as each individual picture. The descriptive statistics facilitated the process of writing about the results of the data that was analyzed (Scherman, 2007). For all the necessary analyses, pathways and graphical representations described in the above sections, a statistical programme, *WINSTEPS* (Section 3.2.8), was used.

4.3.6 Rasch Analysis

In the 1960's, Georg Rasch, a Danish mathematician, introduced a simple logistic model to construct objective measures (Boone & Rogan, 2005). Designed to overcome the problem of defining the difficulty of an item independently of the subject, it also determined the ability of the individual independently of the items. Relevant to this study, the data of the Picture Vocabulary Test can be analyzed independent of the subject (Bond & Fox, 2001; Bush & Schumacker, 1993; Linacre, 1993; McCamey, 2002; Tennant & Conaghan, 2007; Waugh, 1999).

The Rasch model can be applied to analyse dichotomous data and polytomous data (Pallant & Tennant, 2007), and various types of questions or items. Dichotomous data can only be right or wrong, assigned the value of 1 or 0 respectively. There are

also multidimensional Rasch models that deal with more complex forms of assessment, as noted by Rost and Carstensen (2002) and Briggs and Wilson (2003). These forms of assessment, such as the partial credit model and the rating scale model, are not relevant to this study and will therefore not be discussed (de Beer, 2004; Henson, 1999). This study utilized the dichotomous model, also known as the one parameter model or b-parameter model; it deals with the difficulty value of items in an assessment, and focuses on whether the items follow the correct order of difficulty (de Beer, 2004; Dinero & Haertel, 1977).

The learners had to identify various objects. For this study, a correct response was awarded a mark of 1 and an incorrect response 0, known as the '*observed score'* (Fox & Bond, 2001, p. 173). Each question or item became progressively difficult, following the Guttman scale.

4.2.6.1 Unidimensionality

The Rasch model that this study used is a unidimensional measurement model, focusing on one attribute, trait or ability at a time. Thus, items represent only one trait or dominant factor (Henson, 1999). This allows the researcher to develop useful, meaningful and descriptive insight from the analyzed data. The main principle of unidimensionality requires that analytical procedures are to be incorporated to test the degree to which learners participating in the assessment and items fit this idea of a unidimensional line or whether a single trait is being measured (Bond & Fox, 2001). In this study, the measuring of a single trait had already been investigated and implemented when the original PIPS instrument was designed.

By means of a Rasch analysis, evidence can be provided as to whether a particular item over- or under-discriminates, and if any anomalies exist in the ordering of the items. If objects in the Picture Vocabulary Test are too easy or too difficult Rasch analyses will demonstrate exactly how these items performed. Also, if an object is more difficult for one group than another then this item can also be identified. The Rasch model can also provide diagnostic opportunities in which the items can be explored further. This attribute of a Rasch analysis was most beneficial to this study since the order of difficulty of the items were in question (Andrich, de Jong & Sheridan, n.d.). To determine whether anomalies existed with the ordering, 'Fit

Statistics' were used in the study to shed more light on the different objects' difficulties. Bohlig, Fisher, Masters and Bond (1998) argue that misfitting items are not to be thrown out but rather contemplated and explored for possible reasons.

4.3.6.1 Item Characteristic Curve (ICC)

The probability of the learner responding correctly to the item is expressed through an Item Characteristic Curve (ICC). A correct response to an item is dependent on both the learner's ability and the item's characteristics. The ICC is a mathematical function or a visual representation of the learner's ability and the item's characteristics. An ICC has two asymptotes, the upper asymptote is on the vertical axis at 1.0 and the lower asymptote never reaches 0, as seen in Figure 3.1:

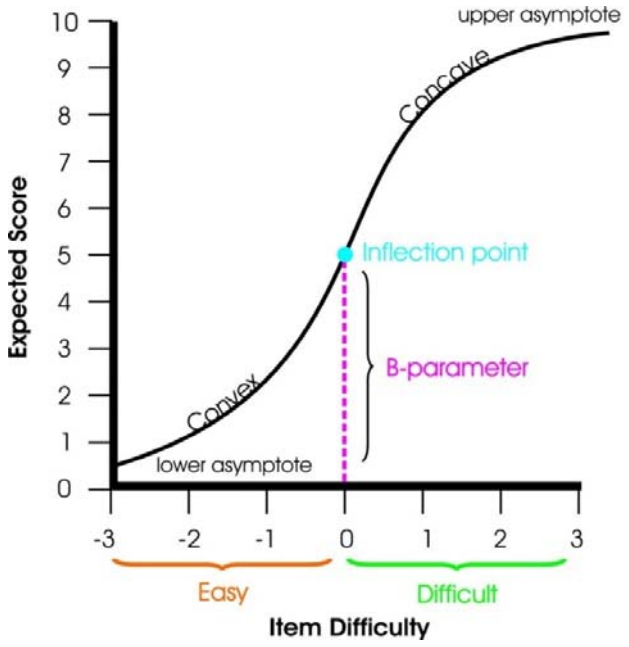


Figure 4.1: Item Characteristic Curve

(de Beer, 2004)

The probability of a correct response to an item by the learner is a continually increasing curve (de Beer, 2004). ICC's differ from one another, with the horizontal location of the inflection of the ability axis shifting more to the right or the left. An inflection occurs when the ICC goes from concave to convex, showing the difficulty level of the item. The ability axis is also known at theta (θ) axis on an ICC. The

horizontal point where the inflection occurs is known as the difficulty level or b-parameter, value or item difficulty (Bond & Fox, 2001). If there are any objects detected that function differently for one group after a DIF analysis has been done, ICC's will be created by Rasch. The ICC's will then show where the inflection occurs and how the objects function.

The b-parameter reflects the point where the ability of the learner to get a correct response is 0.5 or 50%. The greater the value of the b-parameter, the more difficult the item. Theoretically the b-value is from minus infinity to plus infinity (-∞ to +∞), but a value of -2.5 to +2.5 is the typical range. -2.5 indicates a very easy item and +2.5 a very difficult item.

4.3.6.2 The One-Parameter Dichotomous Rasch model

Using the Rasch model that follows a Guttman scale will result in some learners being seen as having more ability than others, and there is a greater probability that the learners with high ability will get the easier items correct. If this is not the case, then the assessment is faulty or has a low level of construct validity (Sick, 2008). Rasch analyses generate separate estimates of each item's difficulty and the learner's ability. These estimates give the researcher a value relative to every individual's ability and every item's difficulty. In other words, a Rasch analysis tells the researcher how the item is functioning relevant to the ability being assessed. It also provides indices to determine if there are items that are spread out or in 'clumps'. The items should move up in difficulty at equal levels and not be grouped on one difficulty level. If this happens in an assessment, the level of construct validity would be in jeopardy since the items do not follow the true Guttman style, each question becoming progressively more difficult (Bond & Fox, 2001).

The Rasch model further provides an opportunity to examine the responses received from the learners to see if they form a pattern that suits the expected outcome. These response patterns from the assessment are tested against what is expected from the specific assessment. The Rasch model is a powerful tool for determining item ordering (Tennant & Conaghan, 2007, p. 1361). In this study, the response patterns

of the learners were compared to the various items' levels of difficulty. This helped determine if the item ordering for the Picture Vocabulary Test was correct.

Rasch also allows the unification of various measurement issues that are required to verify the validity of an assessment. With unification, Rasch measures a single latent trait or ability of a learner and endeavours to specify what occurs when a learner attempts a specific item (Engelhard & Osberg, 1983; Henson, 1999; Wright, 1977). Measuring a single latent trait in this research study involved measuring the learner's ability to identify objects in the pictures used in the Picture Vocabulary Test. If a number or letter had to be added to the Picture Vocabulary Test, a single latent trait could not be measured because now additional abilities or traits were involved (identifying letters and numbers not only objects).

Rasch is useful for reviewing the measurement properties as well as the unidimensionality (see Section 3.2.6) of an assessment. An interval level scale is created by the Rasch model to show the interaction between the learners participating in the assessment and the items used in the assessment (Callingham & Watson, n.d.). Rasch analysis "…provides a complete solution to almost every measurement problem encountered in science" (Wright & Mok, 2004, p. 24).

4.3.6.3 Exploring the data using Rasch

Another reason the Rasch model was chosen was that only a single attribute or latent trait, namely vocabulary, was measured. Each item is expected to contribute meaningfully to the construct being measured, in a hierarchical order from easy to difficult, in the Picture Vocabulary Test. For Picture 1 there were 7 items, Picture 2, 10 items and Picture 3, 5 items, making a total of 22 items. It was important to explore whether the items followed the specifications of hierarchy of item difficulty.

With the help of Rasch analyses, the extent to which the learner performs and the difficulty of the items can be determined along a continuum. Ordinal data is converted to interval data, allowing inferences to be made about the difficulty of the object and to investigate the construct validity (Kyriakides, Kaloyirou & Lindsay, 2006). By fitting the data from the Picture Vocabulary Test to the Rasch model, detailed examination

took place of the level of construct validity, how the items were ordered, the unidimensionality, and whether the items worked the same across all three language groups. An assumption on which the Rasch model is based is that the difference between item difficulty and person ability indicates the probability of a learner being successful with a particular item (Kyriakides et al., 2006). The data of the Picture Vocabulary Test will be explored to see if it follows the Rasch model accurately. A misfit in unidimensionality is indicated by the real standard error. A value of 0.9 or over indicates unidimensionality, while 0.5 and below indicates multidimensionality (Tenant & Pallant, 2006). This was examined to detect any possible misfitting items in the Picture Vocabulary Test.

In the Picture Vocabulary Test a value is given to a response as either 0 (incorrect) or 1 (correct). A 50% chance exists that the learners will get the item correct or incorrect. The probability of a correct response is a logistic function that is determined by the difference in a learner's ability to correctly identify the object and the difficulty in doing so. Items that fit the Rasch model have an item INFIT range of 0.77 to 1.30 (Kyriakides et al., 2006) and a related Z statistic of -2 to +2 (Beaton & Wright, 2005). Any items that do not fall into this range indicate a tendency that they do not follow the expected response pattern of easy or difficult. The response patterns of all three language groups were explored to determine whether any unexpected responses occurred. Furthermore, any items that were experienced as being too difficult or too easy for the learners were identified.

In Rasch analyses the items and the odds ratios are evaluated. Here the odds refer to the probability of successfully answering an item correctly divided by the probability of answering the item incorrectly. The odds ratio is the natural logarithm called natural log-odds, which in turn are referred to as logits (Schumacker, 2004).

Items that do not yield the same results across two or more groups show bias, known as DIF (see Section 3.2.7), which allows comparison of results to be made between various groups (Huang, Church & Katigbak, 1997). The performance of items across the three language groups will be compared in order to detect any items that were not performing in the same way across the groups.

Misfitting items show an unexpected response and an obscured relationship of the probabilities compared to the other items (Lundgren-Nilsson, Grimby, Ring, Tesio, Lawton, Slade, Penta, Tripolski, Biering-Sørensen, Carter, Marincek, Phillips, Simone, & Tennant, 2005). Individual item fit statistics are acceptable within the range ± 3. Any items that are misfits were identified by the Rasch model used in this research study.

4.3.6.4 Fit Statistics

Fit statistics help a researcher detect any discrepancies found between the Rasch Model's expectancy and the actual results of the test that is whether a learner or item's performance is consistent with others (Kyriakides et al., 2006). In order to determine how well the tests data fits the Rasch model's data, chi-square fit statistics were used. With a Rasch analysis, two chi-square ratios are reported: *INFIT (weighted)* and *OUTFIT (unweighted)* Mean Square statistics (Fox & Bond, 2001). Most researchers are more concerned with the INFIT statistic, since it gives more insight into the learner's performance. The learner, whose ability is closer to the item's difficulty, allows for greater understanding about the specific item's performance. *OUTFIT statistics* are concerned with the difference between the expected and observed scores, while with *INFIT statistics* extreme items or persons are detected for targeted items (Tenant & Pallant, 2006, p. 3).

In order for fit statistics to be interpreted, there is a need for experience that is related to that specific measurement context. It is essential to know whether a mean is too large or too small, since each test has its own unique situation. Wright and Linacre (in Fox & Bond, 2001, p. 179) produced a set of general guidelines for researchers, according to whose table a reasonable Item Mean Square range for a multiple choice test is used, from 0.7-1.3 (Bond & Fox, 2001; Tenant & Pallant, 2006).

INFIT is a sum that carries much information, as mentioned above. The statistical information is its "…variance [and] the Standard Deviation (SD) of the estimate squared…" in a Rasch observation (Fox & Bond, 2001, p. 176). To calculate INFIT, each squared standardized residual is weighted by its variance and then added. The total is then divided by the sum of the variances. This produces the same distribution

as the OUTFIT but the differential effects of the weighting can also be seen (Beaton & Wright, 2005).

OUTFIT is the sum of squared standardised residuals, a residual being the difference between the observed score and the expected response. The residual contributes toward misfit in that the greater the residual value, the greater the possibility of misfit. OUTFIT is calculated by squaring each residual, then adding the residuals together and dividing by the number of items to get the mean square (Fox & Bond, 2001).

There are two aspects of fit on which fit statistics focuses. One aspect is the standardised form, known as the *t* statistic, with acceptable values of -2 to +2. The t-statistic is also known as the INFIT *t* and OUTFIT *t.* When the observed data conforms to the Rasch model, the *t* value is near 0 and the SD near 1. If the data is less compatible, the *t* values are greater than +2 or less than -2. When a t-test value has infinite degrees of freedom or the t-statistic has been modified to a unit normal value then ZSTD (standardized as a z-score) is used (Linacre, 2009).

The other aspect is the unstandardised form, known as the 'mean square' or 'average value' of the squared residuals of a specific item. The residual values are the differences between the Rasch model's theoretical expectation of how the item will perform and the actual performance of the item used in the assessment (Tenant & Pallant, 2006). The greater the residuals, the greater the difference between how the item was expected to perform and how it actually performed. All residuals are squared in order to make any minuses into plusses so they can be added in order to give a sum of differences. Therefore INFIT and OUTFIT are always positive, allowing for the mean square fit statistic to be used to monitor the compatibility of the item with the Rasch data (Bond & Fox, 2001). An INFIT mean square of greater than 1 indicates more variation than the Rasch model predicted (underfit). An OUTFIT mean square of less than 1 indicates less variation than was modelled by the Rasch model (overfit).

Item difficulty is estimated from the proportion of learners who succeeded on each item, while person ability is calculated by the proportion of items of which each learner succeeded in. These processes lead to the items being calibrated into logits

and a set of all the learners' measures. For every item used in the assessment, an estimate is given for its difficulty shown in logits.

Any data that was missing in this study was not coded as incorrect because it was interpreted that the learner did not achieve that level of difficulty in the Picture Vocabulary Test. The missing data was not discarded but kept as the missing data can be handled by the WINSTEPS program.

While item difficulty is important item spread along the continuum is also of importance. With the Rasch model reliability indices are generated for both persons and items. The person reliability index indicates the replicability of the order of the persons that could be expected if the sample of learners were given a parallel test measuring the same construct. For person reliability to be generated ability estimates and well targeted items are needed but also a large enough spread of ability across the sample so that a hierarchy of abilities can be measured. Low person reliability indicates that more data has to be collected to reduce error of the estimates. The item reliability index indicates the replicability of items and the placement of items if given to another sample. High item reliability indicates there are items that are more difficult and some that are easier (Bond & Fox, 2001).

4.3.6.5 Item-Learner Map

An Item-Learner Map is a pathway used to represent the development between the items and the learners. It has basic information fundamental to Rasch measurement (Bond & Fox, 2001). In an Item-Learner Map a vertical line can be seen that has X's on the left side, representing the number of learners and their abilities along the variable (y-axis). On the right hand side of the vertical line the objects and their difficulty measures are displayed. Ideally there should be an even spread of items along the variable with no large gaps or clumps. The items should also be lined up with the learners' abilities. The variable (y-axis) measure starts from a minus (easy) value at the bottom to a plus (difficult) value at the top. At the bottom of the Item-Learner Map a '#' is shown with a number. For example '# is 10' if there are 3 next to each other - '# # #' - that means there are 30 learners at that specific measurement of the variable (Linacre, 2009).

4.3.6.6 Item Pathway

An item pathway is a diagrammatical representation of where the items lie along the unidimensional line according to the responses from the learners and the difficulty of the items in an assessment. The learner's abilities are also usually represented on the pathway so that an easy diagnosis can be made of the learner's ability and the item's difficulty (Bond & Fox, 2001).

The location of the items on an item pathway allows for a better understanding of how they function (Bond & Fox, 2001). In this study, the focus is more on the items and therefore the Developmental Pathway in Bond and Fox (2001, p. 22) has been adapted to include only items, not learners, as can be seen in Figure 4.2:

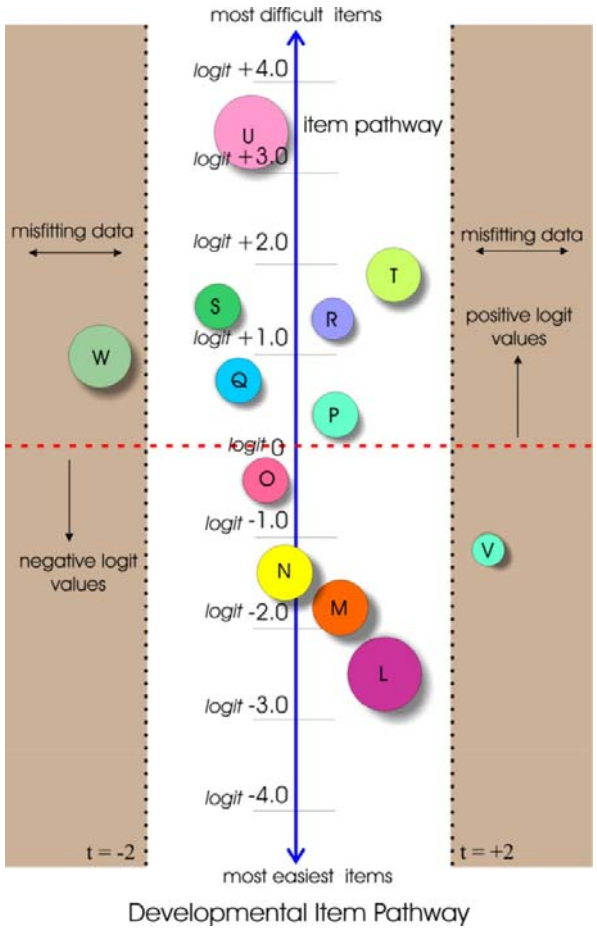
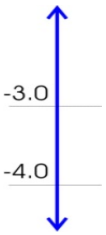


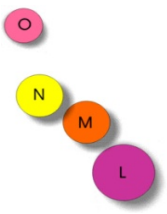
Figure 4.2: Item Development Pathway

(Adapted from Fox and Bond (2001)

The blue line (right) in the Figure 4.2 is the centre and represents the ideal that carries on to infinity and is the unidimensional line along which the items are situated. Interval scales or the dividing segments 'represent the levels of difficulty of the items which are subjected to a log transformation represented along a *'logit scale'*. Each logit unit has a consistent value. The mean of item difficulty is 0, with any item above 0 seen as being difficult and any item with minus (-) as easy (Bond & Fox, 2001, p. 33).



The multi-coloured buttons, or 'stepping-stones', (Bond & Fox, 2001, p. 21) represent all the various items and persons. The pathway at the lower end is typically easier items, such as items L, M and N. Those items at the top are typically more difficult, such as items S, T and U. The distance between the locations of the buttons represents the level of difficulty between the items. The 'greater' the distance to the next button the greater level of difficulty from one item to the next. Looking at the IDP, item N is quite a step away from item O. In an ideal test, the buttons should the same distance apart indicating that each item moves up in difficulty at an equal level. The buttons are also situated at different distances from the blue line, the closer to the line the closer to the ideal of what is meant to be measured in the assessment, while the ones further away are slightly less ideal for what is being measured. As long as they are close enough to the blue line and fall in the pathway, then the assessment is well targeted for what it is meant to be measuring. In addition, some of the buttons are larger than others as the difficulty of the item in the test is located at a specific point and has a 'zone of error' or degree of error associated with it. The smaller the buttons, the smaller the error, which helps the researcher to be more accurate in determining the level of difficulty of that specific item. With the larger buttons, the 'zone of error' is larger, therefore the item's difficulty cannot be determined so accurately and they are not located as precisely (Bond & Fox, 2001).



The vertical dotted black lines on the left and right of the IDP and shaded in brown represents the edges of the pathway. Buttons that fall within these two boundary lines, or the white area, are seen as items that are useful, but items that fall in the light brown shaded areas are seen as problematic, such as items V and W. These items cannot

be interpreted meaningfully in relation to the other items as the possibility exists that they may be measuring another trait (Bond & Fox, 2001).

The dotted horizontal line reflects the point where the items cross over from easy to difficult.

4.3.7 Differential Item Functioning (DIF)

A test has a high level of construct validity when the items perform the same way across different groups. In order to determine this, various methods can be incorporated to test for Differential Item Functioning (DIF) (Mahoney, 2008), the most appropriate for this study being a Rasch analysis.

Within assessments across gender, language and culture it is expected that the items function invariantly, provided that all the learners have the same amount of knowledge for a given subject (Badia, Prieto & Linacre, 2002; O'Neill & McPeek, 1989). This is not always the case and for this reason DIF is a vital source to help identify bias in assessments across dissimilar groups, thereby helping to improve upon the items found in an assessment that displays bias. Items that give different success rates across two or more groups display DIF (Huang, Church & Katigbak, 1997; Tennant & Pallant, 2007).

Numerous studies have been done on DIF, especially with regard to the test performance of different groups, however only in the last decade has the focus moved more towards the differences in test performance among ethnic groups as opposed to sexes. These differences among groups have been extensively reviewed by Green, Crone and Folk (1989), Kim, Cohen and Park (1995) and Wang and Wilson (2005).

When items do not perform in the same way across different groups that have the same abilities or traits DIF occurs, which means that there is a difference in the statistical properties of items. These items are said to 'operate invariantly' (Andrich, 2004, p. 3). DIF is also known as 'bias' (Andrich & Hagquist, 2004; Maller, 2001). A definition found in Smith and Smith (2004, p. 391) regarding bias is said to be the

"differential validity of a given interpretation of a test score for any definable, relevant subgroup of test takers." Also in Smith and Smith (2004, p. 392) a more comprehensive definition is provided of bias as being the "significant and persistent interaction between some (but not all) persons and some (but not all) items." Mahoney (2008, p. 15) elaborates: "Bias…creates a distortion in test results for members of a particular group". With the help of DIF analysis, a statistical procedure, items that may have different meanings for different groups can be investigated, which is often overlooked by conventional processes for reviewing items (Freedle & Kostin, 1990; Scheuneman & Gerritz, 1990).

Through DIF analysis, the researcher is able to monitor whether the level of validity and fairness of the assessment is jeopardised by biased items. The ideal for valid quantitative judgments to take place is for all items to perform the same way across different groups with the same knowledge (Badia, Prieto, & Linacre, 2002). Likewise the construct validity level of an assessment is threatened when items exhibit DIF.

DIF analysis further helps to create a better understanding of the difficulty of an item and the characteristics of the group participating in the assessment, indicating the group's relevant strengths and weaknesses (Hagquist & Andrich, 2004). The reasons some items appear to be biased can be attributed to factors such as ethnicity, exposure to various resources, differing opportunities, background, education, culture, language and life experiences (Green, Crone & Folk, 1989; Maller, 2001; Scheuneman & Gerritz, 1990; Zwick, Donoghue & Grima, 1993; Zwick & Ercikan, 1989).

Many educators discard items that appear to be exceedingly biased in favour of investigating the possible reasons. Three possible factors within a broad environment that can contribute towards bias have been identified by O'Neill and McPeek (1989, p. 256):

1. Surface features or content characteristics of the question
2. Real differences in the groups' knowledge and skills (such as those resulting from different educational experiences)
3. The nature of the criterion used for matching.

These abovementioned factors evidently show that items with a high DIF value are not unfair items but rather items that perform differently across diverse groups with matched knowledge.

DIF can be graphically represented with the help of statistics programmes, such as *WINSTEPS*.

4.3.8 WINSTEPS

WINSTEPS is a programme used to analyse data, developed by people who, on a daily basis, were involved with analyses in the work environment. In the area of educational research, it is helpful with the many applications of the Rasch model, and was designed to "…construct measurement from the responses of a set of persons to a set of items" (Linacre, 2009, p. 29).

There are a number of advantages of using *WINSTEPS* (Bond & Fox, 2001; Linacre, 2009; Scherman, 2007):

- Letters as well as integers can be used
- Easily used with other programs such as Statistical Package for Social Sciences (SPSS) and *EXCEL*
- Data can be analysed from dichotomous, multiple-choice, rating scale or partial credit items
- Missing data can be included in the analysis
- Items and learners are analysed in depth.

With *WINSTEPS*, diagnostic procedures are used to provide information on outliers, unexpected data points and whether the test is not unidimensional. Items and the response structure are calibrated and a central estimate for each learner calculated. These are represented in the form of graphs, plots and tables. For the learner and the items measured, standard error, fit statistics and reports on item or person responses that cause the misfit are also included in the output.

One of the limitations of *WINSTEPS* is that it cannot calculate two or more parameter models, as it was designed specifically for the one-parameter model (Scherman, 2007). However, for the purposes of this study, *WINSTEPS* is ideal.

4.4 ETHICAL CONSIDERATIONS

For the purpose of this study permission by the DOE was attained when the larger SAMP research project went into the field. Clearance was obtained for the project as well as developments from it. Additional letters were submitted to the Ethics Committee for changes that took place. A letter requesting permission to conduct the study was sent to the schools as well as a consent letter to the parents wherein the project was explained, the CEA's contact numbers were provided and the benefits and/or risks of participation stated. The learner was allowed to withdraw or refuse to take part in the study at any time. The letter also ensured confidentiality and anonymity.

4.5 CONCLUSION

In conclusion, this chapter gave a methodological overview of how this research study explored the research questions. By means of this study, the construct validity of the Picture Vocabulary Test was explored, focussing on how the objects represented in the various pictures perform across the three different groups partaking in the study. Since the methodology was based on statistical procedures to explore the answers to the research questions, a Positivist viewpoint was taken, resulting in this research study incorporating Positivism as a research paradigm. The items were investigated apart from the learners or other related aspects, which is a Positivist belief i.e. parts can be studied apart from the whole (Fischer, 1991). The chain of reasoning was lucid and rational and can be inspected by other researchers and they will come to the same conclusion as this study. The Positivist viewpoint of the study lead to the research questions being empirically investigated to reach suitable answers.

The main research question followed an exploratory angle, making the study quantitative. The data was collected at the beginning of the year. The sample

consisted of learners from Afrikaans, English and Sepedi speaking schools in Pretoria. The learners participated in a Picture Vocabulary Test that consisted of 22 objects that they had to identify. The Picture Vocabulary Test is a sub-test that forms part of a larger instrument that was originally from the UK. These objects were arranged from easy to difficult for the learners in the UK. Since learners from different language groups are being assessed with the same instrument the most important and relevant to this study, is the issue of validity. How validity is influenced together with reliability. This chapter concluded with a brief discussion of the ethical considerations for this study. The focus turns to the results of the study, with the data that was analysed and discussed in the chapter to follow.

CHAPTER 5

DATA ANALYSIS

5 DATA ANALYSIS

5.1 INTRODUCTION

This study followed a quantitative approach. As previously mentioned in Chapter 3 (Section 3.1) the focus was on the data received from the learners participating in the Picture Vocabulary Test. The data were statistically analyzed in order to answer the main research question as well as the sub-research questions:

How do objects used in a Picture Vocabulary Test, influence the level of validity?
The objects used in the Picture Vocabulary Test, were investigated in order to get greater insight on how they performed in general as well as across the three different language groups.

These are sub-questions:

What barriers to validity used in a Picture Vocabulary Test can be identified from literature?
From the literature, factors were identified for example language which could possibly be explored further.

To what extent is a unidimensional trait measured by the Picture Vocabulary Test? Rasch analyses were undertaken to explore the fit of the items to an underlying trait, in this case vocabulary. Furthermore, a developmental pathway was explored to further substantiate claims for the measurement of a unidimensional trait.

To what extent do the items in the Picture Vocabulary Test perform the same for the different language groups?

Once the items were thoroughly investigated, further analyses took place in which the item functioning for the three language groups were examined.

How can the identified barriers that decrease the level of validity be minimized?

Suggestions based on the analyses undertaken were generated, for example what the possible solutions could be for items which did not perform as expected for the different language groups.

5.2 SAMPLE

1361 learners participated in this research study. Of these learners, 355 took the Afrikaans test, 562 took the English test and 444 took the Sepedi test. The average age of the learners was 7 years but ranged from 6 to 8 years.

The three language groups and the results of the data from the objects are discussed, (Picture 1 Section 5.2, Picture 2 Section 5.3 and Picture 3 Section 5.4). Then each language group was discussed individually, starting with the Afrikaans learners, (Section 5.2.2, Section 5.3.2, Section 5.4.2) then followed by English (Section 5.2.3, Section 5.3.3, Section 5.4.3) and Sepedi learners (Section 5.2.4, Section 5.2.5, Section 5.2.6).

In this chapter the results of the analyzed data from learner responses from the Picture Vocabulary Test are presented in tables, Item-Learner Maps and Item Development Pathways. There were 1361 learners' responses in total. The learners' answers were captured, as described in Chapter 4 (Section 4.1.1) and analyzed using Rasch modeling.

Rasch analyses were conducted with the purpose of exploring the level of validity of the assessment particularly that of construct validity. Items that do not function correctly do not adhere to the assumptions of the Rasch model and may not be measuring the same construct. One of the assumptions when using the Rasch model is that a single latent trait is being measured (unidimensionality), implying that the

items follow a hierarchical order and will perform the same across different groups with the same knowledge (Beaton & Wright, 2005; Kyriakides, Kaloyirou & Lindsay, 2006).

An important point to take note of is that with Rasch analyses the data that do not provide relevant information are discarded. The reason the data are discarded for this research study is because they are "…not useful discriminators of the substantive sequence under investigation…" (Bond & Fox, 2001, p. 13). Put differently, data that fit the Rasch model perfectly do not provide measurement information about how the items are performing. If a learner gets 10 questions and gets 1 to 6 correct and then 7 to 10 incorrect, he or she is following the Rasch model because each item gets progressively more difficult thereby increasing the probability of getting the next item incorrect. But if the learner gets the first three items incorrect and the last seven correct then his or her data is worth investigating as to why the pattern is the way it is. The analysis of the data received from the SAMP assessment is restricted to the Picture Vocabulary Test, (PVT) consisting of 22 items implying that there are 22 objects that the learner needs to identify. Learners took the test at the beginning of the year and again at the end of the year (Chapter 4, Section 4.1.1). As was mentioned in Chapter 4 (Section 4.1.1) the test consists of three different pictures that were modified to be more appropriate to the South African context. The first picture found in Section 5.2 is a picture of a kitchen, the second picture, found in Section 5.3 is a picture of a view from a bedroom window overlooking a field, and the third picture in Section 5.4 is a child's bedroom.

In order to get a better understanding of all the tables, figures and diagrams a brief explanation is provided.

5.3 UNDERSTANDING RASCH ANALYSES

In the pages to follow, each picture with its related objects was explored and discussed. Item Developmental Pathways are provided to show the exact order, according to difficulty, that the objects followed for each language group. Additionally tables are presented that provide numerical information about the exact difficulty of each of the objects that were identified by the three groups. Item-Learner Maps

indicating the learners abilities compared to the item difficulties are presented as well.

Objects that proved easy to identify have a minus (-) sign in front of the numeric value. The higher the numeric value after the minus sign, the easier the object is to identify. For example, if the picture of the carrot appeared around -4 logits on the object map, it would be considered a very easy item. The difficult items have a plus (+) sign in front of them and the higher the numeric value, the more difficult the object is to identify. Beneath each of the pictures for each individual language group the following can be found:

A learner and item statistics table: In this table the item and learner information is provided. The OUTFIT and INFIT mean square with a value range of 0.7 to 1.3. This table indicates how the learners performed in relation to the items (objects) and the items (objects) difficulty in relation to the learners' abilities. In this table the separation reliability is also provided. If learner and item values fall out of these ranges it indicates that the learners' abilities do not match the item difficulties and visa versa (Linacre, 2005).

Table with object statistics: In this table the items that performed differently to what was expected by the Rasch model are identified. These items have OUTFIT or INFIT mean square values that are above or below the expected value range of 0.7 – 1.3 (Linacre, 2005).

Item-Learner map: This is a vertical line that has X's on the left side that represent the number of learners and their abilities along the variable (y-axis). (This was discussed under the heading Item-Learner Map).

Item Development Pathway (IDP): The objects are displayed along the pathway. Objects near the bottom of the IDP are easier and objects further up are more difficult. The objects with large circles show that they have large standard errors. The size of the circle is depicted by the Standard Error (SE) of each object as described in Chapter 3, Section 3.2.6. Items falling outside the predetermined area not fit the Rasch model and needs to be investigated further. On the IDP values of -2 and +2

are shown these are the t statistics that have been standardised (ZSTD) Items that fall within these values conform to the Rasch model.

Table with object order: This table shows the original PIPS order of objects compared to a specific language group's order of objects. This gives a clearer view of how the difficulties of the objects were experienced by the learners. The objects that follow the same order as the original SAMP and PIPS objects order are highlighted in **light green**. At the beginning of each picture all three languages are discussed, a table with all three language groups' object order is also given. When the objects follow the same difficulty order across all three languages the results are highlighted in **light blue**.

A few important points that act as general guidelines (Linacre, 2005, p. 141) have to be kept in mind when looking at the data of this study:

- Investigate OUTFIT before INFIT
- Mean-square before t standardized (ZSTD)
- High values before low or negative values

Linacre (2005) further mentions that when the mean-square is acceptable then the ZSTD scores can be ignored. The ZSTD asks the question: Does the data fit the model? With Rasch analyses ZSTD becomes over sensitive when the sample is over 300 resulting in exaggerated scores. The sample size for all three language groups is over 300 for this study. Ben Wright (Linacre, 2005, p. 141) gives the following advice regarding ZSTD: "ZSTD is only useful to salvage non significant MNSQ ›1.5, when sample size is small or test length is short". For this reason only the INFIT and OUTFIT mean square will be reported for the items.

5.4 PICTURE 1 – PICTURE OF KITCHEN

In Picture 1 a drawing of a kitchen with various objects can be seen. There are 7 objects that need to be identified by the learners namely **carrots, knife, fork, cupboard, cherries, pan** and **bowl.** The first object is the easiest of all the objects, according to the original PIPS instrument from the UK. The objects become

progressively more difficult as mentioned throughout the study. The picture of the kitchen and the objects in it is presented (Figure 5.1).

Figure 5.1: Picture 1

The objects to be identified by the learners were sequenced in the following order according to the original PIPS instrument: carrots, knife, fork, cupboard, cherries, pan and bowl - with carrots being the easiest to identify and bowl the most difficult (see Appendix A). This order of difficulty and items for Picture 1 are discussed.

5.4.1 Findings across all three language groups for Picture 1

For all the learners for Picture 1 information is provided in Table 5.1 about the learners and items performance (Appendix B).

Table 5.1: Learner & Item statistics for all learners for Picture 1

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	2.18	1.00	-0.20	1.04	0.00	0.00
Items	-3.67	1.00	0.00	1.02	0.20	0.83

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for both learners and items was 1.04 and 1.02, which is close to 1 indicating that there are no extremely difficult or easy items or learners that performed extremely well or bad for Picture 1. The INFIT MNSQ values for both learners and items were 1 which reflects that the objects used were correctly targeted for the learners (Scherman, 2007; Bond & Fox, 2001).

The separation reliability for the learners was 0.00 indicating low reliability. This could be due to the fact that the items were not targeted to a specific group (Linacre, 2005). It is also possible that different ability levels are not adequately distinguished along the continuum as described in Chapter 3 (Section 3.3.2).

In Table 5.2 the items results are provided and then discussed (Appendix C).

Table 5.2: Object statistics for all 3 language groups for Picture 1

Objects	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Pan	1.02	0.3	1.07	1.0	0.80
Cherries	1.02	0.2	1.03	0.2	0.33
Bowl	0.97	-0.5	0.96	-0.6	0.56

The objects carrots, knife, fork, and cupboard do not appear in the table because the information they provide is uninformative and has been excluded by Winsteps automatically (Linacre, 2005).

For ***pan*** the respective OUTFIT and INFIT MNSQ was 1.07 and 1.02 which falls within the expected range of 0.7 to 1.3 (Kyriakides, Kaloyirou & Lindsay, 2006). The Point Measure Correlation, which is the correlation between the observations in the data and the measures of the items or persons producing them, of 0.80 was acceptable and positive meaning that the item is functioning as expected; any items below 0.20 were flagged as possible items that may need reconstructing (Thorndike, 1997; van den Berg & Vorster, 1982). Furthermore, the Point Measure Correlation indicates that the higher the ability of the person the more likely the person will get the item correct. For ***cherries*** the respective INFIT and OUTFIT MNSQ was 1.02 and 1.03, which fall in the expected range. For ***bowl*** the respective INFIT and OUTFIT MNSQ was 0.97 and 0.96, fall in the expected range of 0.7 – 1.3.

An essential exploration is the item-learner targeting. This can be explored by means of Item-Learner Maps which are described in the introduction as well as Chapter 3 (Section 3.2.6). An Item-Learner Map is shown for Picture 1 across all the language groups. Clearly seen from the Item-Learner Map is that the learners' abilities and items difficulties are not targeted correctly. Ideally for every item difficulty there should be corresponding learner ability (Linacre, 2005). The map also clearly illustrates that the learners' abilities are greater than the most difficult item, namely pan. What is cause for concern is the large gap between the items. Ideally there should be objects which get progressively more difficulty with equal gaps between them as opposed to the large gaps found between the objects.

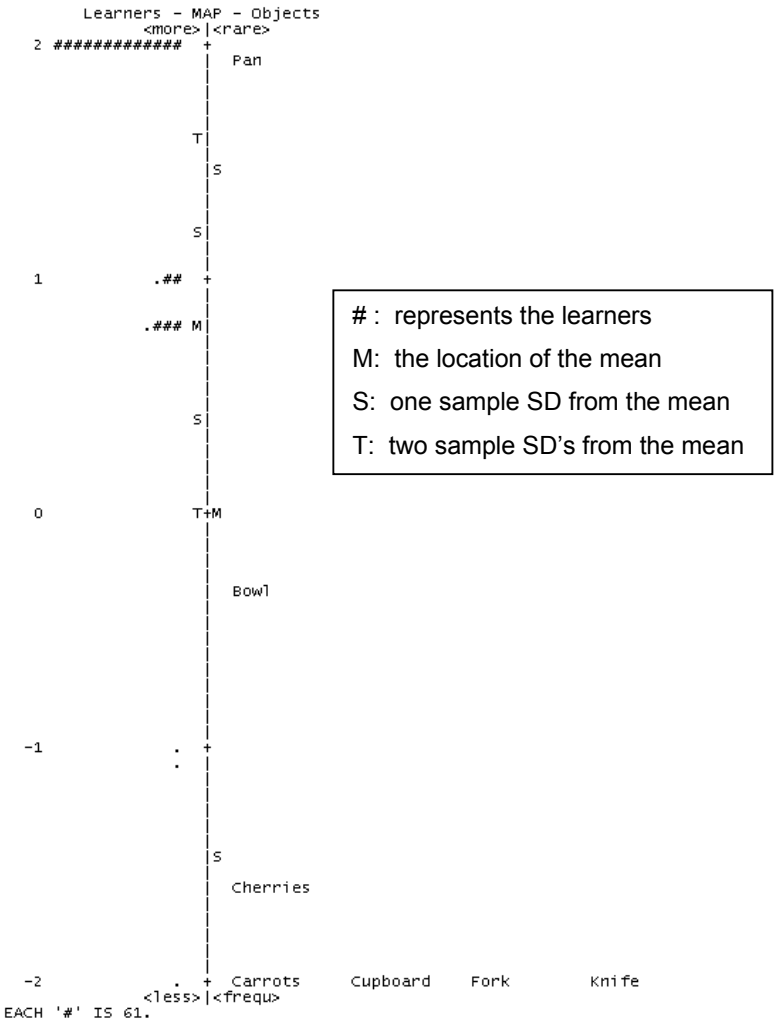


Figure 5.2: Item-Learner Map for all languages – Picture 1

Large gaps in difficulty are seen between cherries, bowl and pan. This could be because these items function differently across the three language groups (Linacre, 2005). Carrots, knife, fork and cupboard fall on the same difficulty level as illustrated in the Item-Learner Map above. These objects were experienced as being easy for the learners to identify indicated by the -2 to the left of these objects.

5.4.2 Afrikaans learners' results for the Baseline assessment of Picture 1

Of the 1361 learners, 355 were Afrikaans learners. However, once the uninformative responses had been deleted only data from 303 Afrikaans learners' were analysed. Any data that had perfect scores or were 0 were not included as they do not provide

any useful information (Linacre, 2005). In Table 5.3 the results are displayed (Appendix D).

Table 5.3: Learner & Item statistics for Afrikaans learners for Picture 1

	Mean*	INFIT MNSQ***	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	5.44	1.55	0.80	0.76	0.10	0.00
Items	0.00	0.88	0.10	0.73	-0.30	0.96

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for the learners of 0.76 fell below the expected value of 1 indicating that the learners' responses were too predictable. The OUTFIT MNSQ for the items is 0.73 indicating overfit or that the answers were too predictable. The INFIT MNSQ value for the learners was 1.55, above the expected value (underfit), indicating noise; the items did not perform as was expected by the Rasch model. The INFIT MNSQ value for the items was 0.88 indicating underfit.

The separation reliability for the learners was 0.00 indicating low reliability. This could be due to the fact that the items were not targeted for the Afrikaans learners and that there is not a range of ability levels represented along the continuum. This indicates that the measurements made about the learners' ability weren't accurate, in other words the abilities of the learners were not accurately matched. The item separation reliability for the items was 0.96 indicating that the objects used in Picture 1 do have varying difficulty levels as discussed in Section 3.3.2 in Chapter 3.

In Table 5.4 the object statistics are shown and discussed (see Appendix E).

Table 5.4: Object statistics for Afrikaans learners for Picture 1

Objects	Logit	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Pan	4.33	1.05	1.50	2.68	9.90	0.71
Bowl	2.67	1.01	0.20	1.03	0.60	0.59
Carrots	-1.52	1.01	0.10	0.14	-2.40	0.34
Cherries	1.55	0.94	-0.30	0.93	-0.30	0.46
Cupboard	-0.85	0.77	-0.60	0.23	-2.10	0.34
Fork	-2.67	0.74	-0.20	0.06	-2.70	0.21
Knife	-3.51	0.67	-0.10	0.01	-5.30	0.17

For ***pan*** the high OUTFIT MNSQ of 2.68 indicates unexpected responses occurred by the learners on this item or that there were only a few random responses by low performers. For ***fork*** the low OUTFIT MNSQ of 0.06 indicates that it is a problematic item needing further exploration .This could be due to unexpected observations, for example being easier than expected (Linacre, 2005). For ***knife*** the low OUTFIT MNSQ of 0.01 indicates unexpected observations that occurred by the learners on this item. *Knife* was experienced as being easier than expected from the Rasch model. The lower INFIT MNSQ of 0.67 indicates that the model predicts the data too well (Linacre, 2005).

Now the item-learner targeting will be explored. This can be explored by means of Item-Learner Maps as described in the introduction. An Item-Learner Map is shown for Picture 1 for the Afrikaans group. The Item-Learner Map shows that the learners' abilities exceeded the items difficulties and were not targeted correctly. Ideally for every item difficulty there should be corresponding learner ability (Linacre, 2005). The map also clearly illustrates that the learners' abilities were greater than the most difficult item, namely **pan** as was seen in the first analysis. Most of the objects were equally dispersed although a large gap is present between **cupboard** and **cherries.**

Ideally there should be objects which get progressively more difficulty with equal gaps between them as opposed to the large gaps found between the objects.

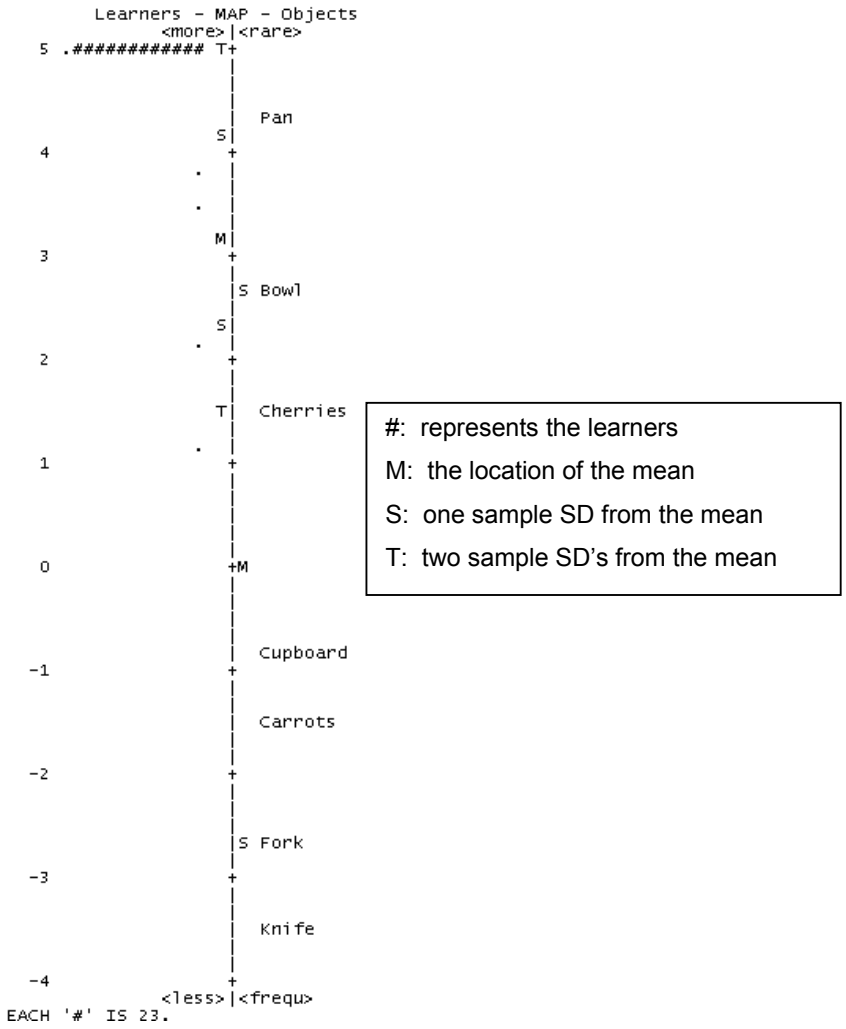


Figure 5.3: Item-Learner Map for Afrikaans – Picture 1

The Item Development Pathway (IDP) Figure 5.4 illustrates how the objects are positioned according to difficulty from the Afrikaans learners' data. Objects near the bottom of the IDP are easier and objects further up are more difficult. The objects are displayed along the pathway. The four objects at the bottom of the pathway are fairly evenly distributed, while there is a large gap between **cupboard** and **cherries. Cherries** and **bowl** are evenly spaced but there is a large gap between **bowl** and **pan.** The objects with large circles show that they have large standard errors (Bond &

Fox, 2001). The size of the circle is depicted by the Standard Error (SE) of each object as described in Chapter 4, Section 4.3.6.6.

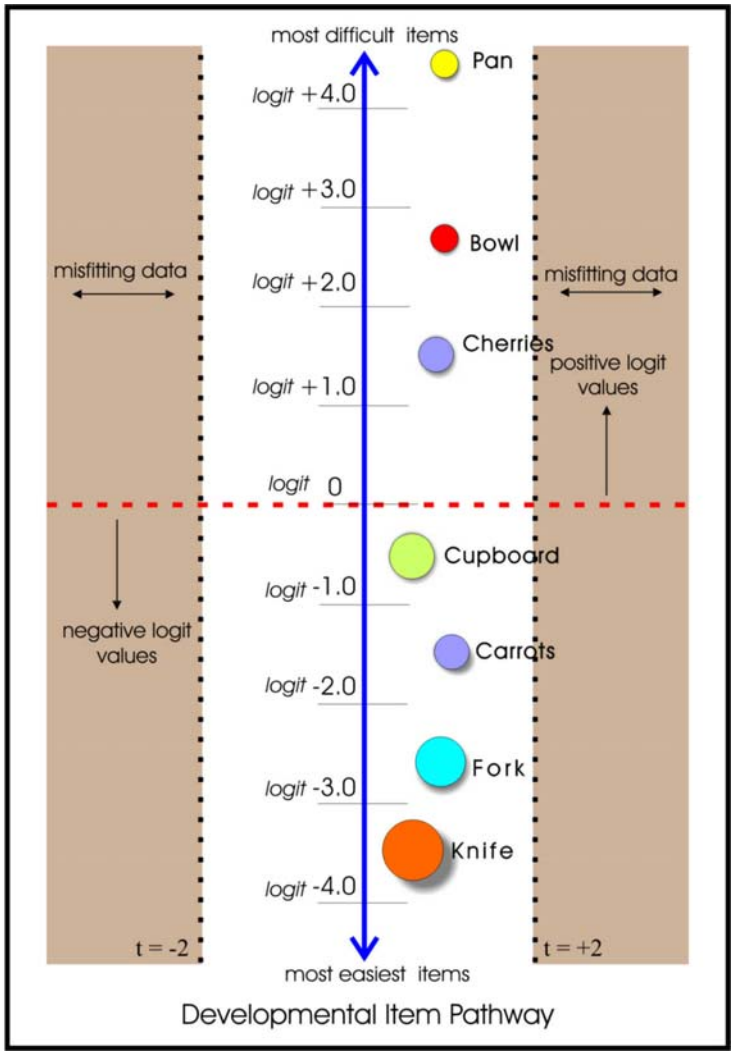


Figure 5.4: Item Development Pathway for Picture 1 – Afrikaans group

Knife has a large circle that indicates that difficulty could not be allocated precisely (Bond & Fox, 2001). **Bowl** has a very small circle which indicates that its difficulty could be allocated rather precisely (Bond & Fox, 2001). Therefore it would be easy to allocate the difficulty of **bowl** to a learner's ability but the same cannot be said for **knife**. The easiest object to identify was the **knife** and the most difficult object was the **pan**. The items were almost equally distributed along the item pathway although some items were lacking in the middle range close to 0. Unidimensionality is important and indicates whether the items are working together to define a single

construct based on the Item Development Pathway and other statistics discussed earlier. In this IDP there were no misfitting items which indicate that a single construct was measured.

Table 5.5 displays the logit values of the objects for Picture 1. The objects are arranged according to order of difficulty for the Afrikaans learners and compared to the original PIPS order (see Appendix F).

Table 5.5: Object order for Picture 1 - Afrikaans group

Original Order	Afrikaans Order (logit values)	Standard Error
1. Carrots	2. Knife (-3.51)	1.02
2. Knife	3. Fork (-2.67)	0.74
3. Fork	1. Carrots (-1.52)	0.38
4. Cupboard	4. Cupboard (-0.85).	0.42
5. Cherries	5. Cherries (+1.55)	0.22
6. Pan	7. Bowl (+2.67)	0.10
7. Bowl	6. Pan (+4.33)	0.10

As can be seen from the above table, there are few similarities between the original PIPS order of objects and the Afrikaans learners' order of objects. The two objects, **cupboard** and **cherries**, are in the same order of difficulty as the original PIPS instrument. The other objects do not follow the original order of difficulty for example **knife** and **fork**.

As clearly indicated the items did not perform in the manner expected for Picture 1 for the Afrikaans learners in comparison to the original PIPS object order of the sub-test from the UK. The English learners' results will now be investigated.

5.4.3 English learners results for Baseline assessment of Picture 1

The English learners made up 562 learners of the total of 1361 learners who participated in the study. Once the responses which were not informative for

measurement purposes were eliminated a total 514 learner responses were analysed. Any data that had perfect scores or were 0 were not included as they do not provide any useful information for measurement (Linacre, 2005). In Table 5.6 the results are displayed (refer to Appendix G).

Table 5.6: Learner & Item statistics for English learners for Picture 1

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	4.90	2.17	1.40	0.93	0.40	0.00
Items	-2.82	1.05	0.10	1.62	4.40	0.97

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for the learners of 0.93 was within the expected range indicating that the learners' responses matched the expected responses. The OUTFIT MNSQ for the items was 1.62 which is above the expected range indicating that unexpected responses occurred by the learners on extremely difficult or easy items for Picture 1 for the English learners. The INFIT MNSQ value for the learners was 2.17 which are above the ideal range showing that unexpected response patterns occurred. The INFIT MNSQ value for the items was 1.05 which falls within the expected range. This reflects that the objects used are correctly targeted for the English learners.

The separation reliability for the learners was 0.00 indicating low reliability. This could be due to the fact that the items were not targeted for the English learners' ability levels as placed on a continuum. The item separation reliability was 0.97 indicating that the objects used in Picture 1 have varying difficulty levels.

In Table 5.7 the object statistics are shown and then discussed (see Appendix H)

Table 5.7: Object statistics for English learners for Picture 1

Objects	Logit	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Cherries	-0.59	1.32	2.80	4.62	8.40	0.04
Fork	-5.01	1.03	0.20	2.32	3.30	0.05
Cupboard	-3.35	1.07	0.30	1.82	1.80	0.10
Pan	0.64	1.12	3.20	1.06	0.80	0.46
Knife	-6.31	1.01	0.30	0.11	-7.50	0.06
Bowl	-0.84	0.91	-1.40	0.92	-0.50	0.45
Carrots	-4.28	0.88	-0.30	0.48	-1.90	0.19

For ***cherries*** the high OUTFIT MNSQ of 4.62 indicates unexpected responses occurred by the learners on this item or noise as there were other sources of variance in the data that could not be modeled. For ***fork*** the high OUTFIT MNSQ of 2.32 indicates unexpected responses occurred on this item. For ***cupboard*** the high OUTFIT MNSQ of 1.82 indicate that unexpected responses occurred on this item.

The Item-Learner targeting is explored. The Item-Learner Map is for Picture 1 for the English group. The Item-Learner Map shows that the learners' abilities greatly exceeded the items difficulties and were not targeted correctly. Ideally for every item difficulty there should be corresponding learner ability (Linacre, 2005). The map also clearly illustrates that the learners' abilities are greater than the most difficult item, namely **pan** as was the case with the Afrikaans learners. Most of the objects were equally dispersed although a large gap is found between **cupboard** and **bowl**, while **bowl** and **cherries** difficulty levels are in close proximity of each other.

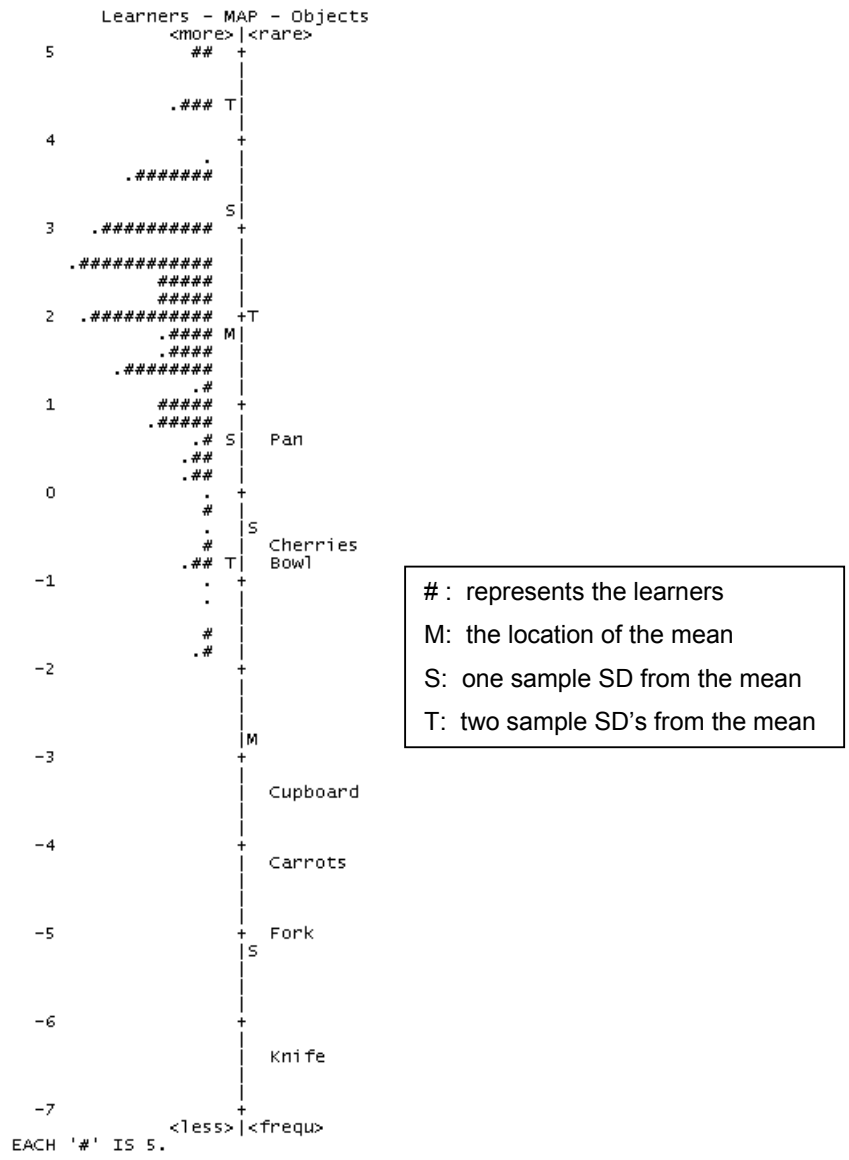


Figure 5.5: Item-Learner Map for English – Picture 1

The performances of the items are indicated in the IDP for the English learners for Picture 1. The items are situated along a pathway that moves from easy to difficult.

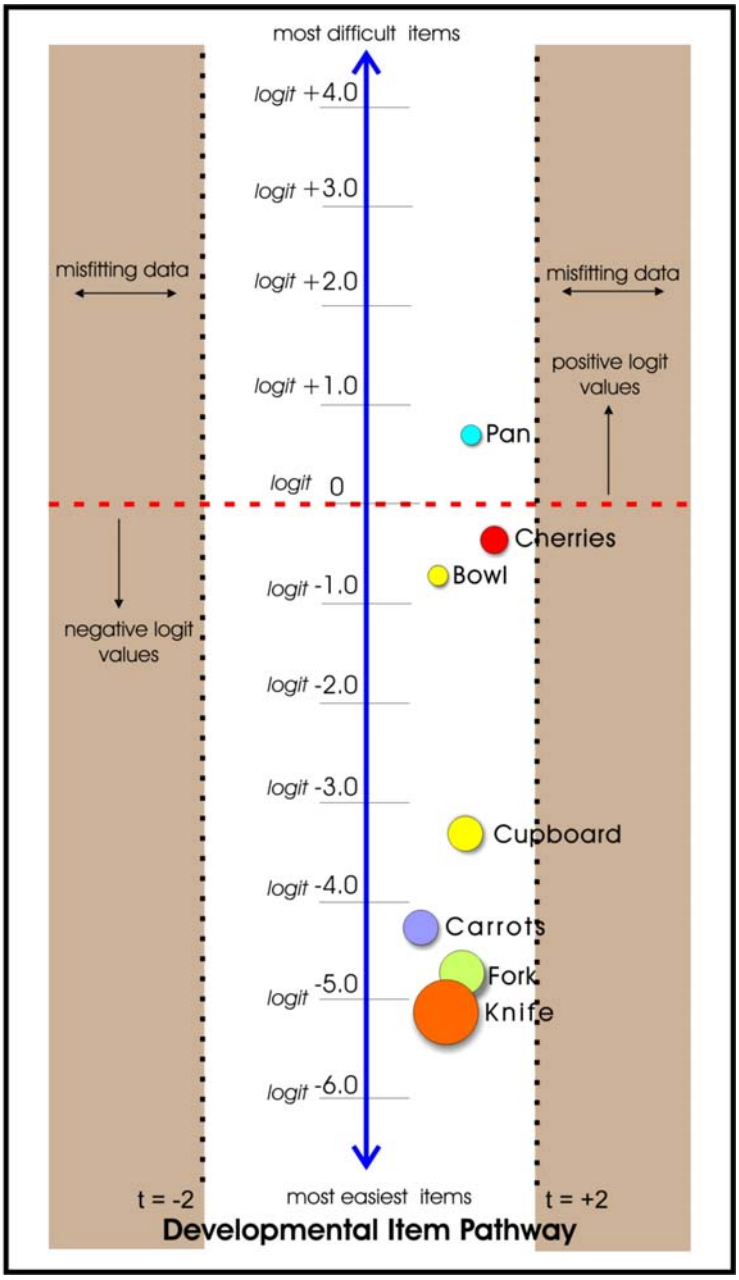


Figure 5.6: IDP for English group – Picture 1

The IDP shows how the objects in Picture 1 were placed according to difficulty order for the English learners. Six of the seven objects were experienced as easy for the English learners with the exception of the **pan**. **Pan** was seen as the most difficult of all the objects with the easiest item being **knife**. The objects are not equally distributed in relation to difficulty; the objects' difficulties are in close proximity to each other. This results in the learners abilities not being accurately measured. It appears

that the objects in Picture 1 are mostly easy identifiable objects with no particularly difficult items for the English learners. Furthermore there is a slight overlap between **knife** and **fork** which makes the ordering of the items unclear. Unidimensionality indicates whether the items are working together to define a single construct based on the item development pathway. There are no misfitting items which gives an indication that a single construct is measured.

Adding to the above information, Table 5.8 shows the logit values of the objects in column 2. The objects are arranged according to order of difficulty and standard error (see Appendix I).

Table 5.8: Object order for Picture 1 - English group

Original Order	English Order (logit value)	Standard Error
1. Carrots	2. Knife (-6.31)	1.01
2. Knife	3. Fork (-5.01)	0.59
3. Fork	1. Carrots (-4.28)	0.35
4. Cupboard	4. Cupboard (-3.35)	0.33
5. Cherries	7. Bowl (-0.84)	0.10
6. Pan	5. Cherries (-0.59)	0.15
7. Bowl	6. Pan (+0.64)	0.07

As can be seen from the above table, there is only one similarity, namely **cupboard** between the original PIPS order of objects and the English learners' order of objects. This means that the objects did not follow the same difficulty order as the objects in the PIPS object order. As can be seen the items did not perform in the manner expected for Picture 1 for the English learners. The Sepedi learners' results will now be investigated.

5.4.4 Sepedi learners results for Baseline assessment of Picture 1

A total of 444 learners from the entire 1361 learners were Sepedi and once the responses not sufficient for measurement were eliminated, a total of 404 learners'

data were explored. Uninformative data that had perfect scores or were 0 were not included since their information was not considered useful (Linacre, 2005). In Table 5.9 the results are displayed (see Appendix J).

Table 5.9: Learner & Item statistics for Sepedi learners for Picture 1

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	3.53	0.51	-0.90	0.25	-0.30	0.00
Items	-0.69	0.99	0.30	0.63	-1.20	0.96

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for the learners of 0.25 is below the expected range indicating that the learners' responses did not match expected responses. The OUTFIT MNSQ for the items and learners was 0.63 and 0.51 respectively which is below the expected range indicating that unexpected responses occurred by the learners on extremely difficult or easy items for Picture 1 for the Sepedi learners. The INFIT MNSQ value for the items was 0.99 which falls within the expected range. This reflects that the objects used are correctly targeted for the Sepedi learners.

The separation reliability for the learners was 0.00 indicating low reliability. This could be due to the fact that the items were not targeted for the Sepedi learners. The items difficulties were also not correctly targeted to the learners' abilities. The item separation reliability was 0.96 indicating that the objects used in Picture 1 do have varying difficulty levels.

In Table 5.10 the object statistics are shown and discussed (see Appendix K).

Table 5.10: Object statistics for Sepedi learners for Picture 1

Objects	Logits	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Pan	3.89	1.04	1.5	1.29	3.20	0.69
Cherries	2.36	1.04	0.50	1.04	0.40	0.45
Cupboard	-1.25	0.98	0.10	0.39	-1.30	0.20
Fork	-3.48	0.98	0.30	0.09	-3.20	0.12
Bowl	2.45	0.95	-1.10	0.92	-1.40	0.58
Carrots	-3.96	0.95	0.30	0.05	-5.00	0.13

Knife does not appear in the table because the information it provides is uninformative and has been excluded by Winsteps automatically.

For ***pan*** the OUTFIT MNSQ of 1.29 shows that it is misperforming for the learners it was targeted for. For ***fork*** the OUTFIT MNSQ of 0.09 indicates unexpected responses occurred on this item and that the observations were too predictable. For ***carrots*** the OUTFIT MNSQ of 0.05 indicates unexpected responses occurred on this item. Some of the items clearly indicate that they did not follow the Rasch model.

The item-learner targeting is explored in the Item-Learner Map for Picture 1 for the English group. The Item-Learner Map shows that the learners' abilities nearly matched the items difficulties but were not targeted exactly. The map also clearly indicates that the learners' abilities are higher than the most difficult item, namely **pan**. Large gaps were found between most of the objects. The difficulty level of **carrots** and **knife** were in close proximity to each other as well as **bowl** and **cherries**. In the Item-Learner Map a lesser amount of Sepedi learners' abilities exceed the difficulty of the items.

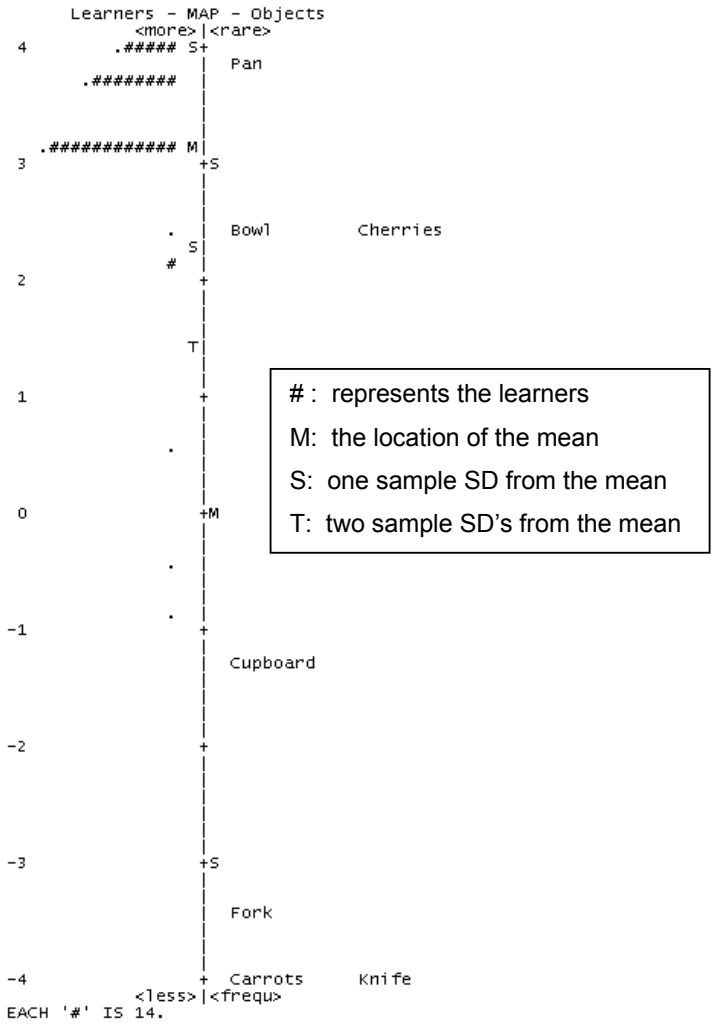


Figure 5.7: Item-Learner Map for Sepedi – Picture 1

The Item Development Pathway for the Sepedi learners is given. The items are situated along a pathway that moves from easy at the bottom to difficult at the top.

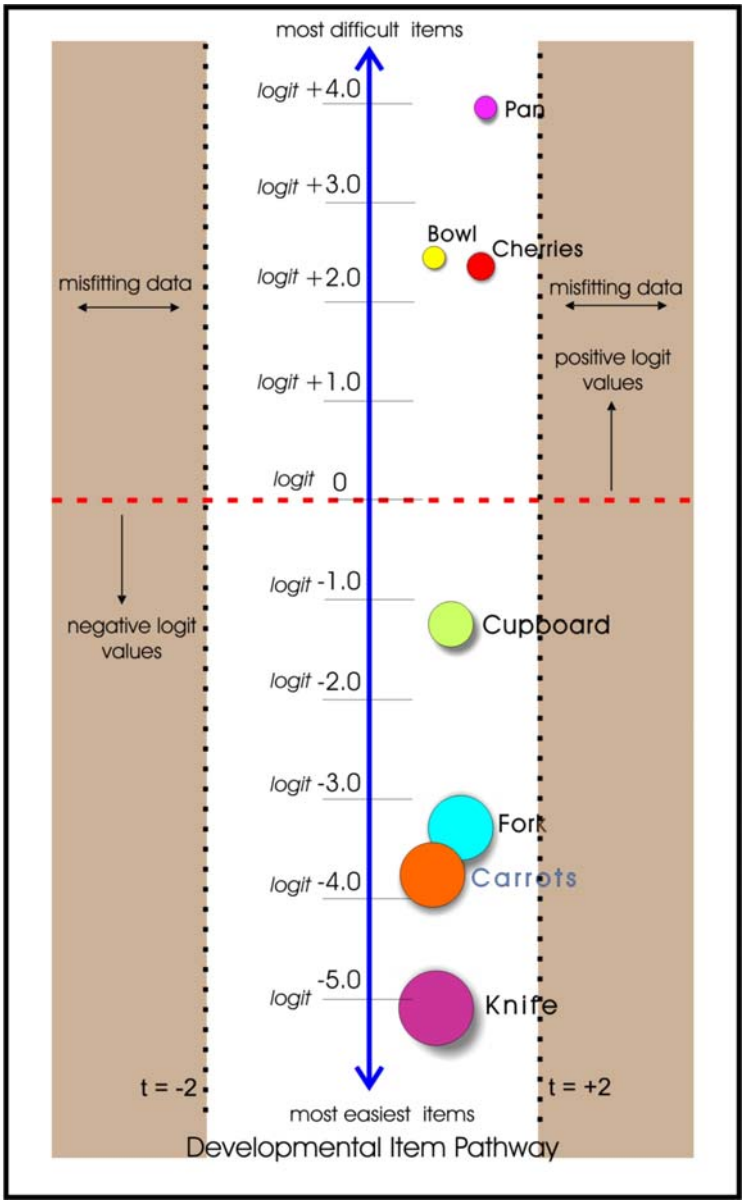


Figure 5.8: IDP for Sepedi group – Picture 1

The Sepedi learners tended to find the objects either easy or difficult. The objects were not distributed equally, and as a result differing gaps were found between the items. Items were found lacking in the middle range close to 0. **Carrots** and **knife** were in close range of difficulty to each other as well as **bowl** and **cherries**. The easiest object to identify for the Sepedi learners was the **knife** and the most difficult was the **pan**. **Knife, carrots** and **fork** had large SE's, showing there may be some uncertainty associated with the estimates. Once again all the items fall within the item

pathway indicating that items are not misfitting and that the same construct is measured.

Table 5.11 reflects the logit values of the objects in column 2. The objects are arranged according to order of difficulty. The Standard Error (SE) for the Sepedi results is given in Column 3 (see Appendix L).

Table 5.11: Object order for Picture 1 - Sepedi group

Original Order	Sepedi Order	Standard Error
1. Carrots	2. Knife (-4.86)	1.82
2. Knife	1. Carrots (-3.96)	1.01
3. Fork	3. Fork (-3.48)	1.01
4. Cupboard	4. Cupboard (-1.25).	0.46
5. Cherries	5. Cherries (+2.36)	0.16
6. Pan	7. Bowl (+2.45)	0.10
7. Bowl	6. Pan (+3.89)	0.10

As can be seen from the above table, there are slight similarities between the original PIPS order of objects and the Sepedi learners' order of objects. The three objects, **fork, cupboard** and **cherries**, were in the same order of difficulty as the original PIPS instrument although the rest of the objects did not follow the original order of difficulty.

5.4.5 Summary of Picture 1 across all groups

Table 5.12 mirrors the order of the objects for all three language groups for Picture 1 as well as the original PIPS order. As mentioned in Section 5.1 above, the cells shaded in **light green** represent the objects that follow the original PIPS instruments' difficulty order. The cells shaded in **light blue** represent the objects that follow the same difficulty order across all three language groups.

Table 5.12: Object order for Picture 1 for all three language groups

Original Order	Afrikaans Order	English Order	Sepedi Order
1. Carrots	2. Knife	2. Knife	2. Knife
2. Knife	3. Fork	3. Fork	1. Carrots
3. Fork	1. Carrots	1. Carrots	3. Fork
4. Cupboard	4. Cupboard	4. Cupboard	4. Cupboard
5. Cherries	5. Cherries	7. Bowl	5. Cherries
6. Pan	7. Bowl	5. Cherries	7. Bowl
7. Bowl	6. Pan	6. Pan	6. Pan

The only object that followed the original difficulty order across all three language groups was **cupboard**. Although there were differences in the order for all three groups, a few similarities were found. The similarities were:

- Across all three language groups: **knife, cupboard** and **pan**.
- Across Afrikaans and Sepedi learners: **cupboard, cherries, bowl** and **pan**.
- Across Afrikaans and English learners: **knife, fork, carrots** and **cupboard**.
- Across English and Sepedi learners**: knife, cupboard** and **pan**.

Very few similarities were found in the order of the objects of the original PIPS instrument compared to the object order of the three language groups. Interestingly there were a number of similarities between the three groups even though these similarities are present with certain objects. More similarities are found among the three groups than from the original order of the UK instrument. But these are only the results for Picture 1.

Next the individual language groups' performance will be discussed for Picture 2. The Afrikaans learners in Section 5.3.2, English in Section 5.3.3 and Sepedi in Section 5.3.4. The learners' abilities and the items' difficulties are shown in a table; misfitting items are identified in a table; learner performance compared to item difficulties is

mapped, items' performance is displayed on an Item Development Pathway and the object order of the PIPS and language groups are displayed in a table.

5.5 PICTURE 2 – PICTURE OF THE OUTDOORS

In Picture 2, the view is from a bedroom window overlooking a field as illustrated. In this picture, the learners have to identify 10 different objects that have also been arranged from easiest to most difficult. The order of the objects is as follows: **butterfly, kite, castle, wasp, pigeon, windmill, tortoise, violin, padlock** and **toadstool** (see Appendix M). This is the original order as found in the PIPS and SAMP assessment.

Figure 5.9: Picture 2

The results of the data analyses of the second picture of the Picture Vocabulary Test will be discussed per language group. Once again, an alphabetical order will be followed starting with the Afrikaans learners in Section 5.5.2, English learners in Section 5.5.3 and the Sepedi learners in Section 5.5.4.

5.5.1 Findings across all three language groups for Picture 2

For all three language groups for Picture 2 information is provided in Table 5.13 about the learners and items performance (see Appendix N).

Table 5.13: Learner & Item statistics for all language groups for Picture 2

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability*
Learners	0.78	0.99	0.10	0.92	0.30	0.53
Items	0.00	1.01	0.00	0.94	-0.60	1.00

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for both learners and items was 0.92 and 0.94 respectively, and are slightly lower than 1 indicating that the responses were too predictable. The INFIT MNSQ values for both learners and items were 0.99 and 1.0 respectively which was close to the expected range. This reflects that the objects used were correctly targeted for the learners. The separation reliability for the learners was 0.78 which was low showing that there is not enough variation in ability levels along the continuum. The item separation reliability for the items was 1.00 indicating that the objects used in Picture 1 do have varying difficulty levels.

In Table 5.14 the object statistics are shown (see Appendix O).

Table 5.14: Object statistics for all language groups for Picture 2

Objects	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Castle	1.18	5.60	1.26	4.40	0.48
Padlock	1.17	5.80	1.17	3.70	0.44
Toadstool	1.09	3.20	1.08	2.00	0.48
Windmill	1.09	2.10	1.09	1.30	0.39
Butterfly	1.04	0.30	0.51	-1.60	0.47
Tortoise	1.03	0.90	1.02	0.30	0.48
Wasp	1.01	0.30	0.98	-0.20	0.55
Kite	0.87	-4.30	0.78	-4.20	0.62
Violin	0.81	-7.10	0.76	-6.0	0.64
Pigeon	0.80	-7.10	0.71	-6.0	0.65

All of the items indicated in Table 5.14 are within the acceptable values for fit, with the exception of **butterfly** with an OUTFIT MNSQ of 0.51. However the INFIT MNSQ is acceptable. For ***castle*** the OUTFIT MNSQ of 1.28 shows that random responses occurred by learners. For ***padlock*** the INFIT and OUTFIT MNSQ of 1.17 is within the predetermined criteria of 0.7 and 1.3. For ***butterfly*** the OUTFIT MNSQ of 0.51 indicates that the observations were too predictable. For ***kite*** the OUTFIT MNSQ and INFIT MNSQ of 0.78 and 0.87 respectively was within the predetermined criteria of fit. For ***violin*** the OUTFIT MNSQ of 0.76 is within the predetermined criteria of fit. For ***pigeon*** the OUTFIT MNSQ of 0.71 was within the predetermined criteria of fit.

Item-Learner targeting is explored by means of Item-Learner Maps. An Item-Learner Map is shown for Picture 2 across all three language groups. Evidently seen from the Item-Learner Map are the learners' abilities and items difficulties are not targeted correctly. Ideally for every item difficulty there should be corresponding learner ability (Linacre, 2005). The map also clearly illustrates that the learners' abilities are greater than the most difficult item, **violin**. A large gap can be seen between **butterfly** and **wasp**. Ideally there should be objects which get progressively more difficulty with

equal gaps between them as opposed to the large gaps found between the objects. In the Item-Learner Map a considerable amount of learners' abilities exceed the difficulties of the items for Picture 2.

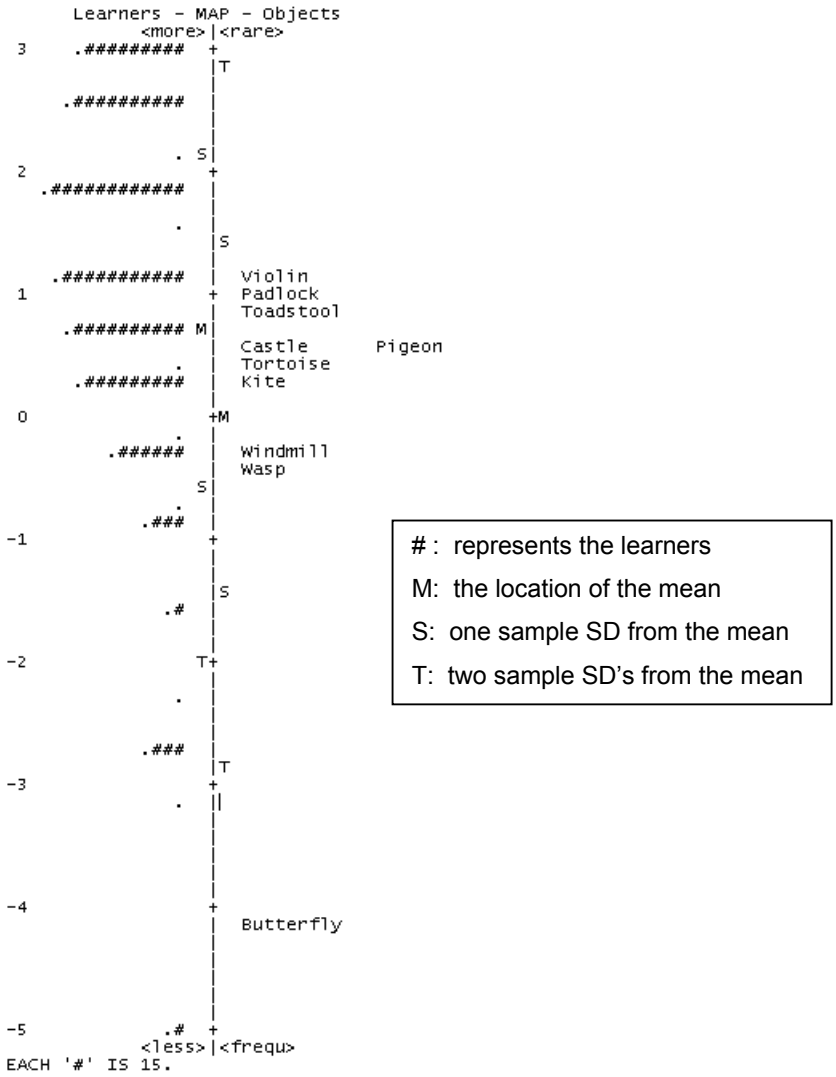


Figure 5.10: Item-Learner Map for all languages– Picture 2

The learners found the majority of objects difficult with the exception of **butterfly** that was seen as an easy item. The objects were not distributed equally, and as a result did not display great variation in difficulty. The majority of items were in the middle range between 1 and 0.

The easiest object to identify for Picture 2 for the learners was **butterfly** and the most difficult was **violin**.

5.5.2 Afrikaans learners results for Baseline assessment of Picture 2

In Table 5.15 the results are displayed regarding the Afrikaans learners' performance for Picture 2. The item-learner statistics, object statistics, Item-Learner Map, Item Development Pathway and the object order are provided (see Appendix P).

Table 5.15: Learner & Item statistics for Afrikaans learners for Picture 2

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	2.34	1.05	0.30	0.97	0.30	0.53
Items	0.00	1.00	-0.10	0.97	-0.40	0.99

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for the learners and items of 0.97 is slightly below 1 indicating that unexpected responses occurred on items. The INFIT MNSQ values for the learner and items respectively were 1.05 and 1.00 which falls within the expected range. The separation reliability for the learners was 0.53 indicating low reliability this could be due to the fact that the items were not targeted for all the ability levels for the Afrikaans learners. The item separation reliability was 0.99, which is close to one indicating that the objects used in Picture 2 do have varying difficulty levels used in Picture 2 do have varying difficulty levels.

In Table 5.16 the object statistics are shown and discussed (see Appendix Q)

Table 5.16: Object statistics for Afrikaans learners for Picture 2

Objects	Logit	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Castle	0.41	1.19	6.00	1.32	6.20	0.48
Padlock	0.87	1.17	5.80	1.16	3.90	0.45
Windmill	-0.49	1.09	2.30	1.10	1.40	0.39
Toadstool	0.76	1.09	3.20	1.08	2.10	0.48
Tortoise	0.33	1.02	0.80	1.00	0.10	0.49
Butterfly	-2.89	1.01	0.20	0.74	-1.10	0.45
Wasp	-0.48	0.97	-0.70	0.96	-0.70	0.56
Kite	0.12	0.87	-4.40	0.81	-4.00	0.62
Pigeon	0.38	0.82	-6.30	0.76	-5.50	0.65
Violin	0.99	0.81	-7.60	0.75	-6.80	0.65

All of the items adhere to the requirements of fit (0.7 – 1.3). For ***castle*** the OUTFIT MNSQ of 1.32 indicates that the learners responded unexpectedly to this item.

The Item-Learner Map is shown for Picture 2 for the Afrikaans group. Evidently seen from the Item-Learner Map are the learners' abilities and items difficulties are not targeted correctly. Ideally for every item difficulty there should be corresponding learner ability (Linacre, 2005). The map also clearly illustrates that the learners' abilities are considerably greater than the most difficult item, **violin**. A large gap can be seen between **butterfly** and **wasp**. Ideally there should be objects which get progressively more difficulty with equal gaps between objects as opposed to the large gaps found between the objects. **Wasp** and **windmill** as well as **castle** and **pigeon** respectively had difficulty levels that were in close proximity to each other. The majority of the objects difficulty was situated between 0 and 1.

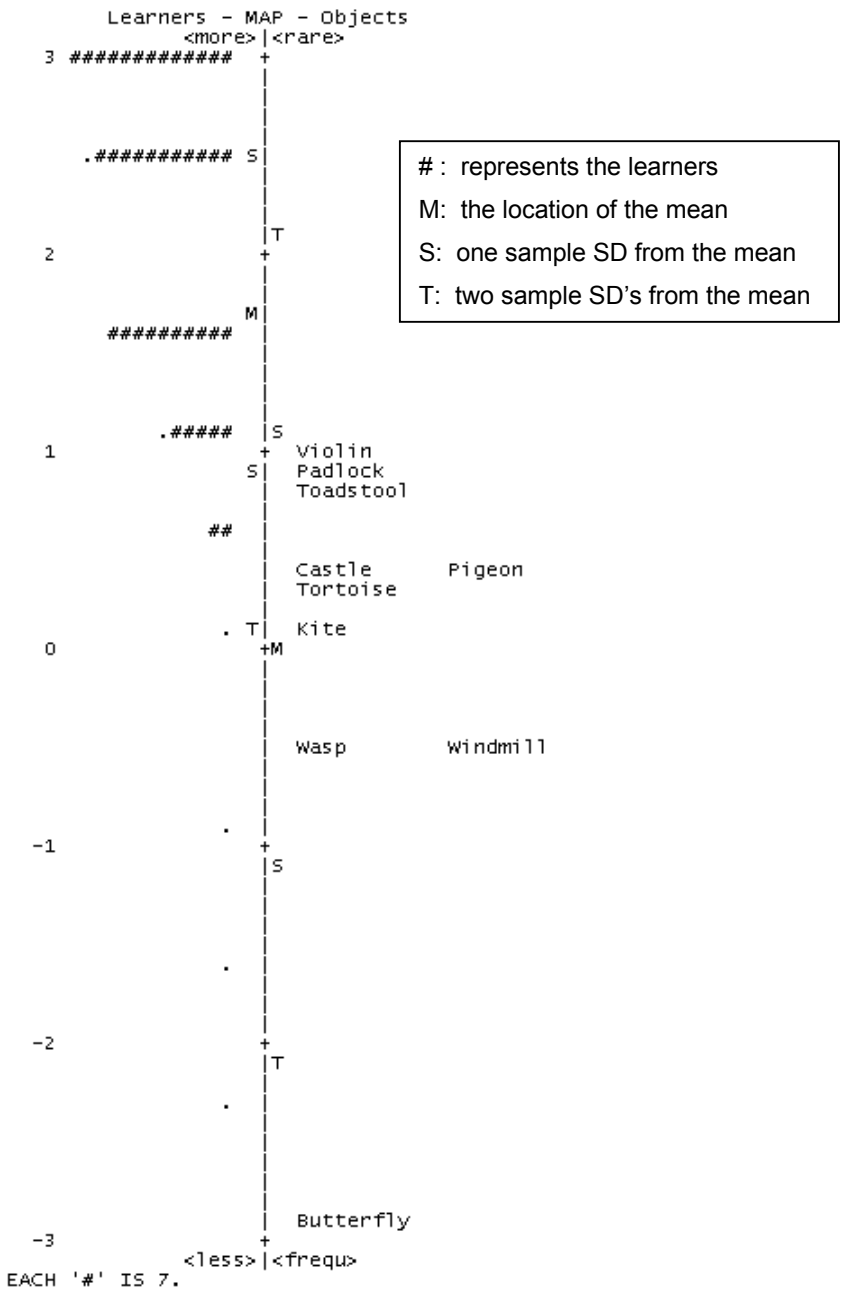


Figure 5.11: Item-Learner Map for Afrikaans– Picture 2

The Item Development Pathway for Picture 2 shows how the objects were placed according to the results of the data from the Afrikaans learners' assessment.

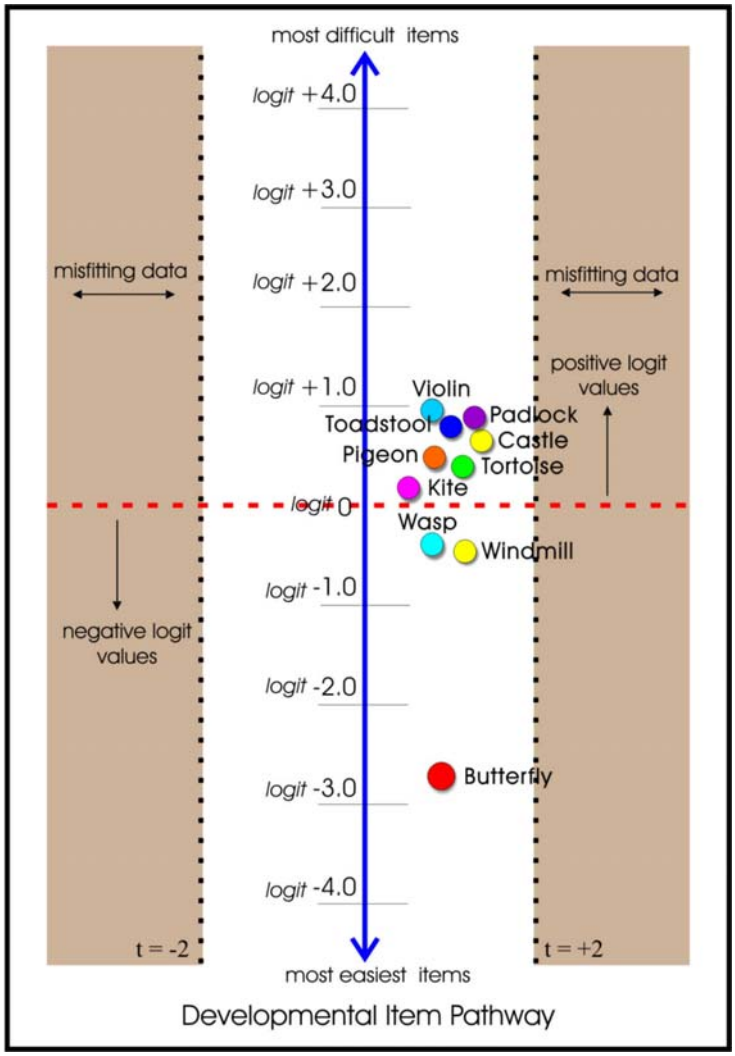


Figure 5.12: IDP for Afrikaans group – Picture 2

The learners found the majority of objects slightly difficult with few exceptions such as **butterfly**, **windmill** and **wasp** that were experienced as easy items. The objects were not distributed equally, and as a result did not display great variation in difficulty. The majority of items were in the middle range. The easiest object to identify for Picture 2 for the Afrikaans learners was **butterfly** and the most difficult was **violin**. A clear progression of items is difficult to establish for the Afrikaans learners as there are a number of items clumped together. Ideally these items would have equal intervals between them. Unidimensionality is an important indicator whether items are working together to define a single construct. For Picture 2 for the Afrikaans learners there are no misfitting items which indicate that a single construct was measured.

Table 5.17 reflects the logit values of the objects in column 2. The objects are arranged according to order of difficulty. The Standard Error (SE) for the Afrikaans results is given in Column 3 (see Appendix R).

Table 5.17: Object order for Picture 2 - Afrikaans group

Original Order	Afrikaans Order	Standard Error
1. Butterfly	1. Butterfly (-2.89)	0.13
2. Kite	6. Windmill (-0.49)	0.08
3. Castle	4. Wasp (-0.48)	0.08
4. WASP	2. Kite (+0.12)	0.07
5. Pigeon	7. Tortoise (+0.33)	0.07
6. Windmill	5. Pigeon (+0.38)	0.07
7. Tortoise	3. Castle (+0.41)	0.07
8. Violin	10. Toadstool (+0.76)	0.07
9. Padlock	9. Padlock (+0.87)	0.07
10. Toadstool	8. Violin (+0.99)	0.07

As can be seen from the above table, there are only two similarities, **butterfly** and **padlock**, found between the original PIPS order of objects and the Afrikaans learners' order of objects. The rest of the objects do not follow the original order of difficulty.

5.5.3 English learners results for Baseline assessment of Picture 2

In Table 5.18 the results are displayed regarding the English learners performance for Picture 2. The item-learner statistics, object statistics, Item-Learner Map, Item Development Pathway and the object order are given (see Appendix S).

Table 5.18: Learner & Item statistics for English learners for Picture 2

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability*
Learners	0.87	0.98	0.00	0.91	0.20	0.53
Items	0.00	1.01	-0.20	1.01	-0.40	0.99

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for both learners and items was 0.91 and 1.01 respectively, which was close to the expected value. However the OUTFIT MNSQ of 0.91 is below 1 indicating overfit. The INFIT MNSQ values for both learners and items were 0.98 and 1.01 respectively, which was close to the expected value. This reflects that the objects used were correctly targeted for the learners. The separation reliability for the learners was 0.53 which is lower than expected. The item separation reliability for the items was 0.99 indicating that the objects used in Picture 2 do have varying difficulty levels. In Table 5.19 the object statistics are shown (see Appendix T).

Table 5.19: Object statistics for English learners for Picture 2

Objects	Logits	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Castle	0.58	1.19	6.30	1.29	5.40	0.48
Butterfly	-3.45	1.07	0.60	1.21	0.80	0.42
Windmill	-0.32	1.10	2.70	1.20	2.90	0.38
Padlock	0.86	1.16	5.50	1.15	3.70	0.45
Toadstool	0.79	1.08	2.80	1.06	1.50	0.49
Tortoise	0.42	1.02	0.60	1.01	0.20	0.49
Wasp	-0.49	0.99	-0.10	0.97	-0.40	0.55
Kite	0.15	0.86	-4.70	0.77	-4.70	0.63
Pigeon	0.43	0.80	-7.10	0.72	-6.20	0.65
Violin	1.01	0.80	-7.90	0.74	-7.00	0.65

Once again, all the items adhere to the requirements of fit, as with the Afrikaans learners (0.7 -1.3).

The Item-Learner Map shows that the learners' abilities and items difficulties are not targeted correctly. For every item difficulty there should be corresponding learner ability (Linacre, 2005). The English learners' abilities far exceed the difficulty of the items for Picture 2. The Item-Learner Map also clearly illustrates that the learners' abilities are considerably greater than the most difficult item, **violin**. A large gap can is found between **butterfly** and **wasp**. **Pigeon** and **tortoise** have difficulty levels in close proximity to each other with no corresponding learners. The majority of the objects difficulty was situated between 0 and 1.

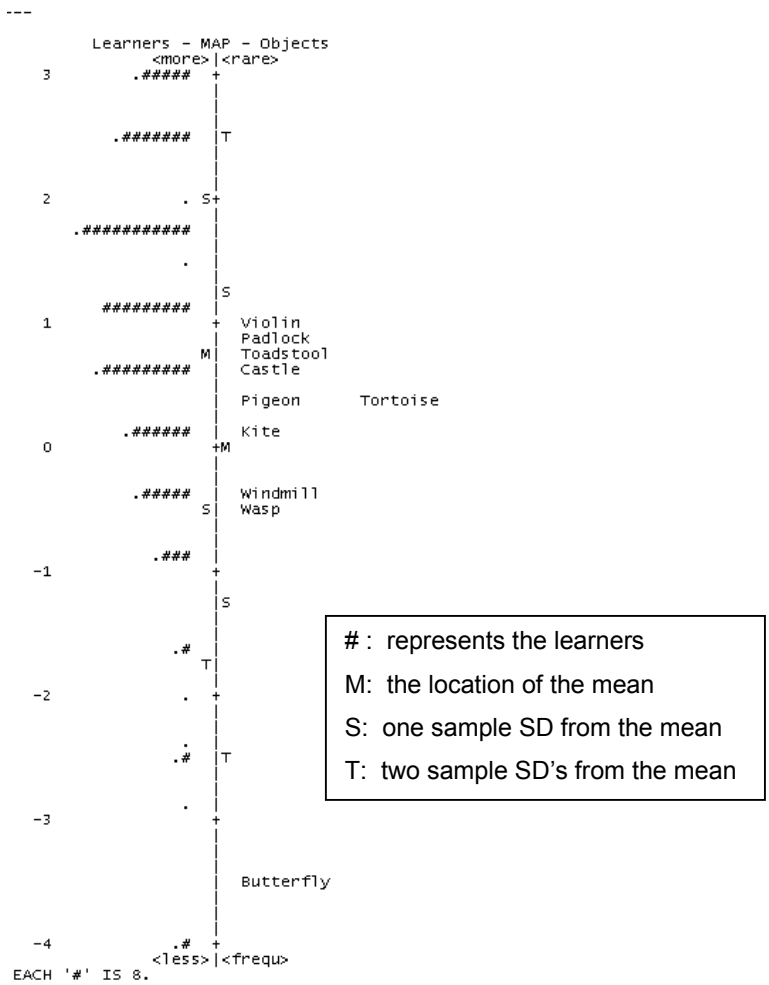


Figure 5.13: Item-Learner Map for Picture 2 - English group

The Item Development Pathway for Picture 2 shows how the objects were placed according to the results of the data from the English learners.

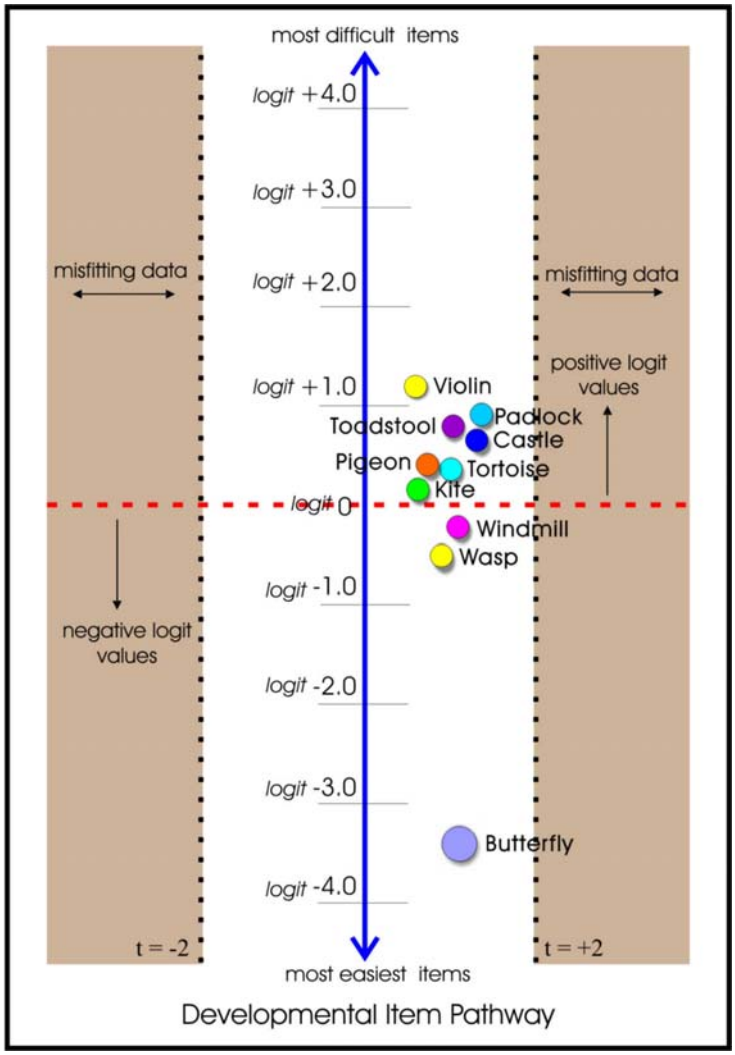


Figure 5.14: IDP for English group – Picture 2

The learners found the majority of objects slightly difficult with few exceptions as **butterfly**, **windmill** and **wasp** that were seen as easy items. The objects are not distributed equally, and as a result do not display great variation in difficulty. The majority of items were in the middle range. The easiest object to identify for Picture 2 for the English learners was **butterfly** and the most difficult was **violin**. The objects do not follow the expected predictions of the Rasch model. Unidimensionality is important and it indicates whether the items are working together to define a single construct based on the Item Development Pathway and other statistics. The above

IDP shows no misfitting items providing evidence that a single construct was measured.

Table 5.20 reflects the logit values of the objects in column 2. The objects are arranged according to order of difficulty. The Standard Error (SE) for the English results is given in Column 3 (see Appendix U).

Table 5.20: Object order for Picture 2 - English group

Original Order	English Order (logit)	Standard Error
1. Butterfly	1. Butterfly (-3.45)	0.17
2. Kite	4. Wasp (-0.49)	0.08
3. Castle	6. Windmill (-0.32)	0.08
4. WASP	2. Kite (+0.15)	0.07
5. Pigeon	7. Tortoise (+0.42)	0.07
6. Windmill	5. Pigeon (+0.43)	0.07
7. Tortoise	3. Castle (+0.58)	0.07
8. Violin	10. Toadstool (+0.79)	0.07
9. Padlock	9. Padlock (+0.86)	0.07
10. Toadstool	8. Violin (+1.01)	0.07

As can be seen from the above table, there are only 2 similarities between the original PIPS order of objects and the English learners' order of objects. The two objects, **butterfly** and **padlock**, are in the same order of difficulty as the original PIPS instrument although the rest of the objects do not follow the original order of difficulty. The Sepedi results are now explored.

5.5.4 Sepedi learners results for Baseline assessment of Picture 2

In Table 5.21 the Sepedi learners' results are displayed for Picture 2. The item-learner statistics, object statistics, Item-Learner Map, Item Development Pathway and the object order are provided (see Appendix V).

Table 5.21: Learner & Item statistics for Sepedi learners for Picture 2

	Mean *	**INFIT MNSQ****	**INFIT ZSTD*****	**OUTFIT MNSQ****	**OUTFIT ZSTD*****	**Separation Reliability**
Learners	0.90	0.93	0.00	0.84	0.00	0.53
Items	0.00	1.00	-0.20	1.00	-0.30	0.99

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for both learners and items was 0.84 and 1.00 respectively. The INFIT MNSQ values for both learners and items were 0.93 and 1.00 respectively. This reflects that the objects used were correctly targeted for the learners. The separation reliability for the learners was 0.53 which is slightly low reliability. The item separation reliability for the items was 0.99 indicating that the objects used in Picture 1 do have varying difficulty levels.

In Table 5.22 the object statistics are shown (see Appendix W).

Table 5.22: Object statistics for Sepedi learners for Picture 2

Objects	**Logits**	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Castle	0.5	1.21	6.80	1.34	6.50	0.47
Windmill	-0.43	1.14	3.30	1.26	3.50	0.36
Padlock	0.92	1.13	4.60	1.12	2.80	0.47
Toadstool	0.78	1.07	2.50	1.06	1.50	0.50
Tortoise	0.32	1.04	1.20	1.02	0.50	0.48
Butterfly	-3.07	1.02	0.20	1.02	0.20	0.44
Wasp	-0.44	0.97	-0.70	0.96	-0.60	0.56
Kite	0.11	0.86	-4.60	0.79	-4.40	0.63
Pigeon	0.38	0.80	-7.00	0.73	-6.20	0.66
Violin	0.94	0.79	-8.10	0.74	-7.30	0.66

There are no misfitting items for the Sepedi learners, as was the case with the Afrikaans and English scenarios presented. The OUTFIT MNSQ of 1.34 for **castle** indicates that unexpected observations by the learners on the item occurred.

The Item-Learner Map is shown for Picture 2 for the Sepedi group. Evidently seen from the Item-Learner Map are the learners' abilities and items difficulties are not targeted correctly. Ideally for every item difficulty there should be corresponding learner ability (Linacre, 2005). The Sepedi learners' abilities exceeded the difficulty of the items for Picture 2. The Item-Learner Map also clearly illustrates that the learners' abilities are considerably greater than the most difficult items, **violin** and **padlock**. A large gap can is found between **butterfly** and **wasp**. **Pigeon** and **tortoise**, as well as **padlock** and **violin** have difficulty levels in close proximity to each other. The majority of the objects difficulty was situated between 0 and 1.

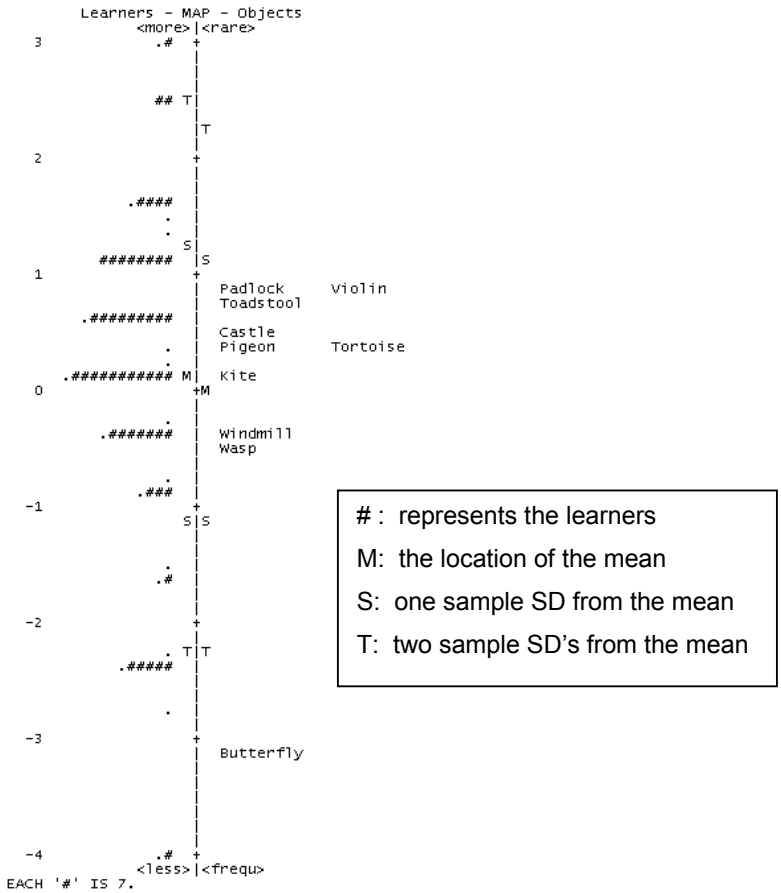


Figure 5.15: Item-Learner Map for Picture 2 - Sepedi group

In the Item Development Pathway the order of objects from the data received from the Sepedi learners is shown.

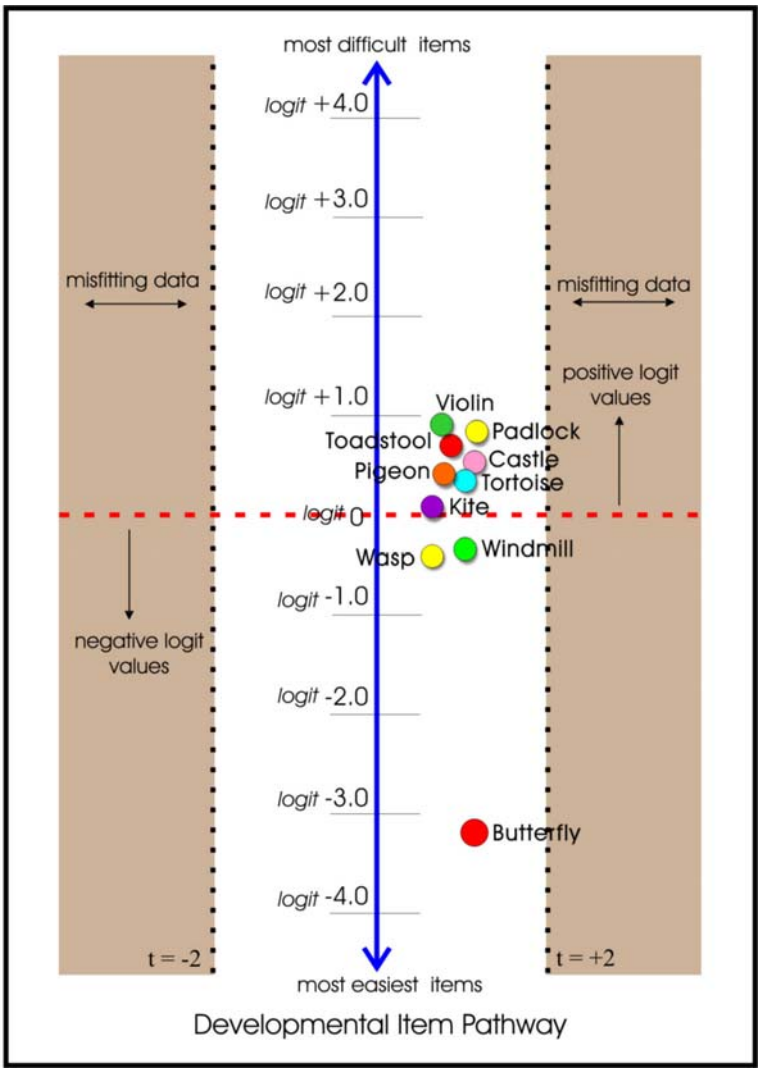


Figure 5.16: IDP for Sepedi group – Picture 2

The learners found the majority of objects slightly difficult with a few exceptions such as **butterfly**, **windmill** and **wasp** that were seen as easy items. The objects are not distributed equally, and as a result do not display great variation in difficulty. The majority of items were in the middle range. The easiest object to identify for Picture 2 for the Sepedi learners was **butterfly** and the most difficult was **violin**. The objects do not follow the expected predictions of the Rasch model. Unidimensionality being

an important assumption of the Rasch model indicates whether the items are workings together to define a single construct. In the above IDP there are no misfitting items which indicate that a single construct is being measured.

Table 5.23 reflects the logit values of the objects in column 2. The objects are arranged according to order of difficulty. The Standard Error (SE) for the English results is given in Column 3 (see Appendix X).

Table 5.23: Object order for Picture 2 - Sepedi group

Original Order	Sepedi Order	Standard Error
1. Butterfly	1. Butterfly (-3.07)	0.15
2. Kite	4. Wasp (-0.44)	0.08
3. Castle	6. Windmill (-0.43)	0.08
4. WASP	2. Kite (+0.11)	0.07
5. Pigeon	7. Tortoise (+0.32)	0.07
6. Windmill	5. Pigeon (+0.38)	0.07
7. Tortoise	3. Castle (+0.50)	0.07
8. Violin	10. Toadstool (+0.78)	0.07
9. Padlock	9. Padlock (+0.92)	0.07
10. Toadstool	8. Violin (+0.94)	0.07

As can be seen from the above table, there are only 2 similarities between the original PIPS order of objects and the Sepedi learners' order of objects. The two objects, **butterfly** and **padlock**, are in the same order of difficulty as the original PIPS instrument although the rest of the objects do not follow the original order of difficulty.

5.5.5 Summary of Picture 2 across all groups

Table 5.24 mirrors the order of the objects for all three language groups for Picture 2. As mentioned in Section 5.1 above, the cells shaded in **light green** represent the objects that follow the original PIPS instruments' difficulty order. The cells shaded in **light blue** represent the objects that follow the same difficulty order across all three language groups.

Table 5.24: Object order for Picture 2 for all three language groups

Original Order	Afrikaans	English	Pedi
1. Butterfly	1. Butterfly	1. Butterfly	1. Butterfly
2. Kite	6. Windmill	4. Wasp	4. Wasp
3. Castle	4. Wasp	6. Windmill	6. Windmill
4. Wasp	2. Kite	2. Kite	2. Kite
5. Pigeon	7. Tortoise	7. Tortoise	7. Tortoise
6. Windmill	5. Pigeon	5. Pigeon	5. Pigeon
7. Tortoise	3. Castle	3. Castle	3. Castle
8. Violin	10. Toadstool	10. Toadstool	10. Toadstool
9. Padlock	9. Padlock	9. Padlock	9. Padlock
10. Toadstool	8. Violin	8. Violin	8. Violin

Only two objects follow the original difficulty order across all three language groups namely **butterfly** and **padlock**.

Although there were differences in the order for all three groups, a few similarities can also be seen:

The similarities:

- Across all three language groups**: butterfly, kite, tortoise, pigeon, castle, toadstool, padlock** and **violin.**
- Afrikaans and Sepedi learners were: **butterfly, kite, tortoise, pigeon, castle, toadstool, padlock** and **violin.**
- Afrikaans and English learners were: **butterfly, kite, tortoise, pigeon, castle, toadstool, padlock** and **violin.**
- English and Sepedi learners were**: butterfly, wasp, windmill, kite, tortoise, pigeon, castle, toadstool, padlock** and **violin.**

For Picture 2 it is rather surprising that the three groups had such a large number of objects in the same order of difficulty. There was a slight resemblance of the object difficulty order for the groups compared to the original PIPS order.

Next the individual language groups' performance is discussed starting in alphabetical order with Afrikaans, English and Sepedi.

5.6 PICTURE 3 – PICTURE OF BEDROOM

In the last picture of the Picture Vocabulary Test, a child's bedroom is portrayed. The learners had to identify 5 different objects ranging from easy to difficult. These objects were **yacht**, **cash**, **microscope**, **jewellery** and **saxophone** (see Appendix Y). This final picture had the most difficult objects to identify (for UK learners) of all three pictures according to the PIPS instrument. The results are given of the order of the objects for the three different language groups, Afrikaans, English and Sepedi.

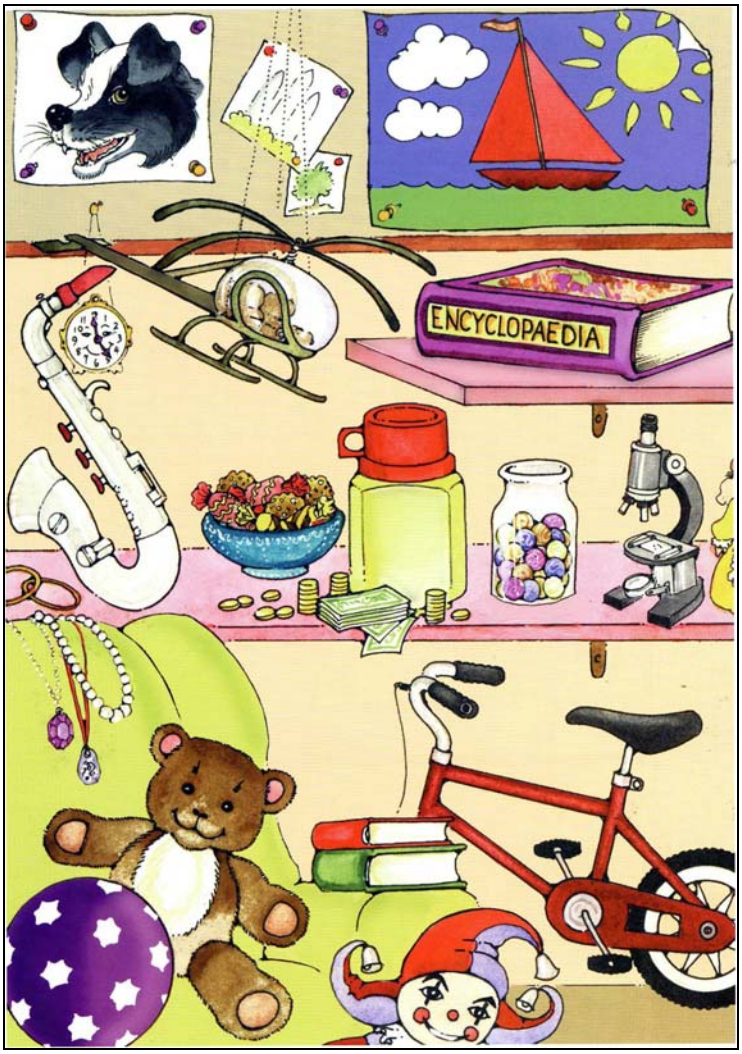


Figure 5.17: Picture 3

5.6.1 Findings across all three language groups for Picture 3

In Table 5.25 the results are displayed of all three language groups for Picture 3. The item-learner statistics, object statistics, Item-Learner Map, and the object order are provided (see Appendix Z).

Table 5.25: Learner & Item statistics for all language groups for Picture 3

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	0.38	1.01	0.10	0.91	0.10	0.00
Items	0.00	1.01	0.30	0.92	-0.60	1.00

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for both learners and items was 0.91 and 0.92 respectively, which is slightly lower than 1 indicating unexpected responses, the learners abilities did not match the items difficulties. The INFIT MNSQ values for both learners and items were 1.01. The separation reliability for the learners was 0.00 which is a low reliability showing that the learners' abilities weren't matched. The item separation reliability for the items was 1.00 indicating that the objects used in Picture 3 do have varying difficulty levels.

In Table 5.26 the object statistics are shown (see Appendix AA).

Table 5.26: Object statistics for all three languages for Picture 3

Objects	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Jewellery	1.14	3.70	1.11	2.00	0.54
Microscope	1.05	1.60	1.01	0.30	0.60
Saxophone	1.04	1.20	1.00	0.00	0.61
Cash	0.93	-1.10	0.67	-2.30	0.55
Yacht	0.86	-4.00	0.79	-3.10	0.67

The items included in this section of the assessment all adhere to the predetermined criteria for fit.

The Item-Learner Map is shown for Picture 3 for all language groups. Evidently seen from the Item-Learner Map is that the learners' abilities and items difficulties are not targeted correctly. Ideally every item difficulty should correspond with learner ability (Linacre, 2005). The learners' abilities exceeded the difficulty of the items for Picture 3. The Item-Learner Map also clearly illustrates that the learners' abilities are considerably greater than the most difficult item which is **saxophone**. Large gaps can be found between the objects with the exception of **saxophone** and **yacht**. In the Item-Learner Map a fair amount of learners' from the three language groups abilities exceed the difficulty of the items for Picture 3.

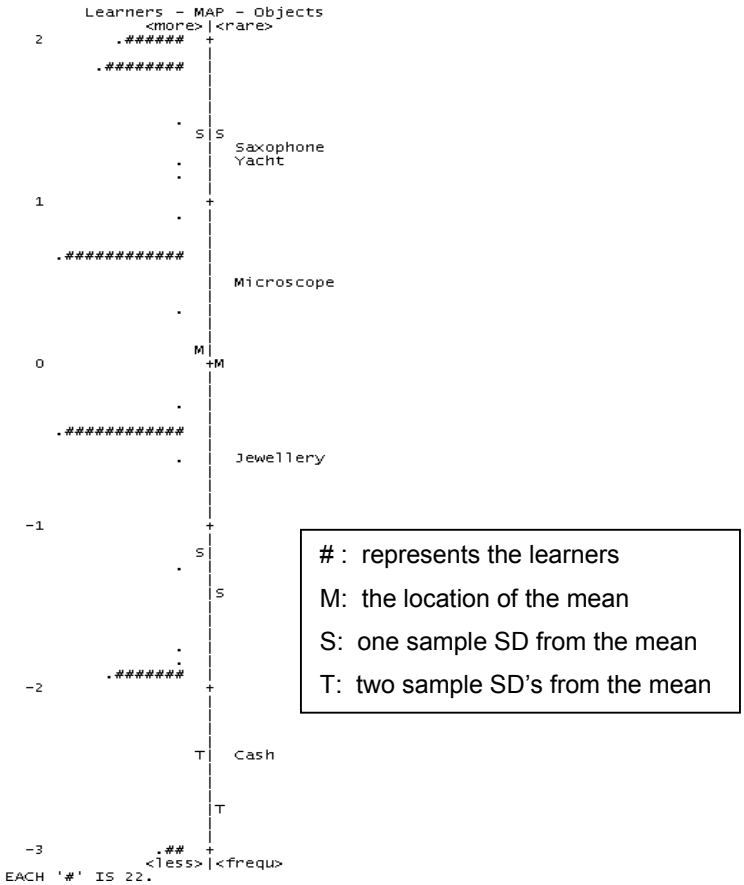


Figure 5.18: Item-Learner Map for all languages– Picture 2

5.6.2 Afrikaans learners results for Baseline assessment of Picture 3

In Table 5.27 the results are displayed of the objects performance for the Afrikaans learners. The item-learner statistics, object statistics, Item-Learner Map, Item Development Pathway and the object order are provided (see Appendix BB).

Table 5.27: Learner & Item statistics for Afrikaans learners for Picture 3

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	0.33	1.05	0.10	1.01	0.10	0.75
Items	1.97	0.98	-0.40	0.95	-0.50	0.99

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for both learners and items were 1.01 and 0.95. The INFIT MNSQ values for both learners and items were 1.05 and 0.98 respectively. The separation reliability for the learners was 0.75 which is low indicating that the ability of the learners was not accurately measured. The item separation reliability for the items was 0.99 indicating that the objects used in Picture 3 do have varying difficulty levels.

In Table 5.28 the object statistics are shown (see Appendix CC).

Table 5.28: Object statistics for Afrikaans learners for Picture 3

Objects	Logits	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Jewellery	1.47	1.13	4.30	1.13	2.40	0.40
Saxophone	2.99	1.10	2.80	1.07	1.30	0.49
Microscope	2.31	0.97	-1.00	0.94	-1.20	0.54
Cash	0.17	0.89	-2.20	0.75	-2.50	0.43
Yacht	2.93	0.82	-5.80	0.87	-2.50	0.64

When inspecting the OUTFIT and INFIT MNSQ it was found that all the item statistics were within the 0.7 and 1.3 range.

The Item-Learner Map is shown for Picture 3 for the Afrikaans group. Evidently seen from the Item-Learner Map is that the learners' abilities and items difficulties are not targeted correctly. The Afrikaans learners' abilities exceeded the difficulty of the items for Picture 3 to a large extent. The Item-Learner Map also clearly illustrates that the learners' abilities are considerably greater than the most difficult item which is **saxophone**. Large gaps can be found between the objects with the exception of **saxophone** and **yacht**. A fair amount of Afrikaans learners' abilities exceeded the difficulty of the items for Picture 3.

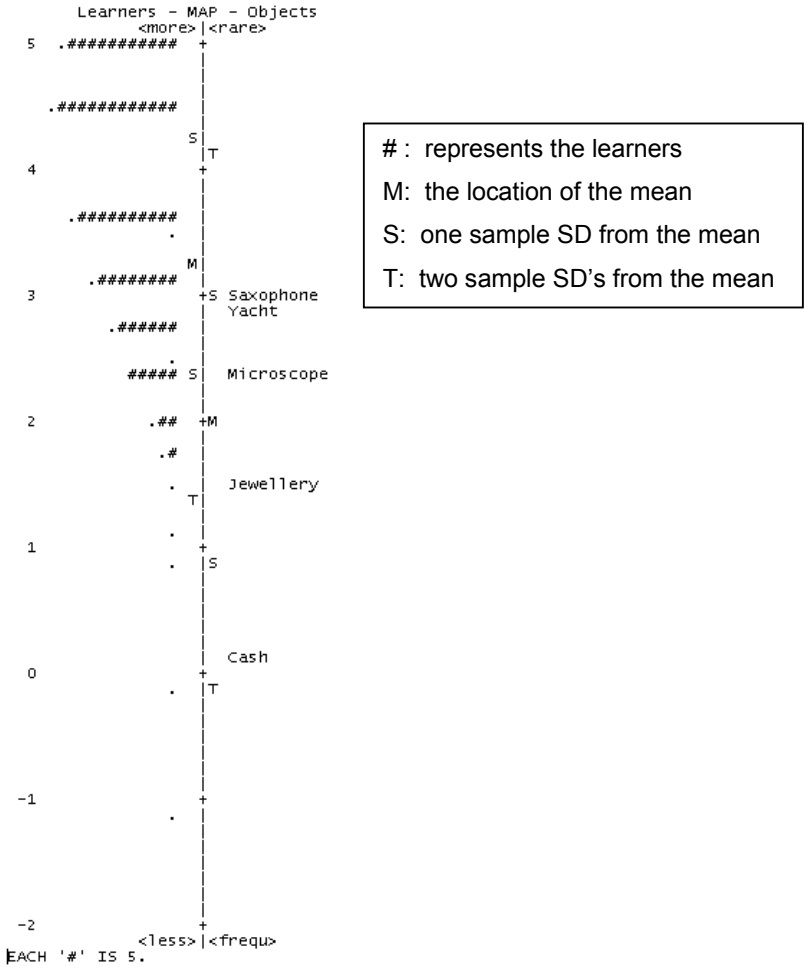


Figure 5.19: Item-Learner Map for Picture 3 - Afrikaans group

The Item Development Pathway demonstrates the order of the objects for Picture 3 for the Afrikaans learners.

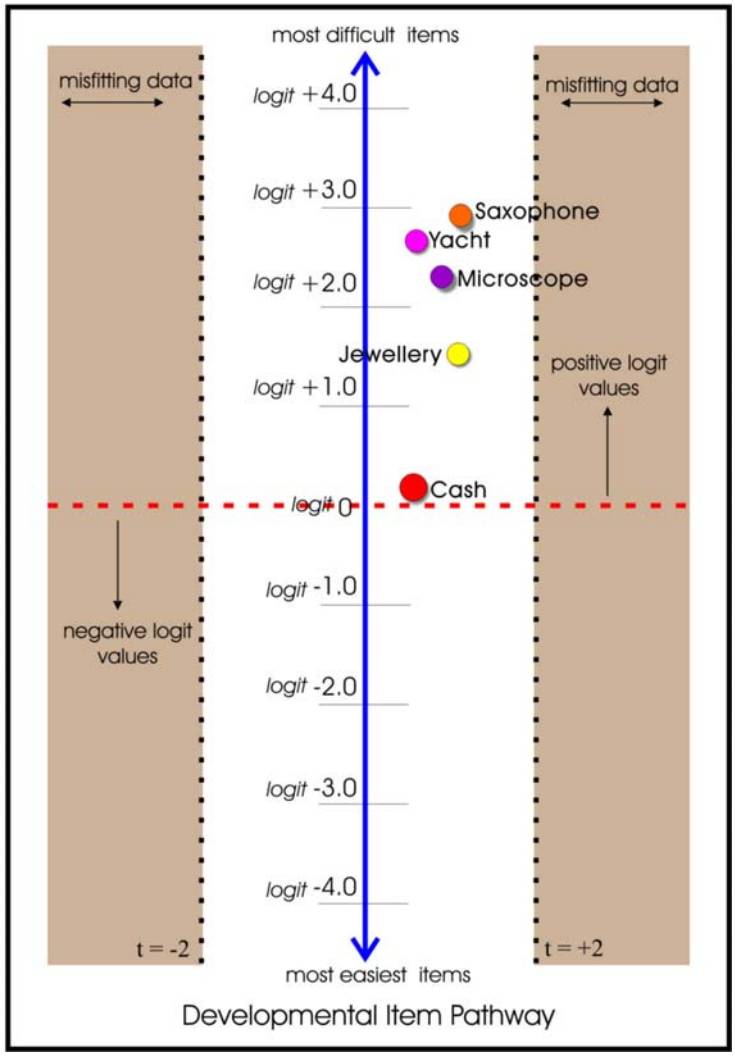


Figure 5.20: IDP for Afrikaans group – Picture 3

The learners found the majority of objects difficult. The objects were distributed but not equally, and as a result do not display large variation in difficulty. The majority of items were in the middle range. The easiest object to identify for Picture 3 for the Afrikaans learners was **cash** and the most difficult was **saxophone**. The above Item Development Pathway shows that there are no misfitting items which give an indication that a single construct is measured.

Table 5.29 reflects the logit values of the objects in column 2. The objects are arranged according to order of difficulty. The Standard Error (SE) for the English results is given in Column 3 (see Appendix CC).

Table 5.29: Object order for Picture 3 - Afrikaans group

Original Order	Afrikaans Order (logit)	Standard Error
1. Yacht	2. Cash (+0.17)	0.09
2. Cash	4. Jewellery (+1.47)	0.07
3. Microscope	3. Microscope (+2.31)	0.07
4. Jewellery	1. Yacht (+2.93)	0.07
5. Saxophone	5. Saxophone (+2.99)	0.07

As can be seen from the above table, there are only 2 similarities between the original PIPS order of objects and the Afrikaans learners' order of objects. The two objects, **microscope** and **saxophone**, are in the same order of difficulty as the original PIPS instrument although the rest of the objects do not follow the original order of difficulty.

As clearly indicated the items did not perform in the manner expected for Picture 3 for the Afrikaans learners.

5.6.3 English learners results for Baseline assessment of Picture 3

In Table 5.30 the results are displayed of the objects performance for the Afrikaans learners. The item-learner statistics, object statistics, Item-Learner Map, Item Development Pathway and the object order are provided (see Appendix EE).

Table 5.30: Learner & Item statistics for English learners for Picture 3

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	0.80	0.96	0.00	0.85	0.10	0.75
Items	1.88	0.96	-1.10	0.89	-1.80	0.99

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for both learners and items were 0.85 and 0.89 and were below 1 indicating that the learners' abilities did not match the items difficulties. The INFIT MNSQ values for both learners and items were 0.96, close to the expected value. This reflects that the objects used are correctly targeted for the learners. The separation reliability for the learners was 0.75 is low which could be because the learners abilities were not matched. The item separation reliability for the items was 0.99 indicating that the objects used in Picture 3 do have varying difficulty levels.

In Table 5.31 the object statistics are shown (see Appendix FF).

Table 5.31: Object statistics for English learners for Picture 3

Objects	Logits	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Jewellery	1.39	1.12	3.90	1.11	1.90	0.40
Saxophone	2.85	1.06	1.80	1.04	0.70	0.51
Microscope	2.26	0.94	-2.00	0.90	-2.30	0.56
Cash	0.04	0.90	-2.00	0.70	-2.90	0.42
Yacht	2.86	0.79	-7.20	0.70	-6.50	0.67

All of the items for this section are in accordance with the criteria for fit.

The Item-Learner Map is shown for Picture 3 for the English group. The Item-Learner Map shows that the learners' abilities and items difficulties are not targeted correctly.

The English learners' abilities exceeded the difficulty of the items for Picture 3 to a large extent. The Item-Learner Map also clearly illustrates that the learners' abilities are considerably greater than the most difficult items which is saxophone and yacht. The gaps found between the objects are not too large with the exception of **saxophone** and **yacht**. Although these gaps need to be investigated further. A fair proportion of English learners' abilities exceeded the difficulty of the items for Picture 3.

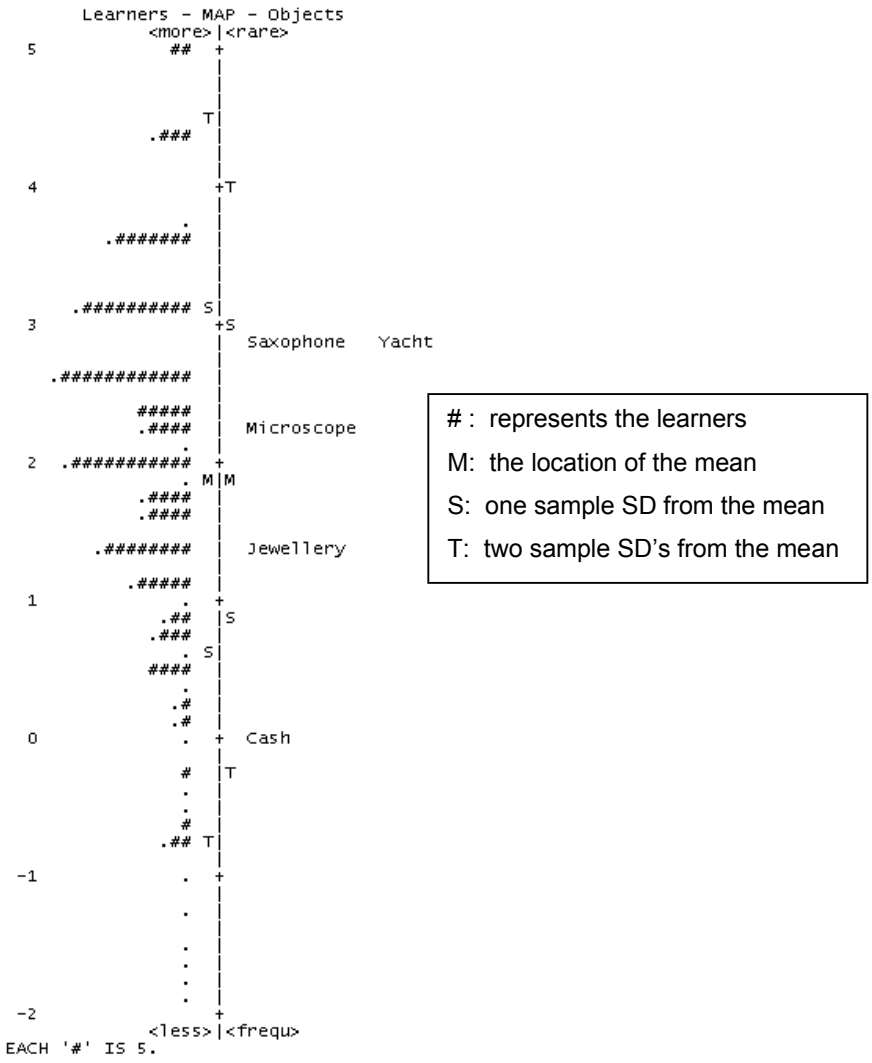


Figure 5.21: Item-Learner Map for Picture 3 - English group

The order of the objects is displayed in the Item Development Pathway for Picture 3 for the English learners.

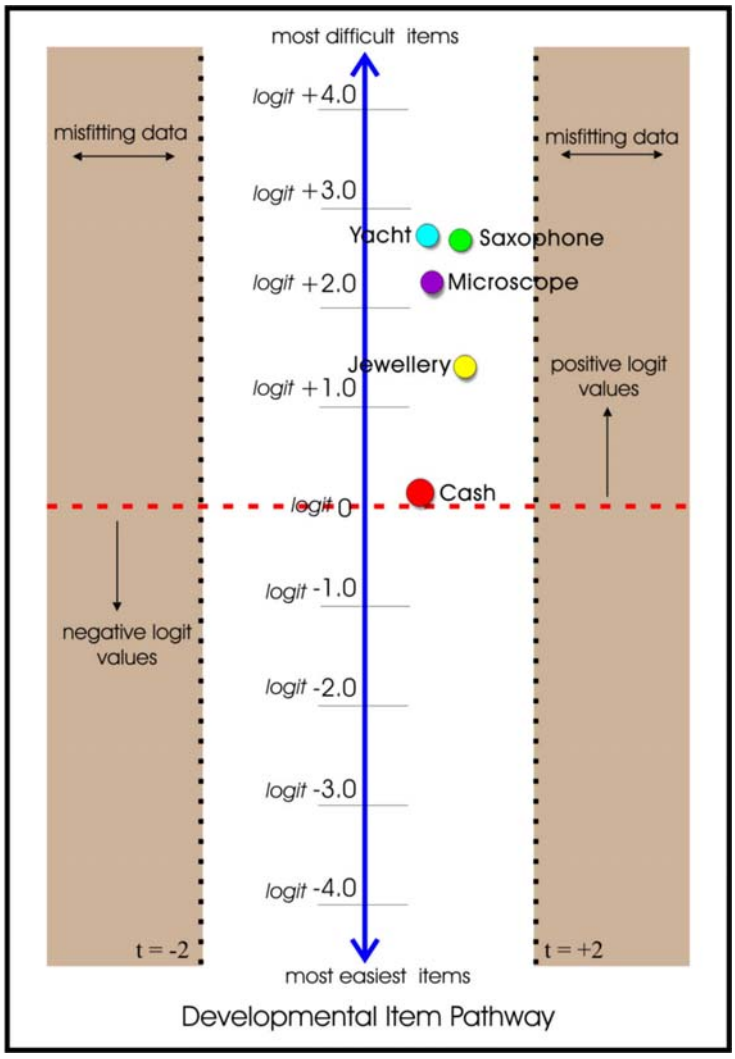


Figure 5.22: IDP for English group – Picture 3

The learners found the majority of objects difficult. The objects are distributed but not equally and as a result do not display the desired variation in difficulty. **Yacht**, **saxophone** and **microscope** have very slight variations in difficulty. The easiest object to identify for Picture 3 for the English learners was **cash** and the most difficult was **yacht**. The objects do not follow the expected predictions of the Rasch model. The above Item Development Pathway shows no misfitting items which gives an indication that a single construct was measured.

Table 5.32 reflects the logit values of the objects in column 2. The objects are arranged according to order of difficulty. The Standard Error (SE) for the English results is given in Column 3 (see Appendix GG).

Table 5.32: Object order for Picture 3 - English group

Original Order	English Order (logit)	Standard Error
1. Yacht	2. Cash (+0.4)	0.09
2. Cash	4. Jewellery (+1.39)	0.07
3. Microscope	3. Microscope (+2.26)	0.07
4. Jewellery	5. Saxophone (+2.85)	0.07
5. Saxophone	1. Yacht (+2.86)	0.07

As can be seen from the above table, there is only 1 similarity between the original PIPS order of objects and the English learners' order of objects namely, **microscope**. The rest of the objects did not follow the original order of difficulty. Taking a look at the reliability of the objects in the Picture Vocabulary Test, it showed that the SE for the objects was very small.

5.6.4 Sepedi learners results for Baseline assessment of Picture 3

In Table 5.33 the results are displayed of the learners and objects performance for Picture 3. The item-learner statistics, object statistics, Item-Learner Map, Item Development Pathway and the object order are provided (see Appendix HH).

Table 5.33: Learner & Item statistics for Sepedi learners for Picture 3

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	-0.61	1.02	0.00	0.95	0.00	0.00
Items	1.77	0.95	-1.30	0.91	-1.20	1.00

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

The OUTFIT MNSQ for both learners and items were 0.95 and 0.91 respectively. The INFIT MNSQ values for both learners and items were 1.02 and 0.95. The separation reliability for the learners was 0.00 which is low that indicates that the learners abilities were not matched with the items difficulties. The item separation reliability for the items was 1.00 indicating that the objects used in Picture 1 do have varying difficulty levels.

In Table 5.34 the object statistics are shown (see Appendix II).

Table 5.34 Object statistics for Sepedi learners for Picture 3

Objects	Logits	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Jewellery	1.32	1.11	3.50	1.08	1.60	0.42
Saxophone	2.79	1.04	1.20	1.00	0.10	0.53
Microscope	2.12	0.95	-1.80	0.92	-1.90	0.55
Cash	-0.09	0.89	-2.20	0.70	-2.80	0.42
Yacht	2.69	0.79	-7.00	0.85	-3.10	0.66

All of the items for this section adhere to the fit criteria.

The Item-Learner Map is shown for Picture 3 for the Sepedi group. The Item-Learner Map shows that the learners' abilities and items difficulties are not targeted correctly. The Sepedi learners' abilities nearly matched the difficulty of the items for Picture 3.

The gaps found between the objects are not equidistant with the exception of **saxophone** and **yacht**. A small proportion of Sepedi learners' abilities exceeded the difficulty of the items for Picture 3. The most difficult item was **saxophone**.

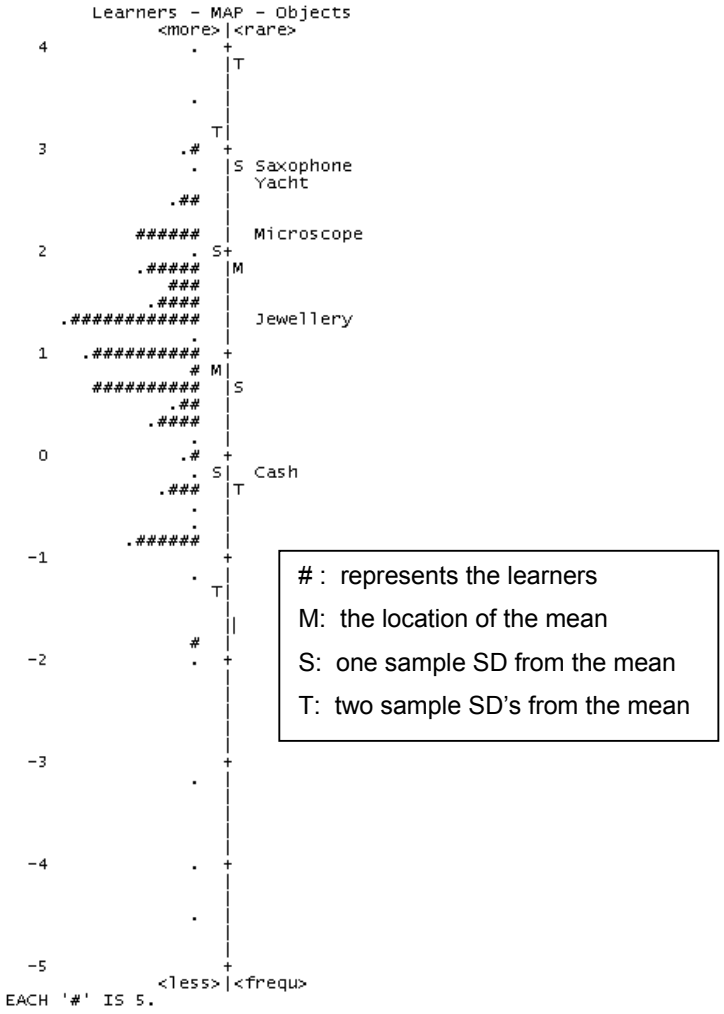


Figure 5.23: Item-Learner Map for Picture 3 - Sepedi group

The Sepedi learners' object difficulty order is illustrated in the Item Development Pathway.

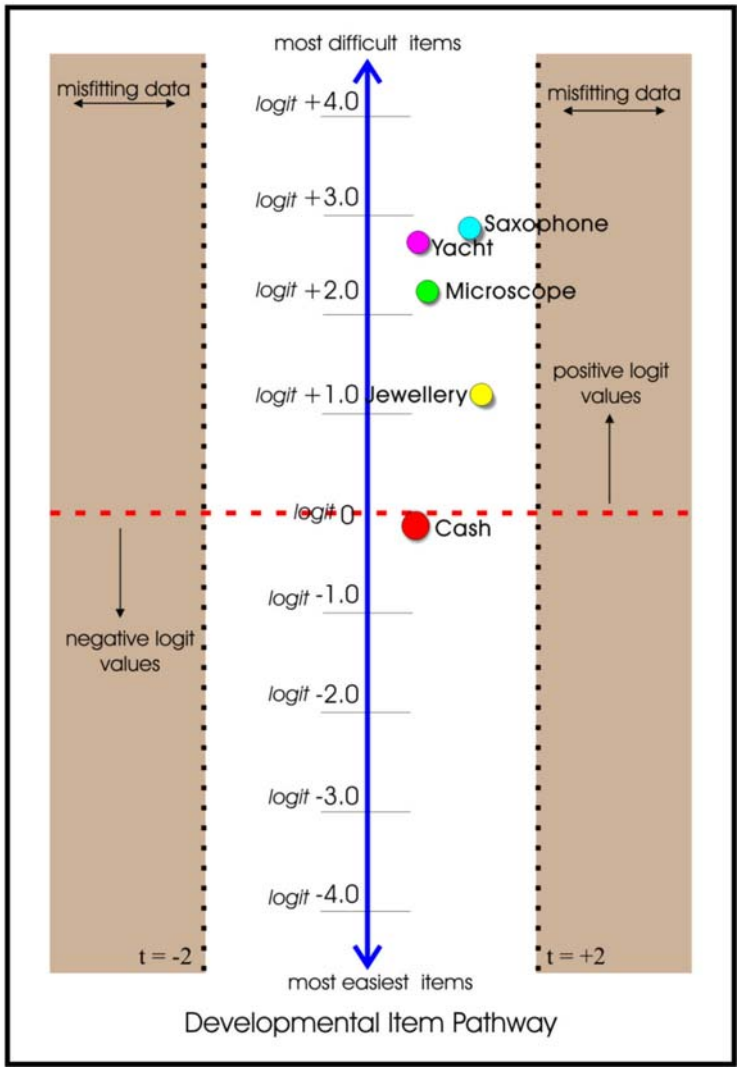


Figure 5.24: IDP for Sepedi group – Picture 3

The learners found the majority of objects difficult with the exception of **cash** that was experienced as fairly easy. The objects are distributed but not in equal increments, and as a result do not display great variation in difficulty. **Saxophone** and **yacht** had slight variations in difficulty. The easiest object to identify for Picture 3 for the Sepedi learners was **cash** and the most difficult was **saxophone**. The objects do not follow the expected predictions of the Rasch model. No misfitting items are shown in the

Item Development Pathway above which indicated that a single construct was measured.

Table 5.35 reflects the logit values of the objects in column 2. The objects are arranged according to order of difficulty. The Standard Error (SE) for the English results is given in Column 3 (see Appendix JJ).

Table 5.35: Object order for Picture 3 - Sepedi group

Original Order	Sepedi Order	Standard Error
1. Yacht	2. Cash (-0.90)	0.09
2. Cash	4. Jewellery (+1.32)	0.07
3. Microscope	3. Microscope (+2.12)	0.07
4. Jewellery	1. Yacht (+2.69)	0.07
5. Saxophone	5. Saxophone (+2.79)	0.07

As can be seen from the above table, there are only 2 similarities between the original PIPS order of objects and the English learners' order of objects namely, **microscope** and **saxophone**. The rest of the objects do not follow the original order of difficulty.

5.6.5 Summary of Picture 3 across all groups

Table 5.36 depicts the difficulty order for all three language groups together with the original order.

Table 5.36: Object order for Picture 3 – for all three groups

Original Order	Afrikaans Order	English Order	Sepedi Order
1. Yacht	2. Cash	2. Cash	2. Cash
2. Cash	4. Jewellery	4. Jewellery	4. Jewellery
3. Microscope	3. Microscope	3. Microscope	3. Microscope
4. Jewellery	1. Yacht	5. Saxophone	1. Yacht
5. Saxophone	5. Saxophone	1. Yacht	5. Saxophone

Only one object follows the original difficulty order across all three language groups namely **microscope**.

Although there were differences in the order for all three groups, a few similarities can also be seen:

- Across all three language groups: **cash** and **jewellery**
- Afrikaans and Sepedi learners were: **cash, jewellery, microscope, yacht** and **saxophone.**
- Afrikaans and English learners were: **cash, jewellery** and **microscope**.
- English and Sepedi learners were: **cash, jewellery and microscope**.

Taking all three pictures into consideration, the pictures that will be most useful across all three language groups would be Picture 2 and Picture 3 with a few modifications to the objects and their order. These two pictures could definitely be worth considering for future assessments across the three language groups.

Next the individual language groups' performance will be discussed starting in alphabetical order.

5.7 ALL OBJECTS DIFFICULTY ORDER ACROSS ALL LANGUAGES

For future assessment purposes it is important to consider how the entire group of objects performed across all three language groups. This helps to provide guidance on how the objects should be arranged to follow the correct difficulty order. In table 5.37 the performance results are displayed for the learners and the items for all three groups. The item-learner statistics, object statistics, Item-Learner Map, and the object order are provided (see Appendix KK).

Table 5.37: Statistics for all objects and learners for all pictures

	Mean*	INFIT MNSQ**	INFIT ZSTD***	OUTFIT MNSQ**	OUTFIT ZSTD***	Separation Reliability
Learners	2.01	1.01	0.00	0.95	0.20	0.75
Items	-1.52	1.00	0.10	0.92	-0.70	0.99

* Mean was set at 1 logit

** Criteria: As close to 1 as possible

*** Criteria: Between +2 and -2

This shows that there are no extremely difficult or easy items or learners that performed extremely well or poor for all three pictures. The INFIT MNSQ values for both learners and items were 1.01 and 1.00 respectively. This reflects that the objects used were correctly targeted for the learners. The separation reliability for the learners was 0.75 which indicates that the abilities of the learners were not accurately matched. The item separation reliability for the items was 0.99 indicating that the objects used in the pictures do have varying difficulty levels.

In Table 5.38 the entire object statistics are shown for all the learners (see Appendix LL).

Table 5.38: Object statistics for all learners for all pictures & all objects

Objects	Model S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE Correlation
Castle	0.07	1.20	6.20	1.43	6.70	0.43
Padlock	0.07	1.22	7.10	1.18	3.90	0.39
Cherries	0.27	1.16	0.80	0.97	0.10	0.20
Pan	0.08	1.16	3.80	1.06	0.80	0.45
Toadstool	0.07	1.15	4.90	1.15	3.20	0.42
Jewellery	0.07	1.14	4.40	1.14	2.60	0.40
Windmill	0.08	1.07	1.60	1.09	1.10	0.36
Saxophone	0.07	1.09	2.50	1.07	1.10	0.50
Tortoise	0.07	1.02	0.60	1.08	1.50	0.45
Wasp	0.08	1.04	1.10	1.04	0.50	0.48
Microscope	0.07	0.97	-1.10	0.94	-1.40	0.55
Bowl	0.11	0.90	-1.30	0.56	-2.70	0.47
Cash	0.09	0.90	-1.80	0.70	-2.70	0.41
Kite	0.07	0.87	-4.10	0.83	-2.90	0.59
Pigeon	0.07	0.83	-5.80	0.75	-4.70	0.62
Butterfly	0.17	0.82	-1.40	0.29	-3.40	0.39
Yacht	0.07	0.80	-6.30	0.70	-5.80	0.66
Violin	0.07	0.71	-9.90	0.64	-9.40	0.68

All of the items fall within the predetermined criteria of 0.7 and 1.3. The exception is the OUTFIT MNSQ for **butterfly** of 0.29. However the INFIT MNSQ is 0.82 and when considering construct validity the INFIT MNSQ is more important.

An essential exploration is the item-learner targeting across all objects for all languages. This was explored by means of an Item-Learner Map. An Item-Learner Map is shown for all the objects across all the language groups. Clearly seen from the Item-Learner Map is that the learners' abilities and items difficulties are not

targeted correctly. Ideally for every item difficulty there should be corresponding learner ability (Linacre, 2005). The map also clearly illustrates that the learners' abilities are higher than the most difficult items, **saxophone** and **yacht**. Very few items have difficulties that are equally dispersed along the variable. Ideally there should be objects which get progressively more difficulty with equal gaps between them as opposed to the large gaps or very small gaps found between the objects. The majority of items are groups near the centre of the Item-Learner Map. Many of the items are in close difficulty range of each other. The items do not fit the Rasch model.

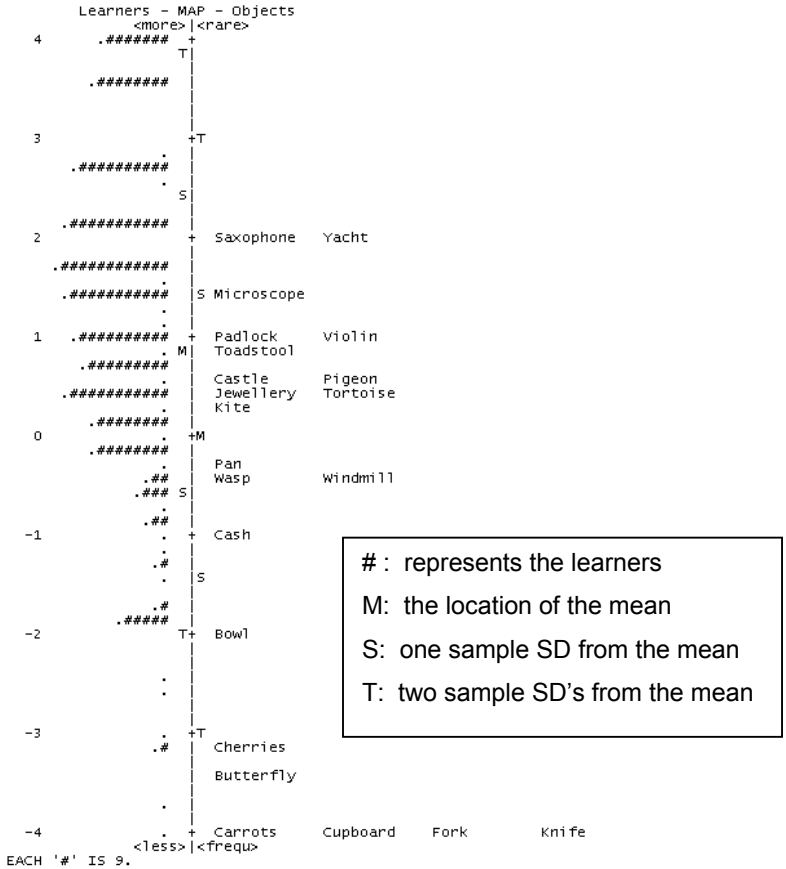


Figure 5.25: Item-Learner Map for Picture 3 for all languages

Table 5.39 depicts the difficulty order for all three language groups together with the original order.

Table 5.39: Order for all the objects across the three language groups

Afrikaans	English	Sepedi
Knife	Knife	Knife
Fork	Fork	Carrots
Carrots	Carrots	Fork
Cupboard	Cupboard	Cupboard
Cherries	Butterfly	Butterfly
Butterfly	Bowl	Bowl
Bowl	Cherries	Cherries
Cash	Cash	Cash
Pan	WASP	Pan
Windmill	Pan	Windmill
Wasp	Windmill	Wasp
Kite	Kite	Kite
Tortoise	Jewellery	Tortoise
Jewellery	Tortoise	Pigeon
Pigeon	Pigeon	Jewellery
Castle	Castle	Castle
Toadstool	Toadstool	Toadstool
Padlock	Padlock	Padlock
Violin	Violin	Violin
Microscope	Microscope	Microscope
Yacht	Saxophone	Yacht
Saxophone	Yacht	Saxophone

There were 22 objects used altogether in the Picture Vocabulary Test. Out of these 22 objects 9 were in the same order of difficulty for all three language groups. This resulted in 41% of the objects following the same difficulty order for all three groups. These objects were: **knife cupboard, cash, kite, castle, toadstool, padlock, violin** and **microscope**. However, if careful consideration is given to the objects that differ in order, it comes to attention that the objects difficulty orders are closely related. It seems to be a matter of the objects being swapped around for example the Afrikaans order would be **yacht - saxophone** and the English order would be **saxophone - yacht**. Seen from this light the differences in order are minor and not as drastic as difference in order from the original instrument and those of the three language groups.

For the Afrikaans and English learners 12 of the objects followed the same order of difficulty. Consequently, 54% of the objects were on the same level of difficulty for both these languages as seen by the objects shaded in **light green**.

Similarly 12 of the objects between the English and Sepedi learners' object difficulty order were also the exact same. However, these objects did not follow the same difficulty order as the Afrikaans and English groups. The Sepedi and English object difficulty order also resulted in 54% of the objects following the same difficulty order (objects highlighted in **light green**).

The groups that had the most similarities with the difficulty order of the objects were the Afrikaans and Sepedi group. For this group 15 out of the 22 objects followed the exact same order of difficulty. This resulted in 68% of the objects following the exact same order for the two language groups.

Turning the focus towards the objects used in the Picture Vocabulary Test, and not the pictures on their own a different deduction can be made. It becomes clearer that that there are a number of objects that can be used across the three different language groups. At present the inferences made about the results of the Picture Vocabulary Test, cannot be valid. The validity level for the Picture Vocabulary Test is not high and will have to be given attention.

When objects are chosen carefully, with thought and contemplation, these specific objects can be used across all three language groups. Instead of having three different orders for the objects for each picture and language group, the objects can be selected and arranged to suit all three languages for the Picture Vocabulary Test, used in the instrument. But careful consideration has to be given to the Standard Error which provides an idea of the uncertainty associated with estimates. Once suitable objects are identified an increase in the reliability of the objects and an increase in the validity level can be expected.

In order to ensure that the objects chosen are good choices a Differential Item Functioning (DIF) analysis has to be done for each object for all three language groups. The results of the DIF analysis are discussed.

5.8 DIFFERENTIAL ITEM FUNCTIONING (DIF) ANALYSIS

Differential Item Functioning (DIF) is a vital source to help identify bias in assessments across dissimilar groups, thereby helping to improve upon the items found in an assessment that displays bias. Items that give different success rates across two or more groups display DIF (Huang, Church & Katigbak, 1997; Tennant & Pallant, 2007). When items do not perform in the same way across different groups that have the same abilities or traits, DIF occurs, which means that there is a difference in the statistical properties of items and then the items operate invariantly (Andrich, 2004). Through DIF analysis, a statistical procedure, the researcher is able to monitor whether the level of validity and fairness of the assessment is jeopardised by biased items. DIF can be uniform in which all ability groups are equally impacted or non-uniform where one group is impacted more than the other groups. DIF could have different meanings namely (Linacre, 2005):

- That one group is performing at the usual ability level and the other group is performing better than usual.
- That one group is performing at the usual ability level and the other group is performing worse than usual.
- That the item is difficult for one group but more difficult for another group.
- That the item is difficult for one group but easier for another group

A graphical representation is given of how each object functioned across the three language groups. Three different lines can be seen on the graph. The blue line represents the English group, the pink/orange line represents the Afrikaans group and the **light green** line represents the Sepedi group (see Appendix MM for all the DIF graphs).

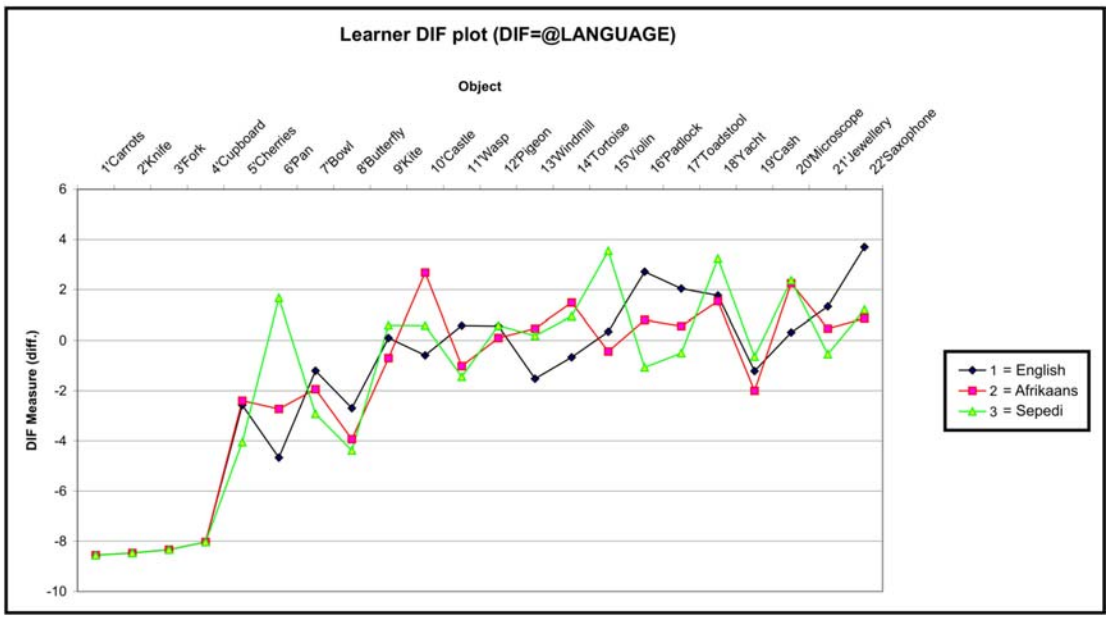


Figure 5.26: Differential Item Functioning Graph

How the objects perform across the 3 language groups is a crucial criterion when deciding what objects should be selected in a Picture Vocabulary Test. Or what objects need to be carefully reconsidered for the Picture Vocabulary Test. As can be seen from the above graph there are some objects that fall on the same level of difficulty for all three groups. But there are also objects that appear to be problematic, in other words they are functioning different from the other groups.

The most noticeable objects being **pan** and **violin** for the Sepedi learners, **castle** for the Afrikaans learners, and **saxophone** for the English learners. These four objects from these language groups are functioning very different when compared to the other language groups. Although it is important to investigate these objects a few facts about DIF have to be taken into consideration. None of the groups had the exact same number of learners; the number of learners differed. There were 355

Afrikaans learners, 562 English learners and 444 Sepedi learners. Furthermore, for a DIF analysis to be successful in selecting objects that aren't performing optimally in certain groups the sample needs to consist of thousands of learners to be able to accurately determine the difficulties of the objects (Linacre, 2009). In the Picture Vocabulary Test, the DIF also appears to be smaller for the other items. As these four items are exhibiting non-inform DIF as one group is impacted more than the other two groups. Furthermore, the DIF effect size is greater than 0.5 logits for these items and therefore further investigation is warranted (Appendix MM).

The Item Characteristic Curves (ICC) was given for the four objects that were not performing correctly according to the DIF analysis. An ICC is a visual representation of the learner's ability and the item's characteristics. An ICC has two asymptotes, the upper asymptote is on the vertical axis at 1.0 and the lower asymptote never reaches 0. The probability of a correct response to an item by the learner is a continually increasing curve (de Beer, 2004). ICC's are discussed in detail in the previous chapter under Section 4.2.6.

In the ICC's displayed the red ICC line indicates how the objects are supposed to perform. The blue ICC line represents the actual performance of the various objects.

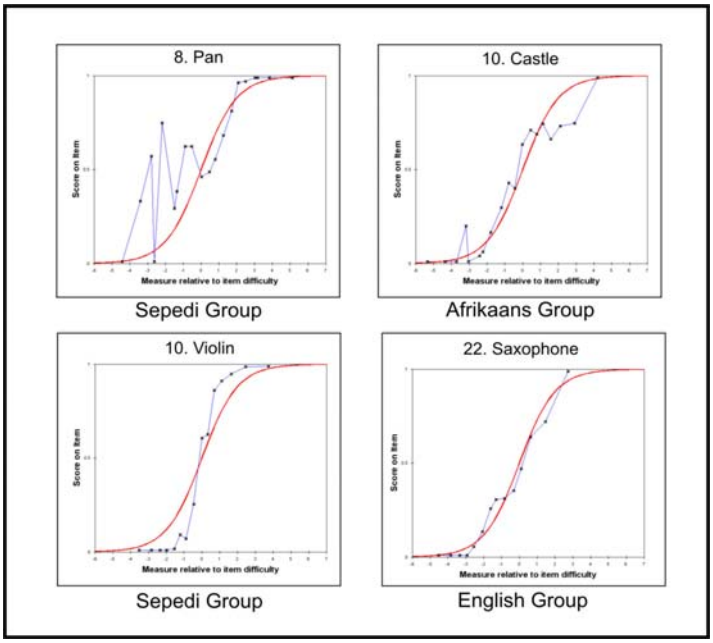


Figure 5.27: Item Characteristic Curves

Pan:

When looking at **pan's** ICC in blue for the Sepedi group and compare it to the red ICC curve it is very erratic. The characteristics of the object seem to jump up and down and do not follow the expected ICC. Only near the end of the curve does the object start to function as it should by almost following the red ICC line. This indicates that the object is not performing as it should for the Sepedi learners compared to the other language groups.

Castle:

For the Afrikaans learners the blue ICC line also did not follow the expected ICC but deviates from the ICC at the bottom and top of the ICC curve. Near the beginning of the curve the blue ICC tend to follow the red ICC curve but then moves on a tangent of its own. The object characteristics do not follow a smooth curved line but rather a rugged path. This indicates that there were some inconsistencies in the item functioning compared to the other language groups. Once again further testing across a much larger group of learners is warranted.

Violin:

For the Sepedi learners it appears that violin did not function optimally. The blue ICC for **violin** follows the red ICC very closely. Although, comparing the ICC of pan to the ICC of violin there are remarkable differences in the shape of the curves. **Violin** seems to appear normal compared to **pan**. There are one or two deviations from the curve that indicate that the characteristics of **violin** are questionable. But only once a larger number of learners are assessed can it really be determined whether **violin** is really a problematic object.

Saxophone:

For the English learners' the object **saxophone** appeared to have questionable characteristics as there are slight deviations from the curve. In comparison to the other items looking at how the blue ICC of **saxophone** follows the red ICC it seems to be functioning relatively correct. But it does seem to reach a flat line nearly halfway up before continuing upwards rather steeply.

5.9 CONCLUSION

This chapter focused on investigating how the objects used in three different pictures functioned. Furthermore, the order of the objects for all three pictures for each language group was compared to the original order of the objects. The results of each language group for each picture were analyzed. These results were used to indicate how the objects were ordered according to difficulty.

Picture 1 had the second highest number of objects (seven in total) that were to be identified by the different language groups. Additionally, Picture 1 was considered to have the objects that were the easiest to be identified by the learners. Only 29% of the objects for the three groups followed the original difficulty order for the objects as designed by CEM.

The largest number of objects that had to be identified by the learners - totaling 10 objects - was present in Picture 2. Furthermore, Picture 2 had the least number of objects falling on the same difficulty level as the original order. Only 20% of the objects identified by the different groups followed the original difficulty order. But this could be due to the fact that Picture 2 had the most objects that needed to be identified by the learners. The objects presented in Picture 2 were all more difficult than the objects presented in Picture 1, according to the SAMP and PIPS assessments.

In Picture 3 only five objects were presented that were to be identified by the three groups. These objects were also meant to be the most difficult to identify for the learners. But the contrary was found when the results for Picture 2 were analyzed. The number of objects which followed the original difficulty order was the most for this picture. A total of 33% of the objects followed the original order. Although it has to be brought under attention that Picture 3 had the least amount of objects that were to be identified by the learners.

With the help of Rasch analyses items that were not functioning correctly were identified in order to determine what aspects of the Rasch model are not being adhered too. This could be items that were either too easy or too difficult or the item

did not follow the hierarchy order. The Item Developmental Pathway's (IDP) were produced for each picture and were accompanied by a table with numerical values. The IDP of each picture provided diagrammatical representations of the order of the objects relevant to the ease of these objects being identified by the learners. The tables provided numerical information about the ease or difficulty the different groups had at identifying the various objects. This was done to determine whether the level of validity was influenced in any noticeable manner.

Furthermore, items that function differently across the three language groups were identified by means of Differential Item Functioning (see Chapter 3, Section 3.2.7). An example would be that all the Afrikaans learners get one specific item incorrect while the English and Sepedi learners answer correct. This item is acting biased and needs to be reconstructed. Now that the data has been thoroughly explored, suggestions can be made about ways to increase the level of construct validity.

With this chapter the conclusion can be made that all three groups had varying amounts of objects that followed the same difficulty order as the original order in the PIPS and SAMP assessment.

Ultimately all of the objects for the three pictures have been properly explored and final conclusions and recommendation can be made, which will be discussed in the final chapter.

CHAPTER 6

REFLECTIONS ON STUDY

6 REFLECTIONS ON STUDY

6.1 INTRODUCTION

In the previous chapter the data were analysed in depth. Crucial to any research is to reflect on what has happened and the findings. In this chapter reflections are provided on various factors that played a role in this study.

Firstly construct validity was reflected upon in Section 6.2. Then the conceptual framework together with the literature review of the study was reflected upon in Section 6.3. In Section 6.4 reflections on the methodology for the study are discussed and the chapter is concluded in Section 6.5.

6.2 REFLECTION ON CONSTRUCT VALIDITY

The purpose of the research documented in this study was to explore how the level of construct validity for the Picture Vocabulary Test used to assess Grade 1 learners, could be increased. The various objects presented in the different pictures were investigated to determine their fit and function in the Picture as objects presented in pictures can influence the validity level in several ways. Learners from different backgrounds may experience difficulties in identifying objects presented in the pictures (see Chapter 2). As indicated in preceding chapters, **validity** is the driving force behind this study, particularly construct validity.

The traditional role of tests, documented in an article written by William (2006), is to judge and classify learners. The greatest concern though is whether these judgements are valid and fair with the veracity of assessment data being questioned.

A need exists to determine actual proficiency compared to demonstrated proficiency, in other words, what the learner knows and can do versus how the learner performs on a test (Wise & DeMars, 2005). These are all issues that can negatively influence the validity level of a test. To recap, a test can never be valid, only the judgements, inferences or interpretation of the results of the test (Downing, 2003).

Moss, Girard and Haniford (2006) agree that when correct interpretations and decisions are made about the results of an assessment, then a high level of validity has been achieved. Correspondingly, Downing and Haladyna (2004, p. 327) write that validity refers to: "…the degree of meaningfulness for any interpretation of a test score". In other words, how true and accurate the inferences are about the test results and that these inferences are backed up by evidence (Briggs, Alonzo, Schwab & Wilson, 2006; Downing, 2003).

Construct validity is an unobservable construct used in a test to assess learners. The higher the construct validity level of a test, the more the test accurately tests the desired trait or knowledge (Pesudovs, Burr, Harley & Elliot, 2007). The focus of construct validity is on the relationships of the sub-processes of a test, that is, the constructs and whether any are being compromised, thereby embracing all forms of validity. To explain further, construct validity includes the relevance of the content, the representativeness of the content and related criteria (Gorin, 2007). Construct validity determines whether the content of the items are useful in providing information. Do the items represent what is being tested? Are the items up to standard in order for the correct inferences to be made? In short, construct validity measures knowledge and skills, followed by the necessary actions or forms of behaviour that are expected to show the specific knowledge or skill (Embretson, 2007). With a high level of construct validity, any claims made on the data can be supported. Explanations can be convincingly provided, including backing that the data fits the model (Mislevy, 2007). By making use of Rasch analyses, all the necessary steps can be taken to ensure that a high level of validity is maintained in a test, so that when certain inferences are made about the results of the test they can be said to be valid and true. The reason for this is that the Rasch analyses can show that a high level of construct validity has been achieved because the data fits the Rasch model and therefore adheres to the assumption of unidimensionality.

By making use of Rasch analyses the construct validity of the Picture Vocabulary Test was thoroughly explored. The Rasch model is a very effective and accurate statistical procedure to determine whether the construct validity level of an assessment is being jeopardised, as it analyzes dichotomous data (Pallant & Tennant, 2007) as used in this study. Any items that were biased or misfitting were detected by the Rasch model. By means of a Rasch analysis, evidence can be provided as to whether a particular item over- or under-discriminates, and if any anomalies exist in the ordering of the items.

Fairness and high levels of validity have become top priorities on assessment agendas (Abedi, 2002). A question is asked as to whether tests are testing what they are supposed to test, or is the test fair across groups and genders? (Alias, 2005). Differing linguistic backgrounds and mismatches between the learners' cultures and the assessment all play a role in lowering validity levels of an assessment. For an assessment to have a high level of validity, items must be used that are equally distributed attributes across various groups - that is they must perform the same across groups (Chen, Gorin, Thomson & Tatsuoka, 2008).

Many threats to validity exist. According to Downing and Haladyna (2004) there are as many threats to validity as there are sources of validity evidence. Threats to validity are any factors that cause interference with assessment data and the meaningful interpretation thereof. In the same way, this research study set out to determine the threats to validity that existed in the Picture Vocabulary Test. Possible threats that can be identified are biased items, too easy or difficult items and flawed item formats. These threats can be detected when exploring the data of an assessment (Downing & Haladyna, 2004).

By means of a Rasch analysis, threats to the level of construct validity for the Picture Vocabulary Test were detected. To determine any threats concerning bias of items a Differential Item Analysis (DIF) was conducted, to test for any differences in item performance across groups with the same abilities, thereby picking up any items that are biased. DIF analysis is a statistical technique used to detect any misfitting items (Wyse & Mapuranga, 2009). Item bias occurs when there is a statistically significant difference in an item's performance across groups. Once biased items have been

detected, changes can be made to ensure future fairness and equity of the relevant assessment. Furthermore, DIF evaluates whether the learners from the different groups have equal opportunities to succeed (Hauger & Sireci, 2008).

Once the data for all three groups was thoroughly explored, misfitting items were detected and threats to validity identified. The results provided evidence that certain items were biased towards certain language groups. The results also pointed out that certain items were too easy for the learners, for example *knife, carrots, fork* and *cash*. Further evidence was provided showing that the items followed did not follow the same pattern of difficulty for the three groups or the order of the original instrument used in the UK. A number of reason can be given as to why these objects did not function as expected. The ability of the learner to accurately identify objects used in a Picture Vocabulary Test is influenced by various factors. Relevant to this study was the visual literacy, language and culture of the learner.

These factors and the role they played in the learners' ability to identify objects in the Picture Vocabulary Test, are briefly discussed.

6.3 REFLECTION ON LITERATURE AND CONCEPTUAL FRAMEWORK

In order to understand the concerns raised about the construct validity level of the Picture Vocabulary Test, factors that could possibly influence validity levels were studied. In the conceptual framework of this study, they were identified and their roles in influencing validity levels explained.

Van de Vijver and Tanzer (2004) state that globalisation has pressurised assessment to have a high level of validity across cultures. However, the greatest concern with an assessment administered across different cultures is whether the results can be interpreted in a similar manner (van de Vijver & Tanzer, 2004). Ross and Ehlers (2001) extend this point further by stating that graphical material (such as objects) used in cross-cultural testing are often seen as part of a universal language, recognised by all with no cultural innuendos. However, cross-cultural research has

been found this as not being the case. Dowse and Ehlers (2001) concur that the target population must be taken into consideration when using pictures for educational reason and thus applying this statement to this study would be to consider the target population when using objects that are to be identified across three different language groups.

The primary literacy for the present era is mainly visual when considering *Wii, X-Box, Playstation, iPod's, DSTV* and similar products. **Visual literacy** is graphically described by Burmark (2010, p. 15) as: "…3-D eyeglasses for the mind. They are the lenses through which we see the meaning – the words and ideas – behind the images". Through pictures and the objects used in them people can interpret what is meant by a picture without having to read any words. Pictures give words and ideas a reality.

The term 'literacy' has also taken on a broader scope that includes visual literacy. It is seen as the complex ability to understand and use symbols of a culture, including media and electronic text, as well as alphabets and numbers, to promote personal and community development, as defined by The Centre for Literacy of Quebec (in Kickbusch, 2001). In order for a person to understand what symbols, objects, pictures, pictograms are, they must be taught (Dowse & Ehlers, 2001). These research studies indicated that if people are taught how to read images their performance in visual perception tests are greatly enhanced. Likewise, with advanced visual literacy, learners, for example, will hear the word *carrot* and be able to identify it when presented in a picture, as in the case of this research study.

The successful identification of objects used in Picture Vocabulary Tests is largely dependent on how many times the learner has been exposed to the particular object, if at all. For the object to be accurately identified the learner has to have been taught to accurately "…interpret three-dimensional reality on a two-dimensional surface" (Dowse & Ehlers, 2001, p. 88). Dowse and Ehlers (2005) further indicate in their health literacy research that using pictograms was highly beneficial for comprehension. When explaining to patients how their medication must be taken, the use of pictograms positively influenced patients' understanding and adherence of the prescription. The same can be said for learners participating in a Picture Vocabulary

Test. When they can successfully accomplish the task of identifying objects in the test they have developed their visually literacy to a satisfactory degree. If the learner's visual literacy is developed and on par with the *objects used in a Picture Vocabulary Test,* the validity level will increase. This is achieved by having access to multiple resources. If the learners have not been visually exposed to the various objects, they will have a lower level of visual literacy. Furthermore, they will have difficulty in accurately identifying the objects presented in the test.

This lower visual literacy creates a *barrier to the inferences about the validity level of the Picture Vocabulary Test*. By providing the opportunity for learners to have access to various resources, their visual literacy can be increased. This is not always possible in South Africa since many schools are in a financial predicament and cannot afford resources, but the successful identification of objects in a test is largely dependent on influences other than just culture and language.

Culture is seen as beliefs, ideas and traditions that are taught from generation to generation, whether efficient or not (Guiso, Sapienza & Zingales, 2006). In the same way, Gutiérrez and Rogoff (2003) see culture as a community of people that have common understandings and traditions that have been extended over generations. These beliefs, ideas, traditions and common understandings form referencing frameworks within a person's mind, which in turn help create meaning from images that are seen. The meaning that is created from images within the viewer's mind is influenced by cultural factors and personal characteristics (Houts, Doak, Doak & Loscalzo, 2006).

These cultural factors and personal characteristics can hinder the interpretation of a 'three-dimensional reality on a two-dimensional surface' as Dowse and Ehlers (2001) noted. Houts et al. (2006) point out that not only are language and other modes of meaning being influenced by culture but also the comprehension of a picture, which is greatly influenced by its cultural relevance: "Pictures are heavily laden with culture-bound conventions that must be learned if they are to be understood" (Houts et al., 2006, p. 180). For this reason, pictures cannot be seen as being universally understood or a universal language.

When contemplating the use of pictures in any form, consideration must be given to the target group. A pilot study was undertaken prior to this research study to determine the most identified objects among the three language groups used in this study. The learners were drawn from rural and urban schools, although the majority were from urban schools. The objects that were to be identified by the learners were taken from a book found in well-known educational bookshops. These objects had been especially drawn and created for Foundation Phase teachers and learners in South Africa to use, duplicate and teach from. The results reflected interesting discoveries.

Leading to the first discovery, learners were asked to identify *sliced bread* as well as a *whole loaf of bread*. These two objects were not situated next to each other on the instrument, but on different pages between other objects. Most of the learners identified the sliced bread with ease but had difficulty identifying the whole loaf. One of the reasons for the misidentification was considered to be that bread is commonly sold in its sliced form but less commonly sold whole. This clearly illustrates that when considering objects that are to be used in assessment for instruction or educational purposes, the target group has to be considered only after appropriate contemplation.

The second discovery was similar to the first. A *mielie* (corn on the cob) was shown with its leaves and stalk and mieliecorns clearly visible, but the learners had difficulty in identifying it. After discussing this with colleagues, several reasons for the learners' experiencing difficulty were formulated. One interesting reason was that the learners only see *mieliepits* (corn) that are found in packaging or tins and very rarely see an entire *mielie* with its leaves and stalk. Yet again, this depends on where a child has grown up. Learners from farming areas may easily identify a *mielie* while urban learners may have greater difficulty, depending on the target group. Further research needs to be carried out to explore whether the reasons given for the responses to the *whole bread* and *mielie* are indeed valid.

Dowse and Ehlers (2001) reported that, in their research using pictograms to help patients better understand how to take their medication, they found that they had to use pictograms specifically designed for the target cultural group in order for better

comprehension. In their study, the target cultural group was IsiXhosa. Not only are IsiXhosa's a cultural group but they also are a specific language group in the country. This strongly indicates that language and culture both form part of, and influence each other. An article by Grant and Wong (2003, p. 390) promotes this point further: "… language and other modes of meaning are dynamic representational resources, constantly being remade by their users as they work to achieve their various cultural purposes".

Language is seen as a powerful means of maintaining and continuing culture and creating social identities (Janks, 2000). From a very early age, children master language in remarkable ways. By the time they are three years old they can have conversations and even make simple jokes. As children grow older they learn more about language and by the time they go to school they add approximately 3,000 additional words to their lexicon a year. However, much of this is dependent on parental input (Ely, 2005).

Language as a tool also aids us in making sense of our experiences, expressing our experiences and transforming our thinking and understanding. Gutierrez, Asato, Santos and Gotando (2002, p. 346) affirm this point by adding that language "…indexes or signals our particular identities and memberships as groups". In addition, the authors impart that language creates an intimate connection with who we are as well as our communities and its practices. In Furstenberg, Levet, English and Maillet (2001, p. 95) the authors describe that in order to understand the aspects of culture one has to "…constantly operate at the intersection of language and culture".

The relationship between culture and language has been debated since Vygotsky started contemplating his numerous concepts and theories in the 1920's. Vygotsky systematised the concept that human activities occur within specific cultural contexts which are mediated by language and other symbol systems. These are known as socio-cultural approaches to learning and development (John-Steiner & Mahn, 1996).

From the above information, it can be concluded that language and culture are interconnected. Neither can deny not being influenced by the other in some way or

another. In this study, culture and language, are seen as inseparable, with both playing a major role in the degree to which objects are correctly identified by a learner. Objects need to be used that have been specially chosen for identification by the target group of learners, with special consideration given to the culture and language.

During and after the study another factor that could act as a barrier to the validity level of the Picture Vocabulary Test was identified as a possibility. This factor was not explored in this study but could be included in any future research. Not only does culture and language impact how *objects used in a Picture Vocabulary Test influence the level of validity* but also the **Socio-Economic Status** (SES) of the learner does. The early years of a child's life can make an immense difference to his or her contribution to society as an adult. In the first five years of a child's life, it is crucial for him or her to receive support: "…in growth in cognition, language, motor skills, adaptive skills and social-emotional functioning" (Grunewald & Rolnick, 2006). Failing to provide support in these areas could lead to, inter alia, school drop-out, crime and poor academic performance (Currie, 2001).

Poverty is a root cause of slums and settlement colonies and impacts on all aspects of a child's development (Nair, 2004). In South Africa, due to economic reasons and the apartheid past, a large majority of the population are forced to live in informal settlements. An informal settlement is usually an unplanned and unauthorised settlement in urban areas and can be visually identified by their temporary structures that are known as 'shacks'. A group of shacks together are known as a 'shantytown' (Huchzermeyer, 2004). A 'shack' is a type of hut that is made from corrugated iron sheets, pressed wood or any other material that will suffice. These shacks are very hot in summer and very cold in winter. An informal settlement may be created either for work opportunities, often situated near industrial areas where the occupants are within walking distance of their workplace, or because of poverty (Adams, Sibanda & Turner, 1999; Smit, 1998). In most circumstances, there is no water or electricity, which means that fires are made to cook food while water is taken either from a stream nearby or from a source that has running tap water. Typical characteristics of these types of settlements, according to Nair (2004, p. 228), are: "… substandard housing, overcrowding, poor water, sanitation and sewage disposal facilities …"

Living in such informal settlements and affected by poverty has a major effect on the academic achievement of children. The development of a child is greatly influenced by their parents' financial situation as explained. Nair (2004. p. 229) states that genes set the limits of achievement and the environment determines whether or not it can be achieved. In poor urban areas, parents tend to be uneducated and unskilled. In most circumstances and traditions, it is the mother's responsibility to look after the children, but in poor urban areas there is a very high chance that the mother is illiterate. If the mother is working, the upbringing of the smaller children is left to an older girl or sibling, who is not attending school because s/he has to look after the children and is also probably illiterate.

Resources, which are available in the home environment, influence how a child performs academically Learners who come from a higher cultural capital group achieve better academically as higher Socio Economic Standards (SES) facilitate the development of higher cultural capital through a broader exposure to objects and resources. Family background also influences the learner's academic performance and achievement. If a leaner comes from a background where his or her parents were unschooled, illiterate, or of a low SES, they stand a chance of performing more weakly academically than their peers. However, learners who attend pre-school are at an advantage academically (Merrell & Tymms 2005a, Roscigno & Ainsworth-Darnell 1999, Teachman 1987).

Comparative studies have been conducted on children from a high socio-economic status (SES) and those from a low SES (Nair, 2004). These studies showed that children from low SES had a lower developmental status than their high SES counterparts. These homes lacked toys that could teach children animal names and how to count. The poor home environment, combined with inadequate provision of toys and other play materials, leads to poor language and fine-motor skills. These studies very clearly indicated that children from poor urban settings lagged behind in their skills development, which in turn had a direct influence on future academic performance at school (Campbell, Ramey, Pungello, Sparling & Miller-Johnson, 2002; Ramey, Campbell, Burchinal, Skinner, Gardner & Ramey, 2000).

The research indicated that in order for children to develop academically the buildings and playgrounds must be safe. The activity rooms should be separate from play rooms. Toys which can teach colour, shape, and size as well as puzzles that develop creativity need to be supplied to the children. Toys and games that promote refined movements ought to be provided. Reading books, musical instruments, a display of children's artwork and toys that teach them to name animals, birds and various other objects are a definite prerequisite. The most important factor towards the successful development of children is having qualified and trained pre-school teachers (Nair, 2004). Optimum nutrition is also said to have a positive effect on academic achievement, and malnutrition opposite negative effect (Glewwe, Jacoby & King, 2001). It is evident from what has been discovered above that children from a high SES have more resources available to them and have more advantages to interact with different educational toys than children from a low SES. Exposure to various resources gives the high SES children a greater academic lead.

There have been studies made of persons who come from a technologically poor environment and have had very little, if any, exposure to objects outside their immediate environment. A number of these studies are cited in the work of Cassidy and Knowlton (1983). One of these interesting studies was on the categorising skills of Kpelle people. The Kpelle, also known as Guerze, as described in the *Encyclopaedia Britannica* (2006), are found in most of Liberia extending into Guinea. They are primarily farmers with a variation of crops such as rice, vegetables, fruit, peanuts, sugar cane and kola nuts. The household consists of a man and his several wives. In the study the Kpelle were asked to sort 20 different objects into what was thought to be 'meaningful categories' or groups. The Kpelle would group a *knife* with an *orange*, a *hoe* with a *potato* and so forth. The Kpelle felt that these were wise ways of categorising the different objects. When asked how people not as wise would do it they grouped the fruit together, the tools together and so on. This grouping was done in the manner originally expected from the Kpelle.

Although the Kpelle were not technologically advanced they had their own idea of what was considered to be wise and educated decisions and what was not. The possibility could exist that if they were more technologically advanced they would follow a different paradigm of grouping. Their preference of thinking is not wrong,

only different, and it may be that someone else, exposed to more technology might group the objects differently. SES influences how people think, act and behave.

These differing levels of SES are a cause for concern when setting up an assessment, especially a Picture Vocabulary Test, since learners with a high SES have certainly been exposed to more objects, being more technologically advanced than those with a low SES. Learners from a low SES also have classrooms with limited resources available (Campbell, Ramey, Pungello, Sparling & Miller-Johnson, 2002).

In this study, the learners from the Sepedi group are mostly from a low SES group. They do have schools built out of bricks but they have limited resources and are situated in areas that have very few resources available.

SES plays a major role in the amount of resources that are available to learners. The greater the number of resources available the greater the chances that more learning will take place. The converse is also true. A learner is more likely to succeed in identifying objects that are presented in a Picture Vocabulary Test when more resources are available to explore and learn from. The number of resources available in turn influences the learner's ability to identify objects presented in the Picture Vocabulary Test. This has a domino effect by either increasing or decreasing the level of validity of the test. The greater the learner's exposure to multiple resources, the greater the chance will be of correctly identifying objects used in the test. Again, the opposite is also true. The lower the SES of the learner the less likely the chances are of being exposed to various resources, which negatively influences the learner's ability to identify different objects.

A rather cynical paraphrased aphorism that can be argued is given by Erickson and Gutierrez (2002, p. 23): "Those who do not know their intellectual history are condemned to repeat it". This position is expanded on by Heckman (2006), who poignantly makes it clear that if young children are not stimulated by their environments they are placed at an early disadvantage. He takes this statement further by stating that children that fall behind may never catch up. Arnold and

Doctorhoff (2003) substantiate the abovementioned authors in his article by commenting that poverty has a negative effect on academic achievement.

Poverty has serious repercussions on a child's cognitive functioning due to poor nutrition, housing and water supply. A synergistic relationship exists between poverty and lack of education (Low, Low, Baumler & Huynh, 2005), but this does not mean that the child will not succeed at school. On the contrary, poverty results in a lack of resources, not academic failure. The lack of resources may result in a child not being able to identify as many objects as one from a higher SES, but that does not mean that the child is intellectually challenged.

The extent to which young children are exposed to objects, situations, visual materials and the surrounding world influences their lives, their perception of their environment and ultimately their visual literacy levels. Depending on their culture and socio-economic status, learners either have the means to broaden their exposure to the world and surroundings or have a limited degree of exposure. In many poorer areas, there is a shortage of resources and educational materials, which limits the learner's exposure to pictures (Arbuckle, 2004). With little or no exposure to various forms of resources and materials, learners are less likely to develop their visual literacy skills. This in turn could lead to difficulty in identifying certain objects because their visual literacy levels are not as developed, thereby increasing the chances of misinterpreting or not identifying the few objects they do see in pictures.

Resources, which are available in the home environment, influence how a child performs academically. Learners who come from a higher cultural capital group achieve better academically as higher Socio Economic Standards (SES) facilitate the development of higher cultural capital through a broader exposure to objects and resources. Family background also influences the learner's academic performance and achievement. If a leaner comes from a background where his or her parents were unschooled, illiterate, or of a low SES, they stand a chance of performing more weakly academically than their peers. However, learners who attend pre-school are at an advantage academically (Merrell & Tymms 2005a, Roscigno & Ainsworth-Darnell 1999, Teachman 1987).

An example related to this study is found in Picture 1, where the learners are asked to identify *cherries*. These fruits are comparatively expensive in South Africa and would most likely be found in more affluent homes, but this does not mean that a child has never been exposed to them in some form or another. However, if resources and money are limited the possibility exists that the learner will not be as successful in identifying *cherries* as other learners.

If learners are presented with objects that the resources from their SES supply they will be able to identify the objects with greater success. If resources are limited then the learner will also be limited to a certain extent.

The conceptual study used in this research study is based on the idea that in order for a Picture Vocabulary Test to have a high level of validity there must be commonalities between the learners' *visual literacy*, *language* and *culture* and the objects used in the test. These three factors influence the learners' performance in a Picture Vocabulary Test as well as its validity level. If any of these three factors do not relate to the items used in the test then the validity level is in serious jeopardy.

The conceptual framework used in this study is shown (see figure 6.1)

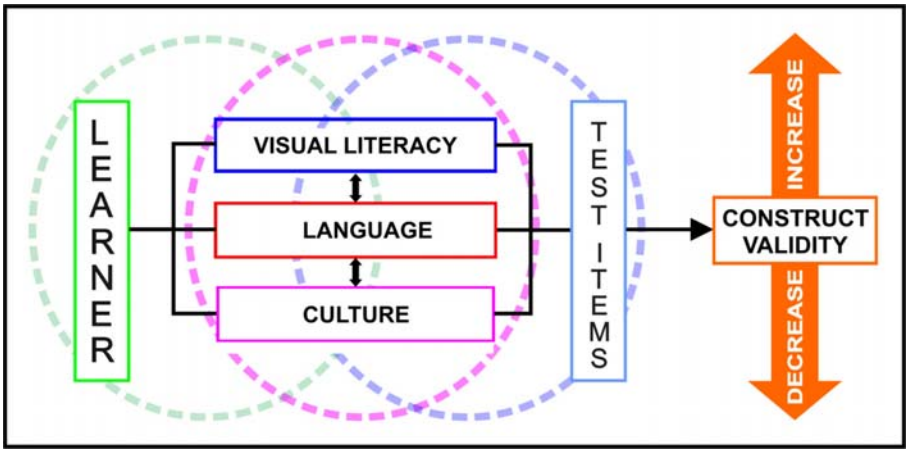


Figure 6.1: Original Conceptual Framework

The conceptual framework could be reconstructed to indicate that language, culture, SES and Visual Literacy of the learner must overlap with the objects used in a

Picture Vocabulary Test. If this happens the objects will perform as expected and the test will have a high validity level. The learners' visual literacy, language, culture and SES must relate to the objects that are used in a test. The adapted conceptual framework is shown in figure 6.2:

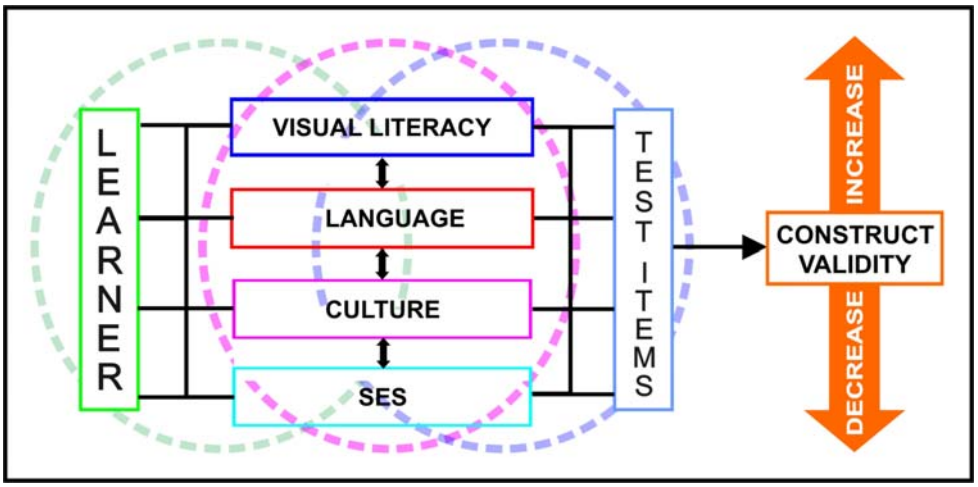


Figure 6.2: Adapted Conceptual Framework

By considering a learner's SES, objects can be identified and incorporated into a Picture Vocabulary Test that he or she can relate to. The common ground that is found between the learners SES, culture, language and visual literacy will help to ensure that valid inferences are made about the learners' results.

6.4 CONCLUSION

As can be seen from the above reflections further research can be done that includes the additional factor of SES. Together with research, continual reflection and improvement has to take place that are essential for successful advancement in any field, and in the study's case in the field of education. By reflection on the various aspects the field is left open for further research that can improve upon this study.

CHAPTER 7

CONCLUSION

"…positive educational change is accomplished locally and it is more like walking through a swamp, testing the ground with each step, than it is like driving on a superhighway or even like building one. To get smarter about working our way in a swamp we need all kinds of research and deliberation, scientific and non-scientific".

- Erickson and Gutierrez (2002, p. 23).

7 CONCLUSION

7.1 INTRODUCTION

The purpose of this final chapter is to revisit the research questions, summarise the processes used to find the answers to them and discuss the results. Validity and the factors that influence the level of validity in relation to objects used in the SAMP Picture Vocabulary Test were explored, before drawing conclusions and making recommendations. In the preceding chapter, the data was analysed and documented. In this chapter, a brief summary of the research design is given in Section 7.2. A discussion of the findings is provided in Section 7.3. The implications for practice are discussed in Section 7.4, and recommendations for future research are offered in Section 7.5. The limitations of the study are discussed in Section 7.6. Lastly, the concluding remarks in Section 7.6 capture the substance and scope of this study. This is done so that an effort can be made to provide an in-depth understanding of the influence objects have on the level of validity in a Picture Vocabulary Test.

7.2 SUMMARY OF THE RESEARCH DESIGN

The purpose of this study was to explore how objects used in a Picture Vocabulary Test influenced the level of construct validity of the test. The theoretical position within which the design of this research study took place was Positivist, making use of a quantitative methodological approach to determine the performance of the objects used in the Picture Vocabulary Test. Positivism is seen from the perspective that science does not need to have a prior sense of the whole to which different parts belong in order to study the different parts (Fisher, 1991).

The sample used in this study was the target population of Grade 1 learners who were in schools whose medium of instruction was Afrikaans, English and Sepedi within Pretoria, Gauteng, South Africa. There were 355 Afrikaans learners, 562 English learners and 444 Sepedi learners. These languages were selected because they are the most dominant in the Pretoria area and the most accessible population for the South African Monitoring in Primary Schools (SAMP) project.

The Picture Vocabulary Test forms part of a larger assessment instrument of the SAMP project. The learners were asked to identify 22 objects that were found in three different pictures. The objects were arranged from easy to difficult, an order originally arranged for an instrument used in the UK. Although the pictures were redrawn and adapted to suit a South African context, the item difficulty order remained the same as for the original instrument. This raised some concern and warranted further exploration. Furthermore the items were explored in terms of construct validity level of the Test, so this study set out to explore whether the concerns were valid and, if so, what suggestions could be made to increase the level of construct validity.

Statistical procedures were followed to analyze the data, use of Rasch analyses. By making using of Rasch analyses, the functions of the items of the test were quantitatively investigated. Rasch analyses are quantitative in nature because of the

statistical procedures used to explore the items in an assessment. Rasch has been used extensively over the years especially as a research tool for researchers since inferences can be verified the adequacy of the instrument and its level of construct validity can also be verified (Callingham & Bond, 2006; Rasch, n.d.; Tennant & Conaghan, 2007).

7.3 SUMMARY OF RESULTS PER RESEARCH QUESTION

The data were explored overall and across the three language groups participating in the Picture Vocabulary Test, so that the research questions could be successfully answered. The factors identified and discussed in the literature review influenced how objects in the Picture Vocabulary Test performed across the language groups. Although this study was limited to only three language groups, evidence illustrated that there were similarities and differences across the languages.

The main research question that guided this study was:

How do objects used in a Picture Vocabulary Test influence the level of validity? This research question was operationalised by means of four specific questions namely:

1. What barriers to validity used in a Picture Vocabulary Test can be identified from literature?
2. To what extent is a unidimensional trait measured by the Picture Vocabulary Test?
3. To what extent do the items in the Picture Vocabulary Test perform the same for the different language groups?
4. How can the identified barriers that decrease the level of validity be minimised?

For the first specific research question the barriers to validity used in a Picture Vocabulary Test can be identified from literature.

As evidenced in the review of literature (Chapter 2) three factors were identified that influence a learner's ability to identify objects, and can be seen as possible barriers to validity. The first factor was language. Within the educational sphere, language has become a multifaceted phenomenon that challenges any educator when knowledge has to be put across to diverse learners. Research has shifted from studying children from one specific language group to those from diverse linguistic societies. As well as studying children learning more than one language at a time, studies have even reported that judgements are passed on children with certain dialects (Garcia, 1993). Incorrect judgements can also be made on one language group's performance compared to another language group when they participate in a test that is not impartial across language groups.

The second factor, *culture*, forms an integral part of each human. Culture is not genetically predisposed from one generation to the next, but rather consists of acquired knowledge, learned patterns of behaviour, attitudes, values, expectations, rituals and rules, giving a person a sense of identity and what his or her history is. Culture and language influence every aspect of society at every level, such as home, school, education and, work, and are an integral part of each human, being a heritage carried with them, be it consciously or subconsciously. As a result, culture and language play a fundamental role in the educational development of the learner (Webb & Read, 2000). It is often taken for granted that pictures are seen as being independent of language and culture, despite the message they are communicating (Hoffman, 2000, p. 35), cultural backgrounds and languages spoken, all of which influence the way pictures and objects are seen and identified by people.

The third factor is *visual literacy* (VL), an exceptionally broad concept. It is used across numerous disciplines, each with its own relevant definition, attributes and expectations of the term. The definition of VL most suited for this study is the ability to accurately identify objects and pictures seen in the past when they reoccur in the present in a similar or altered manner. Sims et al. (2002) argue that because the number of captured visual images is increasing in an age of technology, successful educational outcomes should be at the forefront, with VL cultivated and taught. VL plays a fundamental role in learners' ability to identify objects shown in a Picture Vocabulary Test. The learners must have had past visual experiences associated

with the various objects used in it and the ability to identify the objects. They must be able to differentiate, make sense of and identify the objects displayed, although this can only happen if they have had past exposure to them. The problem then arises that such objects must also be applied in such a manner that they are identifiable by all cultures, where possible (Arbuckle, 2004; Burton, 2004; Sims et al., 2002). All these factors play a role in how objects are perceived, remembered and understood.

The second specific question was: **To what extent is a unidimensional trait measured by the Picture Vocabulary Test?**

The results from this research indicate that a unidimensional trait was measured. This is evidence from the fact that only a few items misfitted for the different language groups and this was mainly in Picture 1 which typically consists of easier objects. Objects included for Picture 2 and Picture 3 all adhered to the predetermined criteria and therefore did not misfit. The conclusion that a unidimensional trait is measured is further supported by the developmental pathways. Furthermore, when all of the items were analysed together none of the items misfitted providing further evidence of unidimensionality and a relatively high level of construct validity. However, what is of concern is that the item person targeting could be improved and this would mean possibly including more items of varying difficulty levels which would be appropriate for the various ability levels along the continuum of vocabulary ability.

Hawthorne and Tomlinson (1997, p. 301) wrote: “Pictures are most effective when their contents are familiar, realistic and depict a single activity”. Carney and Levin (2002) point out that when pictures are used in an appropriate manner, learning can be enhanced. Relevant to this study when objects used in a Picture Vocabulary Test are used in an appropriate manner for the appropriate language group the validity level can be increased. Certain objects performed the same in relation to difficulty across language groups namely ***knife, cupboard, cash, kite, castle, toadstool, padlock, violin*** and ***microscope*****.** This shows that certain objects can be used across groups.

How can objects used in a Picture Vocabulary Test increase the level of validity?

The third specific question addressed in the research is: To what extent do the items in the Picture Vocabulary Test perform the same for the different language groups? The results indicated that there was negligible DIF for the majority of the items. However, four items were identified as exhibiting non-uniform DIF in which one group found the item substantially more difficult than the other groups included in the analysis. These items **pan, castle, violin and saxophone** which were then explored further.

As indicated in the literature review, certain factors were identified that influence learners' ability to identify objects presented in a Picture Vocabulary Test. The identified factors were the visual literacy, culture and language of the learner. If the learner has been given the opportunity to develop his or her visual literacy they would be able to successfully identify objects presented in a Picture Vocabulary Test.

It is important to note that if consideration is given to the cultural background of the learners and their language, considerable improvements could be made to the validity level of the test. Objects must be selected that are familiar to the culture and language of the leaner. A distinct overlap must be seen between the objects that are used in the assessment and those found in the culture of the learners. Additionally, the objects that are used in a Picture Vocabulary Test must be found in the language of the learner.

The final specific question is: How can the identified barriers that decrease the level of validity be minimized?

Possible barriers to the validity level of the Picture Vocabulary Test could be described as follows:

- Objects that did not perform the same across the three different groups, indicating bias. The bias items were **pan, violin, castle** and **saxophone.**

- Objects that were to be identified by all the learners were not aligned with the learners' abilities.
- Objects that were supposed to be more difficult in the context of England were experienced as being easier for the South African learners. Examples were **knife, cupboard** and **cash.**
- Objects that were supposed to be easier in the England context were experienced as being more difficult for the South African learners. Examples were **violin, castle** and **kite**.
- Certain objects again were seen as too easy for the learners' abilities. These objects were **knife, fork** and **carrots**. This could be the result of well developed visual literacy in the learners, resulting in the objects not being challenging enough for identification, or because they are more readily available in the learners' culture.
- The arrangement of the objects' difficulty order across the three language groups did not match the difficulty order. Each language produced differing difficulty orders for the objects used in the Picture Vocabulary Test. For example, the objects order for the Afrikaans learners for difficulty levels 13 - 15 were **tortoise, jewellery** and **pigeon**. For the English learners it was **jewellery, tortoise** and **pigeon**, and for the Sepedi learners it was **tortoise, pigeon** and **jewellery.** This occurred because the learners were from different cultures and language groups. The visual literacy of the learners influences the ability to identify objects.
- Many objects were on the same level of difficulty as other objects, indicating that they did not increase in difficulty in equal increments. Objects on the same difficulty level across the language groups for Picture 1 were **carrots, cupboard, fork** and **knife**. Objects on the same difficulty level across the language groups for Picture 2 were **castle** and **pigeon**. This is the result of objects that were not selected for a specific target group. An instrument intended for learners from the UK is used on learners from South Africa.

The literature review identified how language, culture and visual literacy can also act as barriers to the validity level of a Picture Vocabulary Test. It is often taken for granted that the learners are familiar with the picture and the objects represented in it

(Arbuckle, 2004). When this is taken for granted and the pictures are used in a Picture Vocabulary Test the level of validity becomes questionable, because each person looking at a picture has their own style of reading and interpreting the story represented by it (Moore & Dwyer, 1994; Weber & Mitchell, 1996). By ensuring that the picture and its objects incorporate the learners' culture, language and visual literacy the barriers to validity can be decreased. For this research study it is suggested that the objects be rearranged to fit the order of difficulty for each language group. This could entail reordering the items so that the difficulty order of the objects can follow accordingly in all three pictures. If only one test is preferred then objects must be used that perform the same across all three language groups and equally increase in difficulty. A further suggestion would be to include additional objects so that learners' abilities can be correctly measures.

Furthermore, Rasch analyses generated separate estimates of each item's difficulty and the learner's ability. These estimates give the researcher a value relative to every individual's ability and every item's difficulty. In other words, Rasch analysis tells the researcher how the item is functioning relevant to the ability being assessed. It also provides indices to determine if there are items that are spread out or in 'clumps'. The items should move up in difficulty at equal levels and not be grouped on one difficulty level. If this happens in an assessment, the level of construct validity would be in jeopardy since the items do not follow the true Guttman pattern (Bond & Fox, 2001). In Picture 1 the clumped objects were *carrots, cupboard, fork* and *knife.* For Picture 2 the clumped objects were *kite, tortoise, castle, pigeon, toadstool, padlock* and *violin*. For Picture 3 the clumped objects were *saxophone* and *yacht*. All these objects were experienced as being on similar difficulty levels. Objects that have differing levels of difficulty will help to increase the validity level of the test.

For all three pictures the order of the objects differed from group to group (Section 4.5) from the original order. As can be expected, each language group identifies objects in its own manner, depending on their level of visual literacy, their cultural background, language and socio-economic status as mentioned above.

In summary, as discussed in the literature chapter, culture and language influence the type of environment, and inadvertently the objects, to which learners are

exposed. In addition, the visual literacy of a learner is also influenced by culture, language and available resources. Certain objects are more familiar to certain cultures than to others. When these objects appear in a Picture Vocabulary Test they are more readily identified by learners in which these objects appear in abundance. When this happens these objects are seen as being biased. The objects used in the Picture Vocabulary Test were generalised across all three language groups, which give cause for concern. It cannot be taken for granted that objects will perform the same across different language groups, even though they are from the same country. If objects are not specifically chosen for the intended target group the validity level of the test will be jeopardised.

The overall findings of the Picture Vocabulary Test which the research study explored revealed that although the objects of the original Picture Vocabulary Test from the UK were familiar to learners from South Africa their performance was not the same in relation to difficulty. The objects had different difficulty levels for the learners from South Africa compared to those from the UK. For example, the object *cash* was experienced as an easy item for learners from South Africa (cash is found in Picture 3, the picture with the most difficult objects to be identified by UK learners). Another example was the object *violin* that was situated on Picture 2 for the UK learners but was experienced as being the fourth most difficult object to be identified by learners from South Africa. However, even though the difficulty level differed the items included in the three pictures and the scale for vocabulary did fit and as illustrated by the developmental pathways do form a sound construct. This study made use of a Rasch model that follows a Guttman scale, which the original UK instrument was designed to follow. This is when the items in a test allow a learner to succeed up to a certain difficulty and then the learner fails items above that difficulty level (Linacre, 2005). Using a Rasch model that follows a Guttman scale will result in some learners being seen as having more ability than others, and there is a greater probability that those with high ability will get the easier items correct. If this is not the case, then the assessment is faulty or has a low level of construct validity (Sick, 2008). However, this did not seem to be the case from the results of this research.

7.4 IMPLICATIONS

A high level of validity is the ultimate requisite for assessments that should be labelled trustworthy (Bond, 2003). The findings of this study have far-reaching implications on many persons who want to use images or objects that are fair when used in assessments. The schooling environment has become accountable for fair assessments across languages and cultures. Many classrooms are primarily accommodating multicultural and multilingual learners, and the educators are expected to treat each learner equally (Pendlebury, Lake & Smith, 2009). Assessments used in classrooms are used broadly and not targeted for a specific group of learners. By incorporating fair assessments into multicultural and multilingual classrooms consideration is given to the factors that influence validity levels.

Persons interested in incorporating pictures in an assessment, educational researchers working with any imagery, objects, pictures, illustrations and persons working with policy will find the evidence of the link between objects and visual literacy, culture and language very useful.

For educators and teachers this study offers insight into what role culture and language play in how objects are identified. It also gives a strong indication that visual literacy must be developed to a greater extent by introducing learners to objects found in different cultures. Furthermore, together with culture the language of the learners must be developed so that general knowledge about the surrounding world can be increased. In particular, more time can be spent in educating the learners about other cultures that are indigenous to their country, including objects used by that specific culture.

This study will also be useful to persons interested in designing culturally fair assessments. The factors that influence how objects are identified by the learners are not only relevant to objects in a Picture Vocabulary Test. These factors influence other assessments that are used across different language groups. The research questions identified the barriers to validity and how these barriers can be minimised.

The same barriers and advice can be incorporated into other assessments with the necessary adjustments made relevant to that specific assessment.

The findings can also relate to policy makers that crucial consideration must be given to solitary assessments used across multiple cultures and language. Unless these assessments have high level of construct validity the inferences made about learners' performance can be false. These false inferences negatively influence the learner's future academic performance. It will be in the best interest of policymakers and educators to sit together and determine the way forward regarding solitary assessments used in multicultural and multilingual classrooms. The actions that result from the judgements made of the test results are squarely on policy maker and educators shoulders if the validity level is not exceptionally high.

7.5 RECOMMENDATIONS FOR FUTURE RESEARCH

The goal of this study was to research how the level of validity was influenced by objects; objects that were used to assess learners from three different language groups. The results were investigated and many significant findings resulted from the Picture Vocabulary Test's data. Although the findings are significant there are some limitations. The following recommendations are provided:

The influence of SES has to be explored further. Future research into this subject should include the influence of the learners' socio-economic status and the availability of resources on their performance in an assessment. This study identified three factors that influence learners' achievement abilities in a Picture Vocabulary Test. With future research, extra possible factors can be identified and explored.

The SAMP project should be extended to include additional language groups. In order to truly unpack whether the assessment functions the same for everyone in the population additional language groups need to be included in future cycles of the SAMP project.

The sampling for the project should be carefully considered. The item-person targeting is an essential component of test construction. The sampling procedures have to be revisited to include groups of different abilities along the continuum of the trait under exploration.

A picture bank relevant to the context of South Africa should be developed. Another possible avenue of research is to explore objects and pictures that can be stored in a picture bank that are commonly identifiable across language groups. These objects and pictures can be used in future Picture Vocabulary Tests as well as for other assessments that make use of pictures or objects.

Bias in tests should be explored and made transparent. If the necessary funds are available, schools could invest in a statistical programme such as *Winsteps*. The assessments used across cultures and languages can be explored to detect items that are biased or not functioning properly. This would help that assessments used in classrooms are fair across groups.

Advances in psychometric theory should be included in future studies. This will help to delineate the difficulty of objects as well as the abilities of learners in Grade 1. Furthermore, with the further developments in psychometric theory the nuances within tests can be adequately explored.

7.6 LIMITATIONS TO STUDY

One specific limitation to this study was that only a limited number of objects and their performance was explored in the Picture Vocabulary Test while a larger amount of objects fall within the visual literacy abilities of the learners. Possibly a better idea of objects that could perform correct across all three language groups could be identified. Another limitation is that only three language groups out of the 11 official languages of South Africa were investigated. As mentioned in Section 7.1, these three languages were chosen because they were most dominant in Pretoria, from

where the SAMP project was managed. The study focused on the language groups that were in the nearby vicinity that were easily accessible, resulting in fewer expenses.

Another limitation is that the sample had learners with limited abilities, that is they were all in Grade 1, all around the same age and with approximately the same knowledge. In order to have objects with varying difficulties future samples should include a range of ability groups to ensure that the item targeting is appropriate. Grade Nought learners should be tested to identify which objects fall within their range of difficulty, the same with Grade 1 and Grade 2. By doing this, the range of items that are easy to those that are difficult can be identified and used. This will ensure that the learners with greater abilities will be matched with items with equal or greater difficulty.

7.7 CONCLUSION

The findings of the study expanded the work of previous researchers in the area of visual literacy as well as researchers interested in validity levels in assessments. This investigation revealed that certain objects perform differently across language groups but that the unidimensionality of the construct was upheld. Additionally, a Picture Vocabulary Test designed for one specific group cannot be used across different groups, the reason being that a strong likelihood exists that certain objects will function differently than anticipated for the different groups. Literature on validity indicates that extensive investigation of items must take place before being used in an assessment. Literature in the field of visually literacy clearly shows that culture and language influence how pictures, illustrations, pictograms and diagrams are understood.

8 REFERENCES

Abedi, J. (2002). Standardized Achievement Tests and English Language Learners: Psychometric Issues. *Educational Assessment,* 8(3), 231-257.

Adams, M., Sibanda, S. & Turner, S. (1999). Land Tenure Reform and Rural Livelihoods in Southern Africa. *ODI Natural Resource Perspectives,* 38, 1-15.

Abraham, J. (1996). Positivism, Prejudice and Progress in the Sociology of Education: Who's afraid of Values? British Journal of Sociology of Education, 17(1), 81 – 86.

Acton, G.S. (2003). What is good about Rasch Measurement? *Rasch Measurement Transactions,* 16, 902-903.

Alexander, H.A. (2006). A view from Somewhere: Explaining the Paradigms of Educational Research. *Journal of Philosophy of Education,* 40(2), 205 – 221.

Alias, M. (2005). Assessment of Learning Outcomes: validity and reliability of classroom tests. *World Transactions on Engineering and Technology Education,* 4(2), 235-238.

Andrews, F.M. (1984). Construct Validity and Error Components of Survey Measures: A Structural Modeling Approach. *Public Opinion Quarterly,* 48, 409-442.

Andrich, D., de Jong, J.H.A.L., & Sheridan, B.E. (n.d.). *Diagnostic Opportunities with the Rasch Model for Ordered Response Categories.* Murdoch University, Western Australia.

Andrich, D. (2004, January 20-22). *Detection of Differential Item Functioning Using Analysis of Variance.* Paper presented at the Second International Conference of Measurement in Health Education, Psychology and Marketing: Developments with

Rasch Models, Perth, Australia. Retrieved November 12, 2009, from Pretoria University's database.

Andrich, D., & Hagquist, C. (2004). *Detection of Differential Item Functioning using Analysis of Variance.* Paper presented at the Second International Conference on Measurement in Health, Education, Psychology and Marketing: Developments with Rasch Models, Murdoch University, Perth.

Arbuckle, K. (2004). 'The language of pictures: Visual literacy and print materials for Adult Basic Education and training', *Language Matters,* 35 (2), 445-458.

Archer, E. (2006a). *PIPSSA planning document January 2006.* Pretoria: Centre for Evaluation and Assessment.

Archer, E. (2006b). Suggested changes to PIPS paper-based baseline instrument for South Africa. Pretoria: Centre for Evaluation and Assessment.

Archer, E., Scherman, V., Coe, R., & Howie, S. J. (2010). Finding the best fit: The adaptation and translation of the Performance Indicators for Primary Schools for the South African Context. *Perspectives in Education, 28 (*1), 77-88.

Arnold, D.H., & Doctorhoff, G.L. (2003). The Early Education of Socioeconomically Disadvantaged Children. *Annual Reviews Psychology,* 54, 517 – 545.

Avgerinou, M., & Ericson, J. (1997). A review of the concept of Visual Literacy. *British Journal of Educational Technology,* 28(4), 280-291.

Badia, X., Prieto, L., & Linacre, J.M. (2002). Differential Item and Test Functioning (DIF & DTF). *Rasch Measurement Transactions,* 16(3), 889-894.

Bamford, A. (2003). *The Visual Literacy White Paper*. Adobe Systems Incorporated: UK.

Barnard, A. (1988). Kinship, Language and Production: A Conjectural History of Khoisan Social Structure. *Africa: Journal of the International African Institute,* 58(1), 29-50.

Beaton, D.E., & Wright, J.G. (2005). Development of the Quickdash: Comparison of Three Item-Reduction Approaches. *The Journal of Bone and Joint Surgery,* 87(A), 1038-1046.

Bohlig, M., Fisher, W.P., Masters, G.N. and Bond, T. (1998). Content Validity and Misfitting Items. *Rasch Measurement Transactions,* 12(1), 607.

Bond, T. (2003). Validity and Assessment: A Rasch measurement perspective. *Metodologia de las Ciencias del Comportamiento,* 5(2), 179-194.

Bond, T.G. & Fox, C.M. (2001). Applying the Rasch Model. Fundamental measurement in the Human Sciences. London: Lawrence Erlbaum Associates.

Boone, W. & Rogan, J. (2005). Rigour in Quantitative Analysis: The Promise of Rasch Analysis Techniques. *African Journal of Research in SMT Education.* 9(1), 25-38.

Borsboom, D., Mellenbergh, G.J., & van Heerden, J. (2004). The Concept of Validity. *Psychological Review,* 111(4), 1067-1071.

Briggs, D.C. & Wilson, M. (2003). An Introduction to Multidimensional Measurement using Rasch Models. *Journal of Applied Measurement,* 4(1), 87-100.

Briggs, D.C., Alonzo, A.C., Schwab, C., & Wilson, M. (2006). Diagnostic Assessment with Ordered Multiple-Choice Items. *Educational Assessment,* 11(1), 33-63.

Buchanan, D.R. (1998). Beyond Positivism: humanistic perspectives on theory and research in health education. *Health Education Research, Theory and Practice,* 13(3), 439 – 450.

Burmark, L. (2010). Visual Literacy: The Power of Images for Learning. *OnCue,* 30(4), 14 – 19.

Burton, L. (2004, October 22-23). *The Tide is Turning: Trawling the Tributaries Together.* Paper presented at 2004 Australian School Library Association (NSW) Inc 5th State Conference. Australia.

Bush, M.J. & Schumacker, R.E. (1993, April 12-16). *Quick Norms with Rasch Measurement.* Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, Georgia. Retrieved June 2008 from University of Pretoria database.

Callingham, R. & Watson, J.M. (n.d.). Overcoming Research Design Issues Using Rasch Measurement: The *StatSmart* Project. University of New England.

Callingham, R. & Bond, T. (2006). Research in Measurement Education and Rasch Measurement. *Mathematics Education Research Journal.* 18(2), 1-10.

Campbell, F.A., Ramey, C.T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early Childhood Education: Young Adult Outcomes from the Abecedarian Project. *Applied Developmental Science,* 6(1), 42-57.

Carney, R.N. & Levin, J.R. (2002). Pictorial Illustrations *Still* improve Students' Learning from Text. *Educational Psychology Review,* 14(1), 5-26.

Carter, R., & McCarthy, M. (2004). Talking, Creating: Interactional Language, Creativity, and Context. *Applied Linguistics,* 25(1), 62-88.

Cassidy, M.F., & Knowlton, J.Q. (1983). Visual Literacy: A failed Metaphor? *ECTJ,* 31(2), 67-90.

Cavanagh, R.B., Romanoski, J., Giddings, G.J., Harris, M. & Dellar, G.B. (2003). *Application of Rasch Model and Traditional Statistics to Develop a Measure of Primary School Classroom Learning Culture.* Paper presented at the 2003 Annual

Conference of the Australian Association for Research in Education, Auckland, New Zealand.

Charteris-Black, J. (2002). Second Language Figurative Proficiency: A Comparative Study of Malay and English. *Applied Linguistics,* 23(1), 104-133.

Chen, Y.H., Gorin, J.S., Thomson, M.S., & Tatsuoka, K.K. (2008). Cross-Cultural Validity of the TIMSS-1999 Mathematics Test: Verification of a Cognitive Model. *International Journal of Testing,* 8, 251-271.

Cincotta, R.P.,Wisnewski, J., & Engelman, R. (2000). Human Population in the Biodiversity Hotspots. *Nature,* 404, 990 – 992.

Clark, A.M. (1998). The Qualitative-Quantitative Debate: Moving from Positivism and Confrontation to Post-Positivism and Reconciliation. *Journal of Advanced Nursing,* 27, 1242 – 1249.

Cochran-Smith, M. (1995). Color Blindness and Basket Making Are Not the Answers: Confronting the Dilemmas of Race, Culture, and Language Diversity in Teacher Education. *American Educational Research Journal*, 32(3), 493-522.

Cohen, J.E. (2003). The State of the Planet. Human Population: The Next Half Century. *Science,* 302, 1172 – 1175.

Cohen, L., Manion, L. & Morrison, K. (2000). *Research Methods in Education.* (5th ed.). London: RoutledgeFalmer.

Crawford-Lange, L.M. & Lange, D.L. (1987). Integrating Language and Culture: How to Do It. *Theory into Practice,* 26(4), 258-266.

Crossley, M. (2000). Bridging Cultures and Traditions in the Reconceptualisation of Comparative and International Education. *Comparative Education*, 36(3), 319-332.

Creswell, J.W. (2005). Educational Research. Planning, conducting, and Evaluating Quantitative and Qualitative Research. New Jersey: Prentice-Hall, Inc.

Cronbach, L.J. & Meehl, P.E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin,* 52, 281-302.

Currie, J. (2001). Early Childhood Education Programs. *Journal of Economic Perspectives,* 15(2), 213-238.

De Beer, M. (2004). Use of Differential Item Functioning (DIF) Analysis for Bias Analysis in Test Construction. *SA Journal of Industrial Psychology,* 30(4), 52-58.

Debes, J.L., & Williams, C.M. (1974). The Power of Visuals. *Instructor,* 84(4), 31-38.

DeLoache, J.S. (1991). Symbolic Functioning of Very Young Children: Understanding of Pictures and Models. *Child Development,* 62(4), 736-752.

Department of Education. (2001). Policy on Assessment and Qualifications for Schools in the General Education and Training Band. Pretoria: Government Printers.

Department of Education. (2002a). *Revised National Curriculum Statement, Grades R– 9, Languages.* Pretoria: Government Printers.

Department of Education. (2002b). National Curriculum Statement, General Education and Training, Assessment Guidelines for Foundation Phase Grades R – 3. Pretoria: Government Printers.

Department of Education. (2002c). Revised National Curriculum Statement, Grades R-3, Schools, Foundation Phase, Pretoria: Government Printers.

Department of Education. (2003). Draft white paper on e-education: transforming learning and teaching through information and communication technologies. Pretoria: Government Printers.

Department of Education. (2006). *Education Statistics in South Africa*. Pretoria: Government Printers.

Department of Education. (2008). *Teaching Reading in the Early Grades, A Teacher's Handbook.* Pretoria: Government Printers.

De Witt, M.W., & Booysen, M.I. (1995). *Socialization of the young child.* Van Schaik Publishers, Pretoria, RSA.

Di Leo, J.H. (1983). *Interpreting Children's Drawings.* Brunner/Mazel Publishers, New York, USA.

Dinero, T.E. & Haertel, E. (1977). Applicability of the Rasch Model with Varying Item Discriminations. *Applied Psychological Measurement,* 1(4), 581-592.

Downing, S.M. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education,* 37, 830-837.

Downing, S.M., & Haladyna, T.M. (2004). Validity Threats: Overcoming Interference with Proposed Interpretations of Assessment Data. *Medical Education,* 38, 327-333.

Dowse, R., & Ehlers, M.S. (2001). The Evaluation of Pharmaceutical Pictograms in a Low-Literate South African Population. *Patient Education and Counselling,* 45(2), 87-99.

Dowse, R., & Ehlers, M.S. (2005). Medicine Labels Incorporating Pictograms: Do they influence understanding and adherence? *Patient Education and Counselling,* 58(1), 63-70.

Education in SA, (2009). Retrieved February 3, 2009 from www.southafrica.info/about/education

Eloff, I. & Ebersohn, L. (2004). *Keys to Educational Psychology.* Cape Town: UCT Press.

Ely, R. (2005). Language and Literacy in the School Years. In *The Development of Language,* 395-443. Boston: Allyn and Bacon.

Embretson, S.E. (2007). Construct Validity: A Universal Validity System or Just Another Test Evaluation Procedure? *Educational Researcher,* 36(8), 449-455.

Embretson, S. & Gorin, J. (2001). Improving Construct Validity with Cognitive Psychology Principles. *Journal of Educational Measurement,* 38(4), 343-368.

Engelhard, G & Osberg, D.W. (1983). Constructing a Test Network with a Rasch Measurement Model. *Applied Psychological Measurement,* 7(3), 283-294.

Erickson, F., & Gutierrez, K. (2002). Culture, Rigor, and Science in Educational Research. *Educational Researcher,* 31(8), 21-24.

Fan, X. (1998). Item Response Theory and Classical Test Theory: An Empirical comparison of the Item/Person Statistics. *Educational and Psychological Measurement,* 58(3), 357-381.

Fischer, W.P. (1991, April). *The Rasch Debate: Validity and Revolution in Educational Measurement.* Paper presented at the International Objective Measurement Workshop, Chicago. Retrieved June 5, 2008, from Pretoria University Database.

Fraenkel, J.R. & Wallen, N.E. (2003). *How to Design and Evaluate Research in Education* (5th ed.). Boston: McGraw Hill.

Frederiksen, J.R. & Collins, A. (1989). A Systems Approach to Educational Testing. *Educational Researcher,* 18(9), 27-32.

Freedle, R. & Kostin, I. (1990). Item Difficulty of Four Verbal Item Types and an Index of Differential Item Functioning for Black and White Examinees. *Journal of Educational Measurement,* 27(4), 329-343.

Furstenberg, G., Levet, S., English, K., & Maillet, K. (2001). Giving a Virtual Voice to the Silent Language of Culture: The Cultura Project. *Language, Learning and Technology,* 5(1), 55-102.

Garcia, E.E. (1993). Language, Culture, and Education. *Review of Research in Education,* 19, 51-98.

Garrison, J. (1994). Realism, Deweyan Pragmatism and Educational Research. *Educational Researcher,* 23(1), 5-14.

Gardner, H. (2003, April 21). *Multiple Intelligences After Twenty Years.* Paper presented at the 2003 American Educational Research Association, Chicago, Illinois. Retrieved February 26, 2009, from Pretoria University Database.

Gay, L.R., & Airasian, P. (2003). *Educational Research. Competencies for Analysis and Applications.* (7th ed.). New Jersey: Prentice-Hall, Inc.

Glewwe, P., Jacoby, H.G. and King, E.M. (2001). Early childhood nutrition and academic achievement: a longitudinal analysis. *Journal of Public Economics,* 81, 345-368.

Gorin, J.S. (2007). Reconsidering Issues in Validity Theory. Educational Researcher, 36(8), 456-462.

Grant, R.A., & Wong, S.D. (2003). Barriers to literacy for language-minority learners: An argument for change in the literacy education profession. International Reading Association, 386-394.

Graziano, A.M. & Raulin M.L. (2000). *Research Methods. A process of inquiry.* (4th ed.). Boston: Allyn and Bacon.

Green, B.F., Crone, C.R. & Folk, V.G. (1989). A Method for Studying Differential Distractor Functioning. *Journal of Educational Measurement,* 26(2), 147-160.

Gregory, R.J. (2000). Psychological Testing, History, Principles, and Applications (3rd ed.). Boston: Allyn and Bacon.

Grunewald, R. & Rolnick, A. (2006). *A Proposal for Achieving High Returns on Early Childhood Development.* Federal Reserve Bank of Minneapolis. Retrieved January 21, 2010, from University of Pretoria database.

Guba, E.G. (1990). *The Paradigm Dialog.* Indiana: Sage Publications, Inc.

Guba, E.G., & Lincoln, Y. S. (1994). Handbook of Qualitative Research. Competing Paradigms in Qualitative Research. Indiana: Sage Publications, Inc.

Guiso, L., Sapienza, P., & Zingales, L. (2006). Does Culture Affect Economic Outcomes? The Journal of Economic Perspectives, 20(2), 23-48.

Gutiérrez, K.D. & Rogoff, B. (2003). Cultural Ways of Learning: Individual Traits or Repertoires of Practice. Educational Researcher, 32(5), 19-25.

Gutierrez, K.D., Asato, J., Santos,M., & Gotando, N. (2002). Backlash Pedagogy: Language and Culture and the Politics of Reform. *The Review of Education, Pedagogy, and Cultural Studies,* 24, 335-351.

Habib, A. (1996). Myth of the Rainbow Nation: Prospects for the Consolidation of Democracy in South Africa. *Security Review,* 5(6).

Hagquist, C., & Andrich, D. (2004). Is the Sense of Coherence-Instrument Applicable on Adolescents? A latent trait analysing using Rasch Modeling. *Personality and Individual Differences,* 36, 955-968.

Hart, S.L. (1971). Axiology – Theory of Values. *Philosophy and Phenomenological Research,* 32(1), 29-41.

Hauger, J.B., & Sireci, S.G. (2008). Detecting Differential Item Functioning across Examinees tested in their dominant language and Examinees tested in a Second Language. International Journal of Testing, 8, 237-250.

Hawthorne, K. & Tomlinson, S. (1997). One-to-One teaching with pictures – flashcard health education for British Asians with Diabetes. *British Journal of General Practice,* 47, 301-304.

Heckman. J.J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. Science, 312, 1899 - 1903.

Henson, R.K. (1999). *Understanding the One-Parameter Rasch model of Item Response Theory.* Paper presented at the annual meeting of the South West Educational Research Association, San Antonio. Retrieved June 2008 from University of Pretoria database.

Hobsbawm, E. (1996). Language, Culture and National Identity. *Social Research,* 63(4), 1065 – 1080.

Hoffmann, V. (2000). *Picture supported communication in Africa.* Weikersheim: Margraf Verlag.

Houts, P.S., Doak, C.C., Doak, L.G., & Loscalzo, M.J. (2006). The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Education and Counseling,* 61, 173–190.

Howell, D.C. (2002). *Statistical Methods for Psychology.* (5th ed.). Boston: PWS-Kent Publishing Company.

Howie, S.J., Venter, E., Van Staden, S., Zimmerman, L., Long, C., Scherman, V., & Archer, E. (2007). Progress in International Reading Literacy Study 2006.

Summary report. South African children's reading literacy achievement. Centre for Evaluation and Assessment: University of Pretoria.

Huang, C.D., Church, A.T., & Katigbak, M.S. (1997). Identifying Cultural Differences in Items and Traits: Differential Item Functioning in the NEO Personality Inventory. *Journal of Cross-Cultural Psychology,* 28(2), 192-218.

Huchzermeyer, M. (2004). A Legacy of Control/ The Capital Subsidy for Housing, and Informal Settlement Intervention in South Africa. *International Journal of Urban and Regional Research,* 27(3), 591-612.

Jacobs, L.A., & Giarelli, E. (2001). Jewish Culture, Health Belief Systems, and Genetic Risk of Cancer. *Nursing Forum,* 36(2), 5 – 13.

Janks, H. (2000). Domination, Access, Diversity and Design: a synthesis for critical literacy education. Educational Review, 52(2), 175-186.

Ji, L.J., Zhang, Z., & Nisbett, R.E. (2004). Is it Culture or is it Language? Examination of Language Effects in Cross-Cultural Research on Categorization. *Journal of Personality and Social Psychology,* 87(1), 57-65.

John-Steiner, V., & Mahn, H. (1996). Sociocultural Approaches to Learning and Development: A Vygotskian Framework. Educational Psychologist, 31(3), 191-206.

Kickbusch, I.S. (2001). Health Literacy: Addressing the Health and Education Divide. Health Promotion International, 16(3), 289–297.

Kidd, S.A. (2002). The Role of Qualitative Research in Psychological Journals. *Psychological Methods,* 7(1), 126–138.

Killen, R. 2002. Outcomes-based Education: Principles and Possibilities. *Interpretations*, 35(1), 1-18.

Kim, S.H., Cohen, A.S., & Park, T.H. (1995). Detection of Differential Item Functioning in Multiple Groups. *Journal of Educational Measurement,* 1995, 32(3), 261-276.

Kirsten, M (2004). Multilingual/Multicultural Aspects of Visual Literacy and Interpretation in Multimodal Educational Communication. PhD Thesis, University of the Western Cape.

Kpelle. Encyclopaedia Britannica. (2006). Retrieved February 20, 2009, from Encyclopaedia Britannica Ultimate Reference Suite DVD.

Kuhl, P.K., Tsao, F.M., Liu, H.M., Zhang, Y., & de Boer, B. (2001). Language/Culture/Mind/Brain, Progress at the Margins between Disciplines. *Annals of the New York Academy of Sciences* , 935, 136 – 174.

Kyriakides, L., Kaloyirou, C., & Lindsay, G. (2006). An analysis of the Revised Olweus Bully/Victim Questionnaire using the Rasch measurement model. *British Journal of Educational Psychology,* 76, 781–801.

Lather, P. (2006). Paradigm Proliferation as a Good Thing to Think with: Teaching Research in Education as a wild profusion. *International Journal of Qualitative Studies in Education,* 19(1), 35-57.

Linacre, J.M. (1993). *Generalizability Theory and Many-facet Rasch Measurement.* Paper presented at the 1993 Annual Meeting of the American Eduacational Research Association, Atlanta, Georgia. Retrieved May 2008 from University of Pretoria database.

Linacre, J.M. (2009). WINSTEPS Rasch Measurement Computer Program. Chicago: Winsteps.com. Retrieved July 21, 2009 from http://www.winsteps.com.

Linn, R.L. (1998). State-by-State Comparisons of Achievement: Suggestions for Enhancing Validity. *Educational Researcher,* 17(3), 6-9.

Linn, R.L., & Gronlund, N.E. (2000). *Measurement and Assessment in Teaching.* (8th ed.). New Jersey: Prentice-Hall, Inc.

Loving, C.C. (1997). From the summit of Truth to its slippery slopes: Science education's journey through Positivist-Postmodern Territory. *American Educational Research Journal,* 34(3), 421 – 452.

Low, M.D., Low, B.J., Baumler, E.R., & Huynh, P.T. (2005). Can Education Policy be Health Policy? Implications of Research on the Social Determinants of Health. Journal of Health Politics, Policy and Law. 30(6), 1131–1162.

Lundgren-Nilsson, A., Grimby, G., Ring, H., Tesio, L., Lawton, G., Slade, A., Penta, M., Tripolski, M., Biering-Sørensen, F., Carter, J., Marincek, C., Phillips, S., Simone, A., & Tennant, A. (2005). Cross-Cultural Validity of Functional Independence Measure Items in Stroke: A Study Using Rasch Analysis. *J Rehabil Med,* 37, 23–31.

Mahoney, K. (2008). Linguistic Influences on Differential Item Functioning for Second Language Learners on the National Assessment of Educational Progress. *International Journal of Testing,* 8, 14-33.

Maller, S.J. (2001). Differential Item Functioning in Wisc-III: Item Parameters for Boys and Girls in the National Standardization Sample. *Educational and Psychological Measurement,* 61(5), 793-817.

Maree,J.G. & Fraser, W.J. (Eds.). (2004). *Outcomes Based Assessment.* Sandown: Heinemann Publishers.

Marley,S.C., Levin, J.R. & Glenberg, A.M. (2010). What Cognitive Benefits Does an Activity-Based Reading Stategy Afford Young Native American Readers? *The Journal of Experimental Education,* 78, 395-417.

Mathews, M.R. (2004). Reappraising Positivism and Education: The arguments of Phillip Frankl and Herbert Feigl. *Science & Education,* 13, 7-39.

McCamey, R. (2002). *A Primer on the One-Parameter Rasch model.* Paper presented at the 25th Annual Meeting of the Southwest Educational Research Association, Austin,TX. Retrieved June 2008 from University of Pretoria database.

McDougall, J.K. (2004). Changing mindsets: A study of Queensland primary teachers and the visual literacy initiative. PhD Thesis. Central Queensland University. July 2004.

McMillan, J.H., & Scumacher, S. (2006). *Research in Education, Evidence-Based Enquiry.* (6th ed.). Allyn and Bacon, USA.

McMillan, J.H. & Schumacher, S. (2001). *Research in Education: A Conceptual Introduction* (5th ed.). Boston: Longman

Messick, S (1981). Evidence and Ethics in the Evaluation of Tests. *Educational Researcher,* 10(9), 9-20.

Messick, S. (1989). Meaning and Values in Test Validation, The Science and Ethics of Assessment. *Educational Researcher,* 18(2), 5-11.

Messick, S. (1995). Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry not Score Meaning. *American Psychologist, 50* (9), 741-749.

Mills, J. (2001). Being Bilingual: Perspectives of Third Generation Asian Children on Language, Culture and Identity. *International Journal of Bilingual Education and Bilingualism,* 4(6), 383 – 402.

Mislevy, R.J. (2007). Validity by Design. *Educational Researcher,* 36(8), 463-469.

Merrell, C., & Tymms, P. (2002, September). Tackling challenging children. *Literacy Today.*

Merrell, C., & Tymms, P. (2005a). International Comparisons of Early Lliteracy. *Literacy Today.*

Merrell, C., & Tymms, P. (2005b). The Impact of Early Interventions and Pre-School Experience on the Cognitive Development of Young Children in England. *Paper presented at AERA Annual general meeting,* Montreal.

Moore, D.M., & Dwyer, F.M. (1994). A Spectrum of Visual Learning. *Educational Technology Publications,* 4, 1-18. Retrieved May 26, 2008 from Pretoria University Database.

Moss, P.A. (1994). Can there be Validity without Reliability? *Educational Researcher,* 23(2), 5-12.

Moss, P.A. (1992). Shifting Conceptions of Validity in the Educational Measurement: Implications for Performance Assessment. *Review of Educational Research,* 62(3), 229-258.

Moss, P.A., Girard, B.J., & Haniford, L.C. (2006). Chapter 4: Validity in Educational Assessment. Review of Research in Education, 30, 109 – 162.

Nair, M.K.C. (2004). Early Childhood Development in Deprived Urban Settlements. *Indian Pediatrics,* 41, 227 – 237.

Neuman, W.L. (2003). Social Research Methods: Qualitative and Quantitative Approaches (5th ed.). Boston: Allyn & Bacon.

O'Neill, K.A., & McPeek, W.M. (1989). Item and Test Characteristics That Are Associated with Differential Item Functioning. In *Differential Item Functioning: Theory and Practice, October 1989,* (pp. 255-276). Princeton, New Jersey: Educational Testing Services.

Opie, C. (2005). Doing Educational Research. A Guide to First Time Researchers. London: Sage Publications.

Pallant, J.F. & Tennant, A. (2007). An Introduction to the Rasch measurement model: An Example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology,* 46, 1-18.

Papalia, D.E., Olds, S.W., & Feldman, R.D. (2002). *A Child's World - Infancy through Adolescence.* (9th ed). New York:, McGraw-Hill Higher Education.

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction,* 3(3), 227-238.

Pendlebury, S., Lake, L., & Smith, C. (Eds.). (2009). Children's Right to Basic Education. *South African Child Gauge,* 19-23.

Pesudovs, K., Burr, J.M. Harley, C., & Elliot, D.B. (2007). The Development, Assessment, and Selection of Questionnaires. Optometry and Vision Science, 84(8), 663-674.

Pettersson, R. (1998, October 21-25). *Image functions in Information Design.* Paper presented at the 30th Annual Conference of the International Visual Literacy Association. University of Georgia, Georgia. Retrieved May 26, 2008 from Pretoria University Database.

Phillips, D.C., & Burbules, N.C. (2000). *Postpositivism and Educational Research.* Lanham: Rowman & Littlefield.

Ponterotto, J.G. (2005). Qualitative Research in Counseling Psychology: A primer on Research Paradigms and Philosophy of Science. *Journal of Counseling Psychology,* 32(2), 126 – 136.

Popham, W.J. (2003). *Test Better, Teach Better: The Instructional Role of Assessment* (3rd ed.). Alexandria: ASCD.

Popham, W.J. (1999). *Classroom Assessment: What teachers need to know.* (2nd ed.). Boston: Allyn and Bacon.

Pretorius, E.J. (2009). Putting reading in Northern Sotho on track in the early years: Changing resources, expectations and practices in a high poverty school. *South African Journal of African Languages,* 1, 54-72.

Ramey, C.T., Campbell, F.A., Burchinal, M., Skinner, M.L., Gardner, D.M., & Ramey, S.L. (2000). Persistent Effects of Early Childhood Education on High-Risk Children and Their Mothers. *Applied Developmental Science,* 4(1), 2-14.

Raney, K (1999). Visual Literacy and the Art Curriculum. NSEAD.

Rasch, G. (n.d.). *On General Laws and the Meaning of Measurement in Psychology.* Fourth Berkely Symposium. The Danish Institute of Educational Research, Copenhagen.

Ridgen, C. (1999). 'The Eye of the Beholder' – Designing for Colour-Blind Users. *British Telecommunications Engineering,* 17, 2-6.

Roscigno, V.J., & Ainsworth-Darnell, J.W. (1999). Race, Cultural Capital, and Educational Resources: Persistent Inequalities and Achievement Returns. *Sociology of Education.* 72(3), 158-178.

Rost, J. & Carstensen, C.H. (2002). Multidimensional Rasch Measurement via Item Component Models and Faceted Designs. *Applied Psychological Measurement,* 26(1), 42-56.

Rowntree, D. (1990). *Teaching Through Self-Instruction.* London: Kogan Page.

SAQA. (2009). *Information Provision Manual.* Pretoria, Gauteng. South African Qualifications Authority.

Shanahan, D. (1997). Articulating the Relationship between Language, Literature, and Culture: Toward a New Agenda for Foreign Language Teaching and Research. *The Modern Language Journal,* 81(2), 164-174.

Scherman, V. (2007). *The Validity of Value-Added Measures in Secondary Schools.* PhD Thesis. University of Pretoria. March 2007.

Scherman, V., Archer, E., Howie, S.J., & Lopez, T. (2006). National Research Foundation: Investigating Value-added assessment in South Africa, *Technical report year 2003/2004.* Pretoria: Centre for Evaluation and Assessment.

Scherman, V., Archer, E., & Howie, S.J. (2006). 'Locating the Local Village within the Global Village. Assessment Possibilities and Practical Challenges.' *4th Sub-Regional Conference on Assessment in Education.* Hosted by Umalusi. University of Johannesburg.

Scheuneman, J.D., & Gerritz, K. (1990). Using Differential Item Functioning Procedures to Explore Sources of Item Difficulty and Group Performance Characteristics. *Journal of Educational Measurement,* 27(2), 109-131.

Schieffelin, B.B. & Ochs, E. (1986). Language Socialization. *Annual Review of Anthropology,* 15, 163-191.

Schonborn K.J. & Anderson T.R. (2008). A Model of Factors Determining Students' Ability to Interpret External Representations in Biochemistry. *International Journal of Science Education,* 1-40.

Schonborn K.J. & Anderson T.R. (2006). The Importance of Visual Literacy in the Education of Biochemists. *Biochemistry and Molecular Biology Education.* 34(2), 94-102.

Schonborn, K.J., Anderson, T.R., & Grayson, D.J. (2002). Student Difficulties with the Interpretation of A Textbook Diagram of Immunoglobulin G. *Biochemistry and Molecular Biology Education.* 30(2), 93-97.

Schulze, S. (2003). Views on the combination of quantitative and qualitative research approaches. *Progressio,* 25(2), 8 – 20.

Schumacker, R. E. (2004). Rasch measurement: The Dichotomous Model. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement*, 226–257. Maple Grove: JAM Press.

Sibula, P.M. (2007). Furthering the Aim of Multi-lingualism through Integrated Terminology Development. *Lexikos.* 17, 397-406.

Sick, J. (2008). Rasch measurement in Language Education: Part 1. *Shiken:JALT Testing and Evaluation Newsletter,* 12(1), 1-6.

Silver, H., Strong, R., & Perini, M. (1997, September). Integrating Learning Styles and Multiple Intelligences. *Educational Leadership, pp. 22-27.*

Sims, E., O'Leary, R., Cook, J., & Butland, G. (2002). Visual Literacy: What is it and do we need it to use Learning Technologies effectively? University of Bristol: UK.

Smit, W. (1998). The Rural Linkages of Urban Households in Durban, South Africa. *Environment and Urbanization,* 10(1), 77-87.

Smith, E.V. & Smith, R.M. (2004). Introduction to Rasch Measurement. Theory, Models and Applications. Maple Grove: Jam Press.

Statistics SA (2006) *Mid-year population estimates, South Africa.* Pretoria: Statistics South Africa.

Stokes, S. (2001). Visual Literacy in Teaching and Learning: A Literature Perspective. *Electronic Journal for the Integration of Technology in Education,* 1-7, College of Education, Idaho University.

Sullivan, M., Karlsson, J., & Ware, J.E. (1995). The Swedish SF-36 Health Survey-1. Evaluation of Data Quality, Scaling Assumptions, Reliability and Construct Validity across general Populations in Sweden. *Soc. Sci. Med.* 41(10), 1349-1358.

Teachman, J.D. (1987). Family Background, Educational Resources, and Educational Attainment. *American Sociological Review,* 52(4), 548-557.

Tennant, A. & Conaghan, P.G. (2007). The Rasch Measurement Model in Rheumatology: What is it and Why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism (Arthritis Care and Research),* 57(8), 1358-1362.

Tennant, A.& Pallant, J.F. (2007). DIF matters: A practical approach to test if Differential Item Functioning makes a difference. *Rasch Measurement Transactions*, 20(4), 1082-84.

Tennant, A.& Pallant, J.F. (2006). Unidimensionality Matters! (A Tale of Two Smiths?) *Rasch Measurement Transactions*, 20(1), 1048-51.

Thorndike, R.M. (1997). Measurement and Evaluation in Psychology and Education. Columbus: Ohio.

Tseng, Y.H. (2002). A lesson in culture. *ELT Journal,* 56(1), 11 – 21.

Tymms, P. (1999). Baseline assessment, value-added and the prediction of reading. *Journal of Research in Reading,* 22(1), 27-36.

Tymms, P. (2001). The development of a computer-adaptive assessment in the early years. *Educational and Child Psychology,* 18(3), 20-30.

Tymms, P. (2004). Are standards rising in English primary schools? *British Educational Research Journal,* 30(4), 477-494.

Tymms, P., & Albone, S. (2002). Performance indicators in primary schools. In A.J. Visscher and R. Coe (Eds), *School improvement through performance feedback.* Lisse: Swets & Zeitlinger Publishers (pp. 191-220).

Tymms, P., & Coe, R. (2003). Celebration of the Success of Distributed Research with Schools: the CEM Centre, Durham. *British Educational Research Journal,* 29(5), 639-653.

Tymms, P., & Gallacher, S. (1995). Primary Science: An exploration of differential classroom success. *Research in Science & Technological Education,* 13(2), 155-161.

Tymms, P., Merrell, C., & Henderson, B., (1997). The First Year at School: A Quantitative Investigation of the Attainment and Progress of Pupils. *Educational Research and Evaluation, 3*(2), 101-118.

Tymms, P., Merrell, C., & Henderson, B., (2000). Baseline Assessment and Progress during the First Three Years at School. *Educational Research and Evaluation, 6*(2), 105-129.

Tymms, P., Merrell, C., & Jones, P. (2004). Using Baseline Assessment Data to make International Comparisons. *British Educational Research Journal, 30* (5), 673-689.

Tymms, P., & Wylde, M. (2003, April). *Baseline assessment and monitoring in primary schools.* Paper presented at the Symposium Connectable Processes in Elementary and Primary Section. Bamberg.

Vandeyar, S. & Killen, R. (2003). Has Curriculum Reform in South Africa really changed Assessment Practices, and what promise does the Revised National Curriculum Statement hold? *Perspectives in Education,* 21(1), 119-134.

Ukrainetz, T.A. & Blomquist, C. (2002). The Criterion Validity of Four Tests Compared with a Language Sample. *Child Language Teaching and Therapy,* 18(59), 59-78.

Van den Berg, A.R., & Vorster, J.F. (Eds.). (1982). *Basiese Psigometrika.* Pretoria: V & R Drukkery.

van der Vijver, F.J.R. & Poortinga, Y.H. (1997). Towards an Integrated Analysis of Bias in Cross-Cultural Assessment. *European Journal of Psychological Assessment,* 13 (1), 29-37.

van de Vijver, F., & Tanzer, N.K. (2004). Bias and Equivalence in Cross-Cultural Assessment: An Overview. *Revue européenne de psychologie appliqué*, 54, 119 - 135.

Wang, W. & Wilson, M. (2005). Assessment of Differential Item Functioning in Testlet-Based Items using the Rasch Testlet Model. *Educational and Psychological Measurement,* 65(4), 549-576.

Wardlow, G. (1989). Alternative Methods of Inquiry for Agricultural Education. *Journal of Agricultural Education,* pp. 1–5.

Waugh, R.F. (1999). Approaches to Studying for Students in Higher Education: A Rasch Measurement Model Analysis. *British Journal of Educations Psychology,* 69, 63-79.

Webb, V., & Read, J. (2000). 'The Challenge of Cultural Heterogeneity to Educational Development.' Ethno linguistic Factors, LAOIII Conference.

Weber, S., & Mitchell, C. (1996). Drawing ourselves into teaching: Studying the Images that Shape and Distort Teacher Education. *Teaching and Teacher Education, 12(3), 303-313.*

Wehmeier, S. (Ed.). (2010). *Oxford Advanced Learner's Dictionary.* (7th ed.). Oxford: Oxford University Press.

Whelan, D.L. (2004). Picture Perfect. *School Library Journal.* June, p.48-50.

William, D. (2006). Formative Assessment: Getting the Focus right. *Educational Assessment,* 11(3), 283-289.

Williams, T.L., (2007). "Reading" the painting: Exploring Visual Literacy in the Primary Grades. *The Reading Teacher,* 60 (7), 636-642.

WINSTEPS, (2009). Winsteps Rasch Measurement, Copyright 2009, John M. Linacre. Version 3.68.1.

Wise, S.L., & DeMars, C.E. (2005). Low Examinee Effort in Low-Stakes Assessment: Problems and Potential Solutions. *Educational Assessment,* 10(1), 1-17.

Wright, B.D. (1977). Solving Measurement Problems with the Rasch Model. *Journal of Educational Measurement,* 14(2), 97-116.

Wright, B.D. & Mok, M.M.C. (2004). An Overview of the Family of Rasch Measurement Models. Introduction to Rasch Measurement. *Applied Psychological Measurement,* 30, 443-446.

Wyse, A.E., & Mapuranga, R. (2009). Differential Item Functioning Analysis using Rasch Item Information Functions. *International Journal of Testing,* 9, 333-357.

Yenawine, P. (2003). Jump Starting Visual Literacy. *Art Education.* p. 7-12.

Zwick, R., & Ercikan, K. (1989). Analysis of Differential Item Functioning in the NAEP History Assessment. *Journal of Educational Measurement,* 26(1), 55-66.

Zwick, R., Donoghue, J.R., & Grima, A. (1993). Assessment of Differential Item Functioning for Performance Tasks. *Journal of Educational Measurement,* 30(3), 233-251.

9 APPENDICES

9.1 APPENDIX A – PICTURE 1 ITEMS

Pictures

- Show the picture on the opposite page to the child. Ask him/her to point to some **carrots** and wait for their response.
- Continue this procedure, asking for:

  the knife, (either knife is correct),
  a fork, (either fork is correct),
  a cupboard,
  some cherries,
  a pan, (the wok is also correct),
  a bowl.

- On the Pupil Record Sheet put a line through the bracket underneath each picture correctly identified. If the child could not identify any of the pictures put a line through the bracket underneath 'no score'.
- If the score was less than 3 go straight to **Ideas about Reading** otherwise continue with the next page.

9.2 APPENDIX B – ITEM - LEARNER STATISTICS – ALL LEARNERS

TABLE 6.1 Baseline Picture Vocabulary Test ZOU768WS.TXT Aug 5 11:46 2010
INPUT: 1362 Learners 22 Objects MEASURED: 1221 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 8.01 REL.: .98

Learner STATISTICS: MISFIT ORDER ALL 3 LANGUAGES PVT 1

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
138	5	7	-1.07	1.42	2.87	2.4	7.54	2.5	A .51	.80	33.3	75.7	10010122
696	5	7	-1.07	1.42	2.87	2.4	7.54	2.5	B .51	.80	33.3	75.7	3010061
353	6	7	1.00	1.52	3.32	2.0	5.56	2.2	C .25	.60	33.3	81.3	18010391
362	6	7	1.00	1.52	3.32	2.0	5.56	2.2	D .25	.60	33.3	81.3	18010481
531	6	7	1.00	1.52	3.32	2.0	5.56	2.2	E .25	.60	33.3	81.3	12010992
773	6	7	1.00	1.52	3.32	2.0	5.56	2.2	F .25	.60	33.3	81.3	16010051
775	6	7	1.00	1.52	3.32	2.0	5.56	2.2	G .25	.60	33.3	81.3	16010081
803	6	7	1.00	1.52	3.32	2.0	5.56	2.2	H .25	.60	33.3	81.3	16010411
809	6	7	1.00	1.52	3.32	2.0	5.56	2.2	I .25	.60	33.3	81.3	16010471
817	6	7	1.00	1.52	3.32	2.0	5.56	2.2	J .25	.60	33.3	81.3	16010551
831	6	7	1.00	1.52	3.32	2.0	5.56	2.2	K .25	.60	33.3	81.3	16010741
842	6	7	1.00	1.52	3.32	2.0	5.56	2.2	L .25	.60	33.3	81.3	16010881
953	6	7	1.00	1.52	3.32	2.0	5.56	2.2	M .25	.60	33.3	81.3	13010233
10	5	6	.81	1.64	3.08	1.9	3.08	1.9	N .48	.70	.0	75.5	4010081
30	4	5	.81	1.64	3.08	1.9	3.08	1.9	O .45	.68	.0	75.5	4010301
35	4	5	.81	1.64	3.08	1.9	3.08	1.9	P .45	.68	.0	75.5	4010361
54	4	5	.81	1.64	3.08	1.9	3.08	1.9	Q .45	.68	.0	75.5	4010551
60	5	6	.81	1.64	3.08	1.9	3.08	1.9	R .48	.70	.0	75.5	4010621
110	5	6	.81	1.64	3.08	1.9	3.08	1.9	S .48	.70	.0	75.5	5010293
457	4	5	.81	1.64	3.08	1.9	3.08	1.9	T .45	.68	.0	75.5	8030732
538	4	5	.81	1.64	3.08	1.9	3.08	1.9	U .45	.68	.0	75.5	23010011
544	5	6	.81	1.64	3.08	1.9	3.08	1.9	V .48	.70	.0	75.5	23010081
547	4	5	.81	1.64	3.08	1.9	3.08	1.9	W .45	.68	.0	75.5	23010121
549	5	6	.81	1.64	3.08	1.9	3.08	1.9	X .48	.70	.0	75.5	23010141
551	5	6	.81	1.64	3.08	1.9	3.08	1.9	Y .48	.70	.0	75.5	23010151

552	5	6	.81	1.64	3.08	1.9	3.08	1.9	Z .48	.70	.0	75.5	23010161
	BETTER	FITTING	OMITTED										
1171	6	7	1.00	1.52	.30	-.9	.24	-.4	z .69	.60	100.0	81.3	22010343
1172	6	7	1.00	1.52	.30	-.9	.24	-.4	y .69	.60	100.0	81.3	22010353
1173	6	7	1.00	1.52	.30	-.9	.24	-.4	x .69	.60	100.0	81.3	22010363
1174	6	7	1.00	1.52	.30	-.9	.24	-.4	w .69	.60	100.0	81.3	22010373
1178	6	7	1.00	1.52	.30	-.9	.24	-.4	v .69	.60	100.0	81.3	22010413
1180	6	7	1.00	1.52	.30	-.9	.24	-.4	u .69	.60	100.0	81.3	22010433
1181	6	7	1.00	1.52	.30	-.9	.24	-.4	t .69	.60	100.0	81.3	22010453
1182	6	7	1.00	1.52	.30	-.9	.24	-.4	s .69	.60	100.0	81.3	22010463
1188	6	7	1.00	1.52	.30	-.9	.24	-.4	r .69	.60	100.0	81.3	22010573
1193	6	7	1.00	1.52	.30	-.9	.24	-.4	q .69	.60	100.0	81.3	11010073
1196	6	7	1.00	1.52	.30	-.9	.24	-.4	p .69	.60	100.0	81.3	11010103
1197	6	7	1.00	1.52	.30	-.9	.24	-.4	o .69	.60	100.0	81.3	11010113
1199	6	7	1.00	1.52	.30	-.9	.24	-.4	n .69	.60	100.0	81.3	11010133
1202	6	7	1.00	1.52	.30	-.9	.24	-.4	m .69	.60	100.0	81.3	11010173
1206	6	7	1.00	1.52	.30	-.9	.24	-.4	l .69	.60	100.0	81.3	11010223
1207	6	7	1.00	1.52	.30	-.9	.24	-.4	k .69	.60	100.0	81.3	11010233
1209	6	7	1.00	1.52	.30	-.9	.24	-.4	j .69	.60	100.0	81.3	11010253
1213	6	7	1.00	1.52	.30	-.9	.24	-.4	i .69	.60	100.0	81.3	11010303
1215	6	7	1.00	1.52	.30	-.9	.24	-.4	h .69	.60	100.0	81.3	11010323
1223	6	7	1.00	1.52	.30	-.9	.24	-.4	g .69	.60	100.0	81.3	11010463
1224	6	7	1.00	1.52	.30	-.9	.24	-.4	f .69	.60	100.0	81.3	11010453
1233	6	7	1.00	1.52	.30	-.9	.24	-.4	e .69	.60	100.0	81.3	11010603
1247	6	7	1.00	1.52	.30	-.9	.24	-.4	d .69	.60	100.0	81.3	11010743
1249	6	7	1.00	1.52	.30	-.9	.24	-.4	c .69	.60	100.0	81.3	11010773
1251	6	7	1.00	1.52	.30	-.9	.24	-.4	b .69	.60	100.0	81.3	11010793
1254	6	7	1.00	1.52	.30	-.9	.24	-.4	a .69	.60	100.0	81.3	11010823
MEAN	5.6	6.0	2.18	1.92	1.00	-.2	1.04	.0			77.3	77.7	
S.D.	1.3	1.2	1.16	.23	1.17	1.2	1.39	1.1			39.3	2.9	

9.3 APPENDIX C – PICTURE 1 OBJECT STATISTICS – ALL LANGUAGES

TABLE 10.1 Baseline Picture Vocabulary Test ZOU768WS.TXT Aug 5 11:46 2010
INPUT: 1362 Learners 22 Objects MEASURED: 1221 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 8.01 REL.: .98

Object STATISTICS: MISFIT ORDER FOR ALL 3 LANGUAGES pvt1

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
6	892	1218	1.94	.12	1.02	.3	1.07	1.0	A .80	.81	75.0	75.0	Pan
5	548	565	-1.63	.27	1.02	.2	1.03	.2	B .33	.34	90.2	90.2	Cherries
7	1095	1219	-.31	.12	.97	-.5	.96	-.6	a .56	.55	77.3	76.4	Bowl
MEAN	974.0	1040.7	-3.67	1.11	1.00	.0	1.02	.2			80.8	80.5	
S.D.	205.9	223.5	3.32	.82	.02	.4	.04	.6			6.7	6.9	

9.4 APPENDIX D LEARNER – ITEM STATISTICS AFRIKAANS PICTURE 1

TABLE 6.1 Baseline Picture Vocabulary Test ZOU658WS.TXT Aug 5 15:24 2010
INPUT: 1362 Learners 22 Objects MEASURED: 303 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 4.97 REL.: .96

"????????" Learner STATISTICS: MISFIT ORDER AFR PVT 1

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
531	6	7	3.77	1.38	2.45	1.7	1.62	.8	A .23	.54	71.4	89.6	12010992
457	4	5	3.52	1.52	2.25	1.7	.92	.4	B .46	.64	60.0	87.7	8030732
873	5	6	3.55	1.50	2.19	1.6	.77	.3	C .46	.63	66.7	89.5	6010262
1328	5	6	3.55	1.50	2.19	1.6	.77	.3	D .46	.63	66.7	89.5	17010122
138	5	7	2.09	1.26	2.13	1.6	1.68	.9	E .50	.71	71.4	87.0	10010122
903	6	7	3.77	1.38	1.87	1.2	.70	.2	d .40	.54	71.4	89.6	6010562
241	6	7	3.77	1.38	.39	-.9	.15	-.5	c .66	.54	100.0	89.6	19010012
526	6	7	3.77	1.38	.39	-.9	.15	-.5	b .66	.54	100.0	89.6	12010942
159	4	6	1.14	1.62	.14	-.7	.08	-.7	a .94	.82	100.0	92.8	10010362
MEAN	6.2	6.2	5.44	2.01	1.55	.8	.76	.1			78.6	89.4	
S.D.	.9	.9	.46	.10	.90	1.1	.56	.6			15.5	1.5	

9.5 APPENDIX E –OBJECT STATISTICS AFRIKAANS – PICTURE 1

TABLE 10.1 Baseline Picture Vocabulary Test ZOU148WS.TXT Oct 28 11:28 2009
INPUT: 1362 Learners 22 Objects MEASURED: 303 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 4.97 REL.: .96

Object STATISTICS: AFRIKAANS PICTURE 1 BASELINE MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
6	950	1306	4.33	.10	1.05	1.5	2.68	9.9	A .71	.74	70.5	70.6	Pan
7	1132	1307	2.67	.10	1.01	.2	1.03	.6	B .59	.59	72.7	73.0	Bowl
1	1250	1260	-1.52	.38	1.01	.1	.14	-2.4	C .34	.29	98.2	98.4	Carrots
5	559	587	1.55	.22	.94	-.3	.93	-.3	D .46	.44	87.6	86.9	Cherries
4	906	913	-.85	.42	.77	-.6	.23	-2.1	c .34	.24	98.4	98.0	Cupboard
3	1146	1148	-2.67	.74	.74	-.2	.06	-2.7	b .21	.13	99.5	99.5	Fork
2	1214	1215	-3.51	1.02	.67	-.1	.01	-5.3	a .17	.09	99.8	99.8	Knife
MEAN	1022.4	1105.1	.00	.43	.88	.1	.73	-.3			89.5	89.5	
S.D.	223.4	246.2	2.70	.32	.14	.6	.89	4.5			12.0	11.9	

9.6 APPENDIX F – OBJECT ORDER AFRIKAANS PICTURE 1

TABLE 13.1 Baseline Picture Vocabulary Test ZOU148WS.TXT Oct 28 11:28 2009
INPUT: 1362 Learners 22 Objects MEASURED: 303 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 4.97 REL.: .96

Object STATISTICS: MEASURE ORDER AFRIKAANS PVT1 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
6	950	1306	4.33	.10	1.05	1.5	2.68	9.9	.71	.74	70.5	70.6	Pan
7	1132	1307	2.67	.10	1.01	.2	1.03	.6	.59	.59	72.7	73.0	Bowl
5	559	587	1.55	.22	.94	-.3	.93	-.3	.46	.44	87.6	86.9	Cherries
4	906	913	-.85	.42	.77	-.6	.23	-2.1	.34	.24	98.4	98.0	Cupboard
1	1250	1260	-1.52	.38	1.01	.1	.14	-2.4	.34	.29	98.2	98.4	Carrots
3	1146	1148	-2.67	.74	.74	-.2	.06	-2.7	.21	.13	99.5	99.5	Fork
2	1214	1215	-3.51	1.02	.67	-.1	.01	-5.3	.17	.09	99.8	99.8	Knife
MEAN	1022.4	1105.1	.00	.43	.88	.1	.73	-.3			89.5	89.5	
S.D.	223.4	246.2	2.70	.32	.14	.6	.89	4.5			12.0	11.9	

9.7 APPENDIX G – LEARNER – ITEM STATISTICS ENGLISH PICTURE 1

TABLE 6.1 Baseline Picture Vocabulary Test ZOU592WS.TXT Aug 5 15:30 2010
INPUT: 1362 Learners 22 Objects MEASURED: 514 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 5.28 REL.: .97

"???????1" Learner STATISTICS: MISFIT ORDER ENG PVT1

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
591	2	3	3.12	1.57	2.62	1.8	1.78	.9	A .20	.59	33.3	81.6	23010611
768	2	3	3.12	1.57	2.62	1.8	1.78	.9	B .20	.59	33.3	81.6	3010801
565	3	4	3.13	1.57	2.61	1.8	1.33	.7	C .37	.63	50.0	86.1	23010331
603	3	4	3.13	1.57	2.61	1.8	1.33	.7	D .37	.63	50.0	86.1	23010741
703	3	4	3.13	1.57	2.61	1.8	1.33	.7	E .37	.63	50.0	86.1	3010131
841	3	4	3.13	1.57	2.61	1.8	1.33	.7	F .37	.63	50.0	86.1	16010871
1319	3	4	3.13	1.57	2.61	1.8	1.33	.7	G .37	.63	50.0	86.1	17010031
30	4	5	3.14	1.56	2.59	1.8	1.07	.5	H .42	.65	60.0	88.8	4010301
35	4	5	3.14	1.56	2.59	1.8	1.07	.5	I .42	.65	60.0	88.8	4010361
54	4	5	3.14	1.56	2.59	1.8	1.07	.5	J .42	.65	60.0	88.8	4010551
538	4	5	3.14	1.56	2.59	1.8	1.07	.5	K .42	.65	60.0	88.8	23010011
547	4	5	3.14	1.56	2.59	1.8	1.07	.5	L .42	.65	60.0	88.8	23010121
557	4	5	3.14	1.56	2.59	1.8	1.07	.5	M .42	.65	60.0	88.8	23010241
558	4	5	3.14	1.56	2.59	1.8	1.07	.5	N .42	.65	60.0	88.8	23010251
561	4	5	3.14	1.56	2.59	1.8	1.07	.5	O .42	.65	60.0	88.8	23010281
563	4	5	3.14	1.56	2.59	1.8	1.07	.5	P .42	.65	60.0	88.8	23010311
595	4	5	3.14	1.56	2.59	1.8	1.07	.5	Q .42	.65	60.0	88.8	23010651
713	4	5	3.14	1.56	2.59	1.8	1.07	.5	R .42	.65	60.0	88.8	3010231
716	4	5	3.14	1.56	2.59	1.8	1.07	.5	S .42	.65	60.0	88.8	3010261
721	4	5	3.14	1.56	2.59	1.8	1.07	.5	T .42	.65	60.0	88.8	3010311
737	4	5	3.14	1.56	2.59	1.8	1.07	.5	U .42	.65	60.0	88.8	3010471
738	4	5	3.14	1.56	2.59	1.8	1.07	.5	V .42	.65	60.0	88.8	3010481
739	4	5	3.14	1.56	2.59	1.8	1.07	.5	W .42	.65	60.0	88.8	3010491
744	4	5	3.14	1.56	2.59	1.8	1.07	.5	X .42	.65	60.0	88.8	3010541

749	4	5	3.14	1.56	2.59	1.8	1.07	.5	Y .42	.65	60.0	88.8	3010591
758	4	5	3.14	1.56	2.59	1.8	1.07	.5	Z .42	.65	60.0	88.8	3010691
728	6	7	3.68	1.33	1.91	1.3	.93	.4	z .33	.50	71.4	88.5	3010381
731	6	7	3.68	1.33	1.91	1.3	.93	.4	y .33	.50	71.4	88.5	3010411
733	6	7	3.68	1.33	1.91	1.3	.93	.4	x .33	.50	71.4	88.5	3010431
735	6	7	3.68	1.33	1.91	1.3	.93	.4	w .33	.50	71.4	88.5	3010451
751	6	7	3.68	1.33	1.91	1.3	.93	.4	v .33	.50	71.4	88.5	3010611
767	6	7	3.68	1.33	1.91	1.3	.93	.4	u .33	.50	71.4	88.5	3010791
802	6	7	3.68	1.33	1.91	1.3	.93	.4	t .33	.50	71.4	88.5	16010401
1320	6	7	3.68	1.33	1.91	1.3	.93	.4	s .33	.50	71.4	88.5	17010041
1336	6	7	3.68	1.33	1.91	1.3	.93	.4	r .33	.50	71.4	88.5	17010211
1360	6	7	3.68	1.33	1.91	1.3	.93	.4	q .33	.50	71.4	88.5	17010541
353	6	7	3.68	1.33	1.79	1.2	.79	.3	p .36	.50	71.4	88.5	18010391
362	6	7	3.68	1.33	1.79	1.2	.79	.3	o .36	.50	71.4	88.5	18010481
773	6	7	3.68	1.33	1.79	1.2	.79	.3	n .36	.50	71.4	88.5	16010051
775	6	7	3.68	1.33	1.79	1.2	.79	.3	m .36	.50	71.4	88.5	16010081
803	6	7	3.68	1.33	1.79	1.2	.79	.3	l .36	.50	71.4	88.5	16010411
809	6	7	3.68	1.33	1.79	1.2	.79	.3	k .36	.50	71.4	88.5	16010471
817	6	7	3.68	1.33	1.79	1.2	.79	.3	j .36	.50	71.4	88.5	16010551
831	6	7	3.68	1.33	1.79	1.2	.79	.3	i .36	.50	71.4	88.5	16010741
842	6	7	3.68	1.33	1.79	1.2	.79	.3	h .36	.50	71.4	88.5	16010881
696	5	7	2.16	1.20	1.77	1.4	1.26	.6	g .53	.69	71.4	83.4	3010061
BETTER FITTING OMITTED													
361	5	7	2.16	1.20	.68	-.5	.29	-.2	f .76	.69	100.0	83.4	18010481
592	6	7	3.68	1.33	.41	-.9	.17	-.5	e .62	.50	100.0	88.5	23010621
537	1	3	.03	2.11	.13	-.6	.09	-.7	d .95	.79	100.0	91.6	23010021
580	1	3	.03	2.11	.13	-.6	.09	-.7	c .95	.79	100.0	91.6	23010501
614	3	5	.37	1.81	.11	-.6	.08	-.7	b .96	.85	100.0	93.2	32010291
598	4	5	1.97	2.24	.09	-.7	.05	-.9	a .93	.82	100.0	95.7	23010691
MEAN	5.8	6.0	4.90	1.93	2.17	1.4	.93	.4			67.3	89.0	
S.D.	1.3	1.2	.86	.22	.59	.6	.27	.3			11.2	1.9	

9.8 APPENDIX H –ITEM STATISTICS ENGLISH PICTURE 1

TABLE 10.1 Baseline Picture Vocabulary Test ZOU760WS.TXT Nov 5 13:04 2009
INPUT: 1362 Learners 22 Objects MEASURED: 514 Learners 7 Objects 2 CATS

Learner: REAL SEP.: 1.74 REL.: .75 ... Object: REAL SEP.: 8.93 REL.: .99

Object STATISTICS: MISFIT ORDER English PICTURE 1 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
5	560	617	-.59	.15	1.32	2.8	4.62	8.4	A .04	.34	90.0	90.9	Cherries
3	1156	1159	-5.01	.59	1.03	.2	2.32	3.3	B .05	.08	99.7	99.7	Fork
4	925	935	-3.35	.33	1.07	.3	1.82	1.8	C .10	.16	98.9	98.9	Cupboard
6	956	1310	.64	.07	1.12	3.2	1.06	.8	D .46	.51	73.2	78.8	Pan
2	1222	1223	-6.31	1.01	1.01	.3	.11	-7.5	c .06	.05	99.9	99.9	Knife
7	1165	1311	-.84	.10	.91	-1.4	.92	-.5	b .45	.41	90.7	89.7	Bowl
1	1255	1264	-4.28	.35	.88	-.3	.48	-1.9	a .19	.15	99.3	99.3	Carrots
MEAN	1034.1	1117.0	-2.82	.37	1.05	.7	1.62	.7			93.1	93.9	
S.D.	226.4	236.8	2.40	.31	.14	1.5	1.41	4.5			9.0	7.4	

9.9 APPENDIX I –ITEM ORDER PICTURE 1 ENGLISH

TABLE 13.1 Baseline Picture Vocabulary Test ZOU760WS.TXT Nov 5 13:04 2009
INPUT: 1362 Learners 22 Objects MEASURED: 514 Learners 7 Objects 2 CATS

Learner: REAL SEP.: 1.74 REL.: .75 ... Object: REAL SEP.: 8.93 REL.: .99

Object STATISTICS: MEASURE ORDER English PVT1 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
6	956	1310	.64	.07	1.12	3.2	1.06	.8	.46	.51	73.2	78.8	Pan
5	560	617	-.59	.15	1.32	2.8	4.62	8.4	.04	.34	90.0	90.9	Cherries
7	1165	1311	-.84	.10	.91	-1.4	.92	-.5	.45	.41	90.7	89.7	Bowl
4	925	935	-3.35	.33	1.07	.3	1.82	1.8	.10	.16	98.9	98.9	Cupboard
1	1255	1264	-4.28	.35	.88	-.3	.48	-1.9	.19	.15	99.3	99.3	Carrots
3	1156	1159	-5.01	.59	1.03	.2	2.32	3.3	.05	.08	99.7	99.7	Fork
2	1222	1223	-6.31	1.01	1.01	.3	.11	-7.5	.06	.05	99.9	99.9	Knife
MEAN	1034.1	1117.0	-2.82	.37	1.05	.7	1.62	.7			93.1	93.9	
S.D.	226.4	236.8	2.40	.31	.14	1.5	1.41	4.5			9.0	7.4	

9.10 APPENDIX J – LEARNER – ITEM STATISTICS SEPEDI PICTURE 1

TABLE 6.1 Baseline Picture Vocabulary Test ZOU204WS.TXT Aug 5 15:34 2010
INPUT: 1362 Learners 22 Objects MEASURED: 404 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 4.85 REL.: .96

"????????3" Learner STATISTICS: MISFIT ORDER SEPEDI PVT1

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
989	2	3	3.17	1.50	2.05	1.7	1.37	.7	A .34	.54	33.3	78.1	13010653
1218	4	5	3.17	1.50	2.04	1.7	1.03	.5	B .50	.64	50.0	83.6	11010383
110	5	6	3.20	1.48	1.99	1.6	.82	.3	C .49	.62	60.0	86.6	5010293
985	5	6	3.20	1.48	1.99	1.6	.82	.3	D .49	.62	60.0	86.6	13010613
1027	5	6	3.20	1.48	1.99	1.6	.82	.3	E .49	.62	60.0	86.6	14010273
953	6	7	3.67	1.29	1.63	1.1	.88	.4	F .38	.48	66.7	85.1	13010233
1185	5	6	2.46	1.38	.91	-.2	.39	-.1	G .59	.57	80.0	80.0	22010503
966	5	7	2.17	1.23	.84	-.2	.39	-.1	H .73	.70	83.3	82.1	13010373
972	5	7	2.17	1.23	.84	-.2	.39	-.1	I .73	.70	83.3	82.1	13010453
976	5	7	2.17	1.23	.84	-.2	.39	-.1	J .73	.70	83.3	82.1	13010493
1144	1	3	2.11	1.28	.83	-.3	.72	-.2	K .54	.28	66.7	66.7	22010013
214	5	7	2.17	1.23	.77	-.3	.36	-.1	L .74	.70	83.3	82.1	7010483
619	5	7	2.17	1.23	.77	-.3	.36	-.1	M .74	.70	83.3	82.1	24010053
620	5	7	2.17	1.23	.77	-.3	.36	-.1	N .74	.70	83.3	82.1	24010063
623	5	7	2.17	1.23	.77	-.3	.36	-.1	O .74	.70	83.3	82.1	24010093
673	5	7	2.17	1.23	.77	-.3	.36	-.1	P .74	.70	83.3	82.1	24010653
1015	5	7	2.17	1.23	.77	-.3	.36	-.1	Q .74	.70	83.3	82.1	14010143
1022	5	7	2.17	1.23	.77	-.3	.36	-.1	R .74	.70	83.3	82.1	14010223
1031	5	7	2.17	1.23	.77	-.3	.36	-.1	S .74	.70	83.3	82.1	14010333
1045	5	7	2.17	1.23	.77	-.3	.36	-.1	T .74	.70	83.3	82.1	14010503
1061	5	7	2.17	1.23	.77	-.3	.36	-.1	U .74	.70	83.3	82.1	14010663
83	6	7	3.67	1.29	.50	-.9	.23	-.3	V .57	.48	100.0	85.1	5010013
91	6	7	3.67	1.29	.50	-.9	.23	-.3	W .57	.48	100.0	85.1	5010093
93	6	7	3.67	1.29	.50	-.9	.23	-.3	X .57	.48	100.0	85.1	5010113

102	6	7	3.67	1.29	.50	-.9	.23	-.3	Y .57	.48	100.0	85.1	5010213
108	6	7	3.67	1.29	.50	-.9	.23	-.3	Z .57	.48	100.0	85.1	5010273
BETTER FITTING OMITTED													
1228	5	6	3.20	1.48	.47	-1.0	.20	-.4	z .69	.62	100.0	86.6	11010523
1229	5	6	3.20	1.48	.47	-1.0	.20	-.4	y .69	.62	100.0	86.6	11010533
1232	5	6	3.20	1.48	.47	-1.0	.20	-.4	x .69	.62	100.0	86.6	11010593
1234	5	6	3.20	1.48	.47	-1.0	.20	-.4	w .69	.62	100.0	86.6	11010613
1240	5	6	3.20	1.48	.47	-1.0	.20	-.4	v .69	.62	100.0	86.6	11010663
1242	5	6	3.20	1.48	.47	-1.0	.20	-.4	u .69	.62	100.0	86.6	11010693
1245	5	6	3.20	1.48	.47	-1.0	.20	-.4	t .69	.62	100.0	86.6	11010723
1250	5	6	3.20	1.48	.47	-1.0	.20	-.4	s .69	.62	100.0	86.6	11010783
116	4	6	.57	1.84	.13	-.6	.08	-.7	r .94	.86	100.0	93.3	5010343
202	4	6	.57	1.84	.13	-.6	.08	-.7	q .94	.86	100.0	93.3	7010363
212	4	6	.57	1.84	.13	-.6	.08	-.7	p .94	.86	100.0	93.3	7010463
227	4	6	.57	1.84	.13	-.6	.08	-.7	o .94	.86	100.0	93.3	7010633
240	4	6	.57	1.84	.13	-.6	.08	-.7	n .94	.86	100.0	93.3	7010773
664	4	6	.57	1.84	.13	-.6	.08	-.7	m .94	.86	100.0	93.3	24010533
954	4	6	.57	1.84	.13	-.6	.08	-.7	l .94	.86	100.0	93.3	13010243
1060	4	6	.57	1.84	.13	-.6	.08	-.7	k .94	.86	100.0	93.3	14010653
112	3	5	-.38	2.74	.04	-.9	.04	-.9	j .98	.93	100.0	96.5	5010303
651	3	5	-.38	2.74	.04	-.9	.04	-.9	i .98	.93	100.0	96.5	24010403
1035	3	5	-.38	2.74	.04	-.9	.04	-.9	h .98	.93	100.0	96.5	14010373
1066	3	5	-.38	2.74	.04	-.9	.04	-.9	g .98	.93	100.0	96.5	14010713
1216	3	5	-.38	2.74	.04	-.9	.04	-.9	f .98	.93	100.0	96.5	11010343
175	2	4	-.87	3.46	.04	-.9	.03	-1.0	e .99	.95	100.0	97.1	7010073
206	1	3	-.87	3.46	.04	-.9	.03	-1.0	d .99	.92	100.0	97.1	7010403
951	1	3	-.87	3.46	.04	-.9	.03	-1.0	c .99	.92	100.0	97.1	13010213
975	1	3	-.87	3.46	.04	-.9	.03	-1.0	b .99	.92	100.0	97.1	13010483
1003	2	4	-.87	3.46	.04	-.9	.03	-1.0	a .99	.95	100.0	97.1	14010023
MEAN	4.8	5.7	3.53	1.57	.51	-.9	.25	-.3			98.3	84.5	
S.D.	1.3	1.3	1.12	.36	.23	.4	.12	.2			7.3	4.0	

9.11 APPENDIX K – OBJECT STATISTICS SEPEDI PICTURE 1

TABLE 10.1 Baseline Picture Vocabulary Test ZOU778WS.TXT Oct 28 15:06 2009
INPUT: 1362 Learners BASELINE 22 Objects MEASURED: 404 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 4.85 REL.: .96

Object STATISTICS: MISFIT ORDER SEPEDI PVT1

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
6	988	1318	3.89	.10	1.04	1.5	1.29	3.2	A .69	.71	64.7	64.7	Pan
5	567	623	2.36	.16	1.04	.5	1.04	.4	B .45	.47	74.7	76.4	Cherries
4	943	948	-1.25	.46	.98	.1	.39	-1.3	C .20	.16	98.5	98.5	Cupboard
3	1181	1182	-3.48	1.01	.98	.3	.09	-3.2	c .12	.08	99.8	99.8	Fork
7	1152	1319	2.45	.10	.95	-1.1	.92	-1.4	b .58	.56	73.1	72.1	Bowl
1	1273	1274	-3.96	1.01	.95	.3	.05	-5.0	a .13	.09	99.8	99.8	Carrots
MEAN	1049.7	1129.7	-.69	.66	.99	.3	.63	-1.2			85.1	85.2	
S.D.	227.9	238.3	3.30	.60	.04	.7	.48	2.6			14.6	14.6	

9.12 APPENDIX L – OBJECT ORDER SEPEDI PICTURE 1

TABLE 13.1 Baseline Picture Vocabulary Test ZOU778WS.TXT Oct 28 15:06 2009
INPUT: 1362 Learners 22 Objects MEASURED: 404 Learners 7 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 4.85 REL.: .96

Object STATISTICS: MEASURE ORDER SEPEDI PVT1 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
6	988	1318	3.89	.10	1.04	1.5	1.29	3.2	.69	.71	64.7	64.7	Pan
7	1152	1319	2.45	.10	.95	-1.1	.92	-1.4	.58	.56	73.1	72.1	Bowl
5	567	623	2.36	.16	1.04	.5	1.04	.4	.45	.47	74.7	76.4	Cherries
4	943	948	-1.25	.46	.98	.1	.39	-1.3	.20	.16	98.5	98.5	Cupboard
3	1181	1182	-3.48	1.01	.98	.3	.09	-3.2	.12	.08	99.8	99.8	Fork
1	1273	1274	-3.96	1.01	.95	.3	.05	-5.0	.13	.09	99.8	99.8	Carrots
2	1244	1244	-4.86	1.82	MINIMUM MEASURE				.00	.00	100.0	100.0	Knife
MEAN	1049.7	1129.7	-.69	.66	.99	.3	.63	-1.2			85.1	85.2	
S.D.	227.9	238.3	3.30	.60	.04	.7	.48	2.6			14.6	14.6	

9.13 APPENDIX M– PICTURE 2 ITEMS

- Show the picture on the opposite page to the child.
- Ask him/her to point to the **butterfly** and wait for a response.
- Continue this procedure, asking for:

  the kite,
  the castle,
  the wasp,
  the pigeon, (the robin is wrong)
  the windmill,
  the turtle,
  the violin,
  the padlock,
  the toadstool.

- On the Pupil Record Sheet put a line through the bracket underneath each picture correctly identified.
- If the score was less than 7 go straight to **Ideas about Reading** otherwise continue.

9.14 APPENDIX N – LEARNER – ITEM STATISTICS ALL LANGUAGES PIC 2

TABLE 6.1 Baseline Picture Vocabulary Test ZOU222WS.TXT Aug 5 11:56 2010
INPUT: 1362 Learners 22 Objects MEASURED: 1221 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.06 REL.: .53 ... Object: REAL SEP.: 14.31 REL.: 1.00

Learner STATISTICS: MISFIT ORDER ALL 3 LANGUAGES PVT 2

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
561	5	10	.23	.69	1.32	1.6	8.61	3.8	A-.12	.38	60.0	65.7	23010281
3	5	10	.23	.69	1.27	1.4	8.57	3.8	B-.09	.38	60.0	65.7	4010011
138	1	5	-2.76	1.69	4.51	1.9	6.67	2.3	C-.35	.74	60.0	91.9	10010122
31	4	10	-.26	.72	1.74	2.3	5.91	3.6	D-.23	.45	50.0	70.7	4010321
739	4	10	-.26	.72	1.55	1.8	5.79	3.5	E-.12	.45	70.0	70.7	3010491
1346	4	10	-.26	.72	1.53	1.7	5.76	3.5	F-.11	.45	70.0	70.7	17010351
722	4	10	-.26	.72	1.43	1.5	5.60	3.4	G-.04	.45	70.0	70.7	3010321
213	1	5	-2.76	1.69	4.20	1.9	2.93	1.4	H-.06	.74	60.0	91.9	7010473
767	1	5	-2.76	1.69	4.20	1.9	2.93	1.4	I-.06	.74	60.0	91.9	3010791
1025	1	5	-2.76	1.69	4.20	1.9	2.93	1.4	J-.06	.74	60.0	91.9	14010253
1060	1	5	-2.76	1.69	4.20	1.9	2.93	1.4	K-.06	.74	60.0	91.9	14010653
747	3	10	-.84	.81	1.63	1.3	3.49	2.5	L .05	.52	60.0	79.3	3010571
547	3	10	-.84	.81	1.58	1.3	3.42	2.5	M .07	.52	60.0	79.3	23010121
721	1	10	-3.14	1.48	3.11	1.6	1.92	1.0	N .08	.68	80.0	94.6	3010311
768	1	10	-3.14	1.48	3.11	1.6	1.92	1.0	O .08	.68	80.0	94.6	3010801
689	1	10	-3.14	1.48	3.10	1.6	1.85	1.0	P .09	.68	80.0	94.6	3010681
728	2	10	-1.64	1.01	2.42	1.6	2.62	1.6	Q-.05	.61	70.0	88.5	3010381
733	2	10	-1.64	1.01	2.40	1.6	2.54	1.5	R-.03	.61	70.0	88.5	3010431
714	2	10	-1.64	1.01	2.39	1.6	2.46	1.5	S-.02	.61	70.0	88.5	3010241
744	2	10	-1.64	1.01	2.39	1.6	2.46	1.5	T-.02	.61	70.0	88.5	3010541
764	2	10	-1.64	1.01	2.39	1.6	2.46	1.5	U-.02	.61	70.0	88.5	3010751
1241	2	10	-1.64	1.01	2.33	1.6	2.30	1.4	V .03	.61	70.0	88.5	11010683
78	9	10	2.64	1.07	1.19	.5	2.19	1.1	W-.09	.15	90.0	90.0	4010811
370	9	10	2.64	1.07	1.19	.5	2.19	1.1	X-.09	.15	90.0	90.0	18010561

380	9	10	2.64	1.07	1.19	.5	2.19	1.1	Y	-.09	.15	90.0	90.0	18010681
793	9	10	2.64	1.07	1.19	.5	2.19	1.1	Z	-.09	.15	90.0	90.0	16010311
780	6	10	.71	.69	1.48	2.4	1.40	.7		.03	.33	30.0	64.5	16010151
	BETTER FITTING OMITTED													
615	1	5	-2.76	1.69	.15	-.7	.09	-.7	z	.98	.74	100.0	91.9	24010013
629	1	5	-2.76	1.69	.15	-.7	.09	-.7	y	.98	.74	100.0	91.9	24010163
679	1	5	-2.76	1.69	.15	-.7	.09	-.7	x	.98	.74	100.0	91.9	24010723
703	1	5	-2.76	1.69	.15	-.7	.09	-.7	w	.98	.74	100.0	91.9	3010131
731	1	5	-2.76	1.69	.15	-.7	.09	-.7	v	.98	.74	100.0	91.9	3010411
941	1	5	-2.76	1.69	.15	-.7	.09	-.7	u	.98	.74	100.0	91.9	13010093
950	1	5	-2.76	1.69	.15	-.7	.09	-.7	t	.98	.74	100.0	91.9	13010203
963	1	5	-2.76	1.69	.15	-.7	.09	-.7	s	.98	.74	100.0	91.9	13010333
970	1	5	-2.76	1.69	.15	-.7	.09	-.7	r	.98	.74	100.0	91.9	13010433
995	1	5	-2.76	1.69	.15	-.7	.09	-.7	q	.98	.74	100.0	91.9	13010733
1003	1	5	-2.76	1.69	.15	-.7	.09	-.7	p	.98	.74	100.0	91.9	14010023
1021	1	5	-2.76	1.69	.15	-.7	.09	-.7	o	.98	.74	100.0	91.9	14010213
1035	1	5	-2.76	1.69	.15	-.7	.09	-.7	n	.98	.74	100.0	91.9	14010373
1036	1	5	-2.76	1.69	.15	-.7	.09	-.7	m	.98	.74	100.0	91.9	14010383
1054	1	5	-2.76	1.69	.15	-.7	.09	-.7	l	.98	.74	100.0	91.9	14010593
1061	1	5	-2.76	1.69	.15	-.7	.09	-.7	k	.98	.74	100.0	91.9	14010663
1195	1	5	-2.76	1.69	.15	-.7	.09	-.7	j	.98	.74	100.0	91.9	11010093
1210	1	5	-2.76	1.69	.15	-.7	.09	-.7	i	.98	.74	100.0	91.9	11010263
1219	1	5	-2.76	1.69	.15	-.7	.09	-.7	h	.98	.74	100.0	91.9	11010393
1228	1	5	-2.76	1.69	.15	-.7	.09	-.7	g	.98	.74	100.0	91.9	11010523
1236	1	5	-2.76	1.69	.15	-.7	.09	-.7	f	.98	.74	100.0	91.9	11010623
1251	1	5	-2.76	1.69	.15	-.7	.09	-.7	e	.98	.74	100.0	91.9	11010793
1330	1	5	-2.76	1.69	.15	-.7	.09	-.7	d	.98	.74	100.0	91.9	17010141
1240	1	4	-2.65	1.76	.15	-.6	.11	-.6	c	.98	.75	100.0	90.7	11010663
590	1	4	-2.65	1.76	.15	-.6	.11	-.6	b	.98	.75	100.0	90.8	23010601
580	1	3	-2.24	2.06	.11	-.6	.10	-.7	a	1.00	.80	100.0	91.3	23010501
MEAN	6.4	9.6	1.04	.99	.99	.1	.92	.3				75.9	75.7	
S.D.	2.6	1.3	1.82	.42	.36	.6	.61	.4				12.8	9.6	

9.15 APPENDIX O – ITEM STATISTICS ALL LANGUAGES PIC 2

TABLE 10.1 Baseline Picture Vocabulary Test ZOU222WS.TXT Aug 5 11:56 2010
INPUT: 1362 Learners 22 Objects MEASURED: 1221 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.06 REL.: .53 ... Object: REAL SEP.: 14.31 REL.: 1.00

Object STATISTICS: MISFIT ORDER ALL 3 LANGUAGES PVT 2

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
10	717	1218	.62	.07	1.18	5.6	1.26	4.4	A .48	.56	67.3	71.5	Castle
16	635	1134	1.02	.07	1.17	5.8	1.17	3.7	B .44	.53	61.1	69.4	Padlock
17	658	1131	.90	.07	1.09	3.2	1.08	2.0	C .48	.53	64.7	69.5	Toadstool
13	885	1134	-.35	.08	1.09	2.1	1.09	1.3	D .39	.43	75.7	77.9	Windmill
8	1175	1220	-4.13	.22	1.04	.3	.51	-1.6	E .47	.47	97.8	97.8	Butterfly
14	747	1133	.44	.07	1.03	.9	1.02	.3	e .48	.50	71.0	71.1	Tortoise
11	899	1218	-.40	.08	1.01	.3	.98	-.2	d .55	.55	78.6	78.9	Wasp
9	787	1219	.25	.07	.87	-4.3	.78	-4.2	c .62	.56	78.6	73.3	Kite
15	619	1133	1.10	.07	.81	-7.1	.76	-6.0	b .64	.54	77.3	69.3	Violin
12	731	1214	.54	.07	.80	-7.1	.71	-6.0	a .65	.56	78.8	72.0	Pigeon
MEAN	785.3	1175.4	.00	.09	1.01	.0	.94	-.6			75.1	75.1	
S.D.	158.2	42.4	1.46	.04	.13	4.5	.22	3.5			9.7	8.2	

9.16 APPENDIX P - LEARNER – ITEM STATISTICS AFRIKAANS PICTURE 2

TABLE 6.1 Baseline Picture Vocabulary Test ZOU138WS.TXT Aug 5 15:23 2010
INPUT: 1362 Learners 22 Objects MEASURED: 303 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.05 REL.: .53 ... Object: REAL SEP.: 13.14 REL.: .99

"????????" Learner STATISTICS: MISFIT ORDER AFR PVT2

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
138	1	5	-2.28	1.35	2.43	1.5	3.37	1.6	A-	.36	.56	60.0	85.9	10010122
420	9	10	2.50	1.07	1.18	.5	2.10	1.1	B-	.15	.16	90.0	90.0	8010352
908	9	10	2.50	1.07	1.18	.5	2.10	1.1	C-	.15	.16	90.0	90.0	6010612
878	8	10	1.65	.82	1.28	.7	1.55	.8	D-	.09	.22	80.0	80.0	6010312
437	8	10	1.65	.82	1.24	.7	1.46	.7	E-	.04	.22	80.0	80.0	8030512
513	8	10	1.65	.82	1.24	.7	1.46	.7	F-	.04	.22	80.0	80.0	12010742
408	8	10	1.65	.82	1.24	.7	1.44	.7	G-	.03	.22	80.0	80.0	8010232
443	8	10	1.65	.82	1.24	.7	1.44	.7	H-	.03	.22	80.0	80.0	8030572
526	8	10	1.65	.82	1.24	.7	1.44	.7	I-	.03	.22	80.0	80.0	12010942
256	8	10	1.65	.82	1.23	.6	1.44	.7	J-	.03	.22	80.0	80.0	19020152
151	8	10	1.65	.82	1.22	.6	1.43	.7	K-	.02	.22	80.0	80.0	10010262
242	8	10	1.65	.82	1.22	.6	1.43	.7	L-	.02	.22	80.0	80.0	19010022
253	8	10	1.65	.82	1.22	.6	1.43	.7	M-	.02	.22	80.0	80.0	19020122
436	8	10	1.65	.82	1.22	.6	1.43	.7	N-	.02	.22	80.0	80.0	8030502
440	8	10	1.65	.82	1.22	.6	1.43	.7	O-	.02	.22	80.0	80.0	8030542
858	8	10	1.65	.82	1.22	.6	1.43	.7	P-	.02	.22	80.0	80.0	6010102
884	8	10	1.65	.82	1.22	.6	1.43	.7	Q-	.02	.22	80.0	80.0	6010372
915	8	10	1.65	.82	1.22	.6	1.41	.7	R-	.02	.22	80.0	80.0	6010682
457	6	10	.58	.69	1.32	1.6	1.27	.6	S	.05	.33	50.0	64.4	8030732
422	8	10	1.65	.82	1.13	.4	1.31	.6	T	.06	.22	80.0	80.0	8010372
156	8	10	1.65	.82	1.10	.4	1.28	.6	U	.09	.22	80.0	80.0	10010322
146	7	10	1.07	.72	1.26	1.0	1.24	.6	V	.05	.28	70.0	70.0	10010212
520	7	10	1.07	.72	1.26	1.0	1.24	.6	W	.05	.28	70.0	70.0	12010882
506	7	10	1.07	.72	1.26	1.0	1.23	.5	X	.05	.28	70.0	70.0	12010612

241	7	10	1.07	.72	1.25	1.0	1.23	.5	Y .06	.28	70.0	70.0	19010012
450	7	10	1.07	.72	1.25	1.0	1.23	.5	Z .06	.28	70.0	70.0	8030662
246	8	10	1.65	.82	.93	.0	.76	.1	z .31	.22	80.0	80.0	19010062
	BETTER FITTING OMITTED												
502	7	10	1.07	.72	.93	-.2	.79	.0	y .36	.28	70.0	70.0	12010552
889	2	10	-1.62	.92	.85	.0	.93	.2	x .58	.50	90.0	85.6	6010422
158	6	10	.58	.69	.92	-.4	.82	-.1	w .41	.33	70.0	64.4	10010352
282	9	10	2.50	1.07	.92	.2	.60	.1	v .27	.16	90.0	90.0	19050412
316	9	10	2.50	1.07	.92	.2	.60	.1	u .27	.16	90.0	90.0	19080772
410	9	10	2.50	1.07	.92	.2	.60	.1	t .27	.16	90.0	90.0	8010252
412	9	10	2.50	1.07	.92	.2	.60	.1	s .27	.16	90.0	90.0	8010272
416	9	10	2.50	1.07	.92	.2	.60	.1	r .27	.16	90.0	90.0	8010312
433	9	10	2.50	1.07	.92	.2	.60	.1	q .27	.16	90.0	90.0	8010482
464	9	10	2.50	1.07	.92	.2	.60	.1	p .27	.16	90.0	90.0	8040742
873	5	10	.11	.68	.91	-.4	.83	-.2	o .46	.37	60.0	64.8	6010262
481	8	10	1.65	.82	.91	-.1	.73	.1	n .32	.22	80.0	80.0	12010312
495	8	10	1.65	.82	.91	-.1	.73	.1	m .32	.22	80.0	80.0	12010472
508	8	10	1.65	.82	.91	-.1	.73	.1	l .32	.22	80.0	80.0	12010642
525	8	10	1.65	.82	.91	-.1	.73	.1	k .32	.22	80.0	80.0	12010932
1347	7	10	1.07	.72	.87	-.4	.73	-.1	j .41	.28	70.0	70.0	17010362
524	6	10	.58	.69	.87	-.7	.77	-.2	i .45	.33	70.0	64.4	12010922
273	7	10	1.07	.72	.86	-.4	.72	-.1	h .41	.28	70.0	70.0	19040322
900	8	10	1.65	.82	.85	-.2	.64	.0	g .38	.22	80.0	80.0	6010532
903	3	10	-.92	.78	.84	-.3	.74	-.4	f .61	.46	80.0	77.5	6010562
141	7	10	1.07	.72	.83	-.6	.70	-.2	e .44	.28	70.0	70.0	10010162
496	7	10	1.07	.72	.83	-.6	.70	-.2	d .44	.28	70.0	70.0	12010482
862	6	10	.58	.69	.72	-1.7	.64	-.5	c .58	.33	90.0	64.4	6010152
154	5	10	.11	.68	.71	-1.6	.64	-.7	b .63	.37	80.0	64.8	10010302
479	5	10	.11	.68	.71	-1.6	.64	-.7	a .63	.37	80.0	64.8	12010282
MEAN	8.5	10.0	2.34	1.18	1.05	.3	.97	.3			80.1	80.6	
S.D.	1.4	.3	1.17	.47	.14	.4	.28	.3			10.6	8.9	

9.17 APPENDIX Q - ITEM STATISTICS AFRIKAANS PICTURE 2

TABLE 10.1 Baseline Picture Vocabulary Test ZOU898WS.TXT Oct 28 11:31 2009
INPUT: 1362 Learners 22 Objects MEASURED: 303 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.05 REL.: .53 ... Object: REAL SEP.: 13.14 REL.: .99

Object STATISTICS: MISFIT ORDER AFRIKAANS PVT2 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
10	769	1306	.41	.07	1.19	6.0	1.32	6.2	A .48	.57	67.3	72.0	Castle
16	666	1199	.87	.07	1.17	5.8	1.16	3.9	B .45	.53	61.2	69.5	Padlock
13	930	1199	-.49	.08	1.09	2.3	1.10	1.4	C .39	.44	75.0	77.5	Windmill
17	688	1196	.76	.07	1.09	3.2	1.08	2.1	D .48	.53	64.9	69.5	Toadstool
14	778	1198	.33	.07	1.02	.8	1.00	.1	E .49	.50	70.6	70.8	Tortoise
8	1213	1308	-2.89	.13	1.01	.2	.74	-1.1	e .45	.44	93.9	93.8	Butterfly
11	942	1306	-.48	.08	.97	-.7	.96	-.7	d .56	.55	78.5	77.7	Wasp
9	828	1307	.12	.07	.87	-4.4	.81	-4.0	c .62	.56	78.3	73.0	Kite
12	774	1302	.38	.07	.82	-6.3	.76	-5.5	b .65	.57	78.3	72.1	Pigeon
15	641	1198	.99	.07	.81	-7.6	.75	-6.8	a .65	.54	77.7	69.4	Violin
MEAN	822.9	1251.9	.00	.08	1.00	-.1	.97	-.4			74.6	74.5	
S.D.	161.6	53.9	1.07	.02	.13	4.5	.19	3.9			8.7	7.1	

9.18 APPENDIX R - ITEM ORDER AFRIKAANS PICTURE 2

TABLE 13.1 Baseline Picture Vocabulary Test ZOU898WS.TXT Oct 28 11:31 2009
INPUT: 1362 Learners 22 Objects MEASURED: 303 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.05 REL.: .53 ... Object: REAL SEP.: 13.14 REL.: .99

Object STATISTICS: MEASURE ORDER AFRIKAANS PVT2 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
15	641	1198	.99	.07	.81	-7.6	.75	-6.8	.65	.54	77.7	69.4	Violin
16	666	1199	.87	.07	1.17	5.8	1.16	3.9	.45	.53	61.2	69.5	Padlock
17	688	1196	.76	.07	1.09	3.2	1.08	2.1	.48	.53	64.9	69.5	Toadstool
10	769	1306	.41	.07	1.19	6.0	1.32	6.2	.48	.57	67.3	72.0	Castle
12	774	1302	.38	.07	.82	-6.3	.76	-5.5	.65	.57	78.3	72.1	Pigeon
14	778	1198	.33	.07	1.02	.8	1.00	.1	.49	.50	70.6	70.8	Tortoise
9	828	1307	.12	.07	.87	-4.4	.81	-4.0	.62	.56	78.3	73.0	Kite
11	942	1306	-.48	.08	.97	-.7	.96	-.7	.56	.55	78.5	77.7	Wasp
13	930	1199	-.49	.08	1.09	2.3	1.10	1.4	.39	.44	75.0	77.5	Windmill
8	1213	1308	-2.89	.13	1.01	.2	.74	-1.1	.45	.44	93.9	93.8	Butterfly
MEAN	822.9	1251.9	.00	.08	1.00	-.1	.97	-.4			74.6	74.5	
S.D.	161.6	53.9	1.07	.02	.13	4.5	.19	3.9			8.7	7.1	

9.19 APPENDIX S – LEARNER - ITEM STATISTICS PICTURE 2 ENGLISH

TABLE 6.1 Baseline Picture Vocabulary Test ZOU704WS.TXT Aug 5 15:31 2010
INPUT: 1362 Learners 22 Objects MEASURED: 514 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.06 REL.: .53 ... Object: REAL SEP.: 13.92 REL.: .99

"???????1" Learner STATISTICS: MISFIT ORDER ENG PVT2

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
561	5	10	.16	.69	1.31	1.6	4.49	3.4	A-.10	.38	50.0	65.1	23010281
3	5	10	.16	.69	1.25	1.3	4.44	3.3	B-.06	.38	70.0	65.1	4010011
31	4	10	-.32	.72	1.72	2.2	3.43	3.1	C-.24	.43	40.0	70.0	4010321
739	4	10	-.32	.72	1.50	1.7	3.27	2.9	D-.11	.43	60.0	70.0	3010491
1346	4	10	-.32	.72	1.50	1.6	3.27	2.9	E-.10	.43	60.0	70.0	17010351
722	4	10	-.32	.72	1.42	1.4	3.12	2.8	F-.03	.43	60.0	70.0	3010321
767	1	5	-2.50	1.48	2.86	1.6	2.05	1.0	G-.02	.65	60.0	88.9	3010791
721	1	10	-2.89	1.32	2.22	1.3	1.52	.8	H .09	.58	80.0	92.7	3010311
768	1	10	-2.89	1.32	2.22	1.3	1.52	.8	I .09	.58	80.0	92.7	3010801
78	9	10	2.55	1.07	1.19	.5	2.20	1.1	J-.13	.15	90.0	90.0	4010811
370	9	10	2.55	1.07	1.19	.5	2.20	1.1	K-.13	.15	90.0	90.0	18010561
380	9	10	2.55	1.07	1.19	.5	2.20	1.1	L-.13	.15	90.0	90.0	18010681
793	9	10	2.55	1.07	1.19	.5	2.20	1.1	M-.13	.15	90.0	90.0	16010311
1300	9	10	2.55	1.07	1.19	.5	2.20	1.1	N-.13	.15	90.0	90.0	20010511
689	1	10	-2.89	1.32	2.18	1.3	1.32	.7	O .13	.58	80.0	92.7	3010681
747	3	10	-.89	.79	1.55	1.2	2.13	1.7	P .07	.49	60.0	78.6	3010571
728	2	10	-1.64	.96	2.11	1.5	2.00	1.3	Q-.05	.56	70.0	87.2	3010381
733	2	10	-1.64	.96	2.09	1.4	1.87	1.1	R-.02	.56	70.0	87.2	3010431
714	2	10	-1.64	.96	2.08	1.4	1.87	1.1	S-.02	.56	70.0	87.2	3010241
744	2	10	-1.64	.96	2.08	1.4	1.87	1.1	T-.02	.56	70.0	87.2	3010541
764	2	10	-1.64	.96	2.08	1.4	1.87	1.1	U-.02	.56	70.0	87.2	3010751
547	3	10	-.89	.79	1.49	1.1	2.04	1.6	V .11	.49	60.0	78.6	23010121
602	9	10	2.55	1.07	1.17	.5	1.88	1.0	W-.09	.15	90.0	90.0	23010721
783	9	10	2.55	1.07	1.17	.5	1.88	1.0	X-.09	.15	90.0	90.0	16010181
784	9	10	2.55	1.07	1.17	.5	1.88	1.0	Y-.09	.15	90.0	90.0	16010191

804	9	10	2.55	1.07	1.17	.5	1.88	1.0	Z-.09	.15	90.0	90.0	16010421
780	6	10	.63	.69	1.43	2.3	1.36	.7	.01	.32	30.0	63.8	16010151
	BETTER FITTING OMITTED												
39	2	10	-1.64	.96	.71	-.2	.50	-.5	z .76	.56	90.0	87.2	4010411
544	2	10	-1.64	.96	.71	-.2	.50	-.5	y .76	.56	90.0	87.2	23010081
562	2	10	-1.64	.96	.71	-.2	.50	-.5	x .76	.56	90.0	87.2	23010301
1322	5	10	.16	.69	.71	-1.7	.63	-.5	w .59	.38	90.0	65.1	17010061
691	5	10	.16	.69	.70	-1.7	.63	-.5	v .59	.38	90.0	65.1	3010011
726	5	10	.16	.69	.70	-1.7	.63	-.5	u .59	.38	90.0	65.1	3010361
1336	5	10	.16	.69	.70	-1.7	.63	-.5	t .59	.38	90.0	65.1	17010211
601	4	10	-.32	.72	.66	-1.3	.58	-.8	s .68	.43	80.0	70.0	23010711
1309	4	10	-.32	.72	.66	-1.3	.58	-.8	r .68	.43	80.0	70.0	20010601
62	3	10	-.89	.79	.65	-.8	.50	-.9	q .75	.49	80.0	78.6	4010641
809	3	10	-.89	.79	.65	-.8	.50	-.9	p .75	.49	80.0	78.6	16010471
569	1	10	-2.89	1.32	.26	-.8	.10	-.7	o .93	.58	100.0	92.7	23010371
35	1	5	-2.50	1.48	.22	-.8	.14	-.5	n .97	.65	100.0	88.9	4010361
361	1	5	-2.50	1.48	.22	-.8	.14	-.5	m .97	.65	100.0	88.9	18010481
552	1	5	-2.50	1.48	.22	-.8	.14	-.5	l .97	.65	100.0	88.9	23010161
563	1	5	-2.50	1.48	.22	-.8	.14	-.5	k .97	.65	100.0	88.9	23010311
567	1	5	-2.50	1.48	.22	-.8	.14	-.5	j .97	.65	100.0	88.9	23010351
579	1	5	-2.50	1.48	.22	-.8	.14	-.5	i .97	.65	100.0	88.9	23010481
584	1	5	-2.50	1.48	.22	-.8	.14	-.5	h .97	.65	100.0	88.9	23010541
591	1	5	-2.50	1.48	.22	-.8	.14	-.5	g .97	.65	100.0	88.9	23010611
614	1	5	-2.50	1.48	.22	-.8	.14	-.5	f .97	.65	100.0	88.9	32010291
703	1	5	-2.50	1.48	.22	-.8	.14	-.5	e .97	.65	100.0	88.9	3010131
731	1	5	-2.50	1.48	.22	-.8	.14	-.5	d .97	.65	100.0	88.9	3010411
1330	1	5	-2.50	1.48	.22	-.8	.14	-.5	c .97	.65	100.0	88.9	17010141
590	1	4	-2.37	1.54	.21	-.8	.16	-.5	b .97	.66	100.0	87.3	23010601
580	1	3	-1.95	1.78	.16	-.6	.14	-.5	a1.00	.73	100.0	87.8	23010501
MEAN	6.3	9.7	.87	.94	.98	.0	.91	.2			75.7	75.0	
S.D.	2.5	1.2	1.69	.38	.27	.6	.47	.5			12.2	9.3	

9.20 APPENDIX T - ITEM STATISTICS ENGLISH PICTURE 2

TABLE 10.1 Baseline Picture Vocabulary Test ZOU286WS.TXT Nov 5 13:49 2009
INPUT: 1362 Learners 22 Objects MEASURED: 514 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.06 REL.: .53 ... Object: REAL SEP.: 13.92 REL.: .99

Object STATISTICS: MISFIT ORDER Eng PVT2 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
10	769	1310	.58	.07	1.19	6.3	1.29	5.4	A .48	.57	67.0	71.6	Castle
8	1247	1312	-3.45	.17	1.07	.6	1.21	.8	B .42	.44	96.3	96.3	Butterfly
13	935	1214	-.32	.08	1.10	2.7	1.20	2.9	C .38	.44	72.7	76.8	Windmill
16	706	1214	.86	.07	1.16	5.5	1.15	3.7	D .45	.53	61.6	69.7	Padlock
17	720	1211	.79	.07	1.08	2.8	1.06	1.5	E .49	.53	65.3	69.7	Toadstool
14	797	1213	.42	.07	1.02	.6	1.01	.2	e .49	.50	71.2	71.1	Tortoise
11	974	1310	-.49	.08	.99	-.1	.97	-.4	d .55	.55	79.2	79.1	Wasp
9	856	1311	.15	.07	.86	-4.7	.77	-4.7	c .63	.56	76.8	73.6	Kite
12	799	1306	.43	.07	.80	-7.1	.72	-6.2	b .65	.57	79.2	72.4	Pigeon
15	674	1213	1.01	.07	.80	-7.9	.74	-7.0	a .65	.54	78.1	69.5	Violin
MEAN	847.7	1261.4	.00	.08	1.01	-.2	1.01	-.4			74.7	75.0	
S.D.	161.5	48.4	1.24	.03	.14	4.7	.20	4.0			9.3	7.7	

9.21 APPENDIX U - ITEM ORDER ENGLISH PICTURE 2

TABLE 13.1 Baseline Picture Vocabulary Test ZOU286WS.TXT Nov 5 13:49 2009
INPUT: 1362 Learners 22 Objects MEASURED: 514 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.06 REL.: .53 ... Object: REAL SEP.: 13.92 REL.: .99

Object STATISTICS: MEASURE ORDER Eng PVT2 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
15	674	1213	1.01	.07	.80	-7.9	.74	-7.0	.65	.54	78.1	69.5	Violin
16	706	1214	.86	.07	1.16	5.5	1.15	3.7	.45	.53	61.6	69.7	Padlock
17	720	1211	.79	.07	1.08	2.8	1.06	1.5	.49	.53	65.3	69.7	Toadstool
10	769	1310	.58	.07	1.19	6.3	1.29	5.4	.48	.57	67.0	71.6	Castle
12	799	1306	.43	.07	.80	-7.1	.72	-6.2	.65	.57	79.2	72.4	Pigeon
14	797	1213	.42	.07	1.02	.6	1.01	.2	.49	.50	71.2	71.1	Tortoise
9	856	1311	.15	.07	.86	-4.7	.77	-4.7	.63	.56	76.8	73.6	Kite
13	935	1214	-.32	.08	1.10	2.7	1.20	2.9	.38	.44	72.7	76.8	Windmill
11	974	1310	-.49	.08	.99	-.1	.97	-.4	.55	.55	79.2	79.1	Wasp
8	1247	1312	-3.45	.17	1.07	.6	1.21	.8	.42	.44	96.3	96.3	Butterfly
MEAN	847.7	1261.4	.00	.08	1.01	-.2	1.01	-.4			74.7	75.0	
S.D.	161.5	48.4	1.24	.03	.14	4.7	.20	4.0			9.3	7.7	

9.22 APPENDIX V – LEARNER - ITEM STATISTICS PICTURE 2 SEPEDI

TABLE 6.1 Baseline Picture Vocabulary Test ZOU796WS.TXT Aug 5 15:36 2010
INPUT: 1362 Learners 22 Objects MEASURED: 404 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.06 REL.: .53 ... Object: REAL SEP.: 13.32 REL.: .99

"????????3" Learner STATISTICS: MISFIT ORDER SEPEDI PVT2

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
213	1	5	-2.34	1.39	2.37	1.4	1.79	.9	A-.02	.59	60.0	86.9	7010473
1025	1	5	-2.34	1.39	2.37	1.4	1.79	.9	B-.02	.59	60.0	86.9	14010253
1060	1	5	-2.34	1.39	2.37	1.4	1.79	.9	C-.02	.59	60.0	86.9	14010653
86	9	10	2.51	1.07	1.18	.5	2.01	1.0	D-.13	.15	90.0	90.0	5010043
1241	2	10	-1.63	.93	1.79	1.2	1.41	.7	E .07	.52	70.0	86.2	11010683
687	3	10	-.91	.78	1.22	.6	1.43	.9	F .27	.47	80.0	77.9	24016193
959	3	10	-.91	.78	1.22	.6	1.43	.9	G .27	.47	80.0	77.9	13010293
120	4	10	-.36	.71	1.34	1.2	1.32	.8	H .16	.42	50.0	69.6	5010383
1226	4	10	-.36	.71	1.34	1.2	1.32	.8	I .16	.42	50.0	69.6	11010503
635	5	10	.13	.68	1.31	1.6	1.22	.6	J .14	.37	50.0	64.8	24010233
994	4	10	-.36	.71	1.30	1.1	1.27	.7	K .19	.42	50.0	69.6	13010723
662	3	10	-.91	.78	1.17	.5	1.29	.7	L .32	.47	80.0	77.9	24010513
933	2	10	-1.63	.93	.92	.1	1.27	.6	M .51	.52	90.0	86.2	13010143
1178	6	10	.59	.69	1.27	1.5	1.22	.5	N .10	.32	50.0	64.2	22010413
677	4	10	-.36	.71	1.25	.9	1.20	.6	O .23	.42	50.0	69.6	24010703
233	4	10	-.36	.71	1.25	.9	1.22	.6	P .23	.42	50.0	69.6	7010703
1028	8	10	1.66	.81	1.08	.3	1.23	.6	Q .11	.22	80.0	80.0	14010293
1052	8	10	1.66	.81	1.08	.3	1.23	.6	R .11	.22	80.0	80.0	14010573
1163	8	10	1.66	.81	1.08	.3	1.23	.6	S .11	.22	80.0	80.0	22010263
1200	8	10	1.66	.81	1.08	.3	1.23	.6	T .11	.22	80.0	80.0	11010143
999	5	10	.13	.68	1.22	1.2	1.13	.4	U .21	.37	50.0	64.8	13010783
955	3	10	-.91	.78	1.15	.5	1.21	.6	V .35	.47	80.0	77.9	13010253
87	6	10	.59	.69	1.20	1.1	1.11	.4	W .17	.32	50.0	64.2	5010053
618	6	10	.59	.69	1.20	1.1	1.11	.4	X .17	.32	50.0	64.2	24010043
657	6	10	.59	.69	1.20	1.1	1.11	.4	Y .17	.32	50.0	64.2	24010463

1047	6	10	.59	.69	1.20	1.1	1.11	.4	Z .17	.32	50.0	64.2	14010523
	BETTER	FITTING	OMITTED										
200	1	5	-2.34	1.39	.26	-.9	.17	-.6	z .97	.59	100.0	86.9	7010343
202	1	5	-2.34	1.39	.26	-.9	.17	-.6	y .97	.59	100.0	86.9	7010363
218	1	5	-2.34	1.39	.26	-.9	.17	-.6	x .97	.59	100.0	86.9	7010523
222	1	5	-2.34	1.39	.26	-.9	.17	-.6	w .97	.59	100.0	86.9	7010583
226	1	5	-2.34	1.39	.26	-.9	.17	-.6	v .97	.59	100.0	86.9	7010623
615	1	5	-2.34	1.39	.26	-.9	.17	-.6	u .97	.59	100.0	86.9	24010013
629	1	5	-2.34	1.39	.26	-.9	.17	-.6	t .97	.59	100.0	86.9	24010163
679	1	5	-2.34	1.39	.26	-.9	.17	-.6	s .97	.59	100.0	86.9	24010723
941	1	5	-2.34	1.39	.26	-.9	.17	-.6	r .97	.59	100.0	86.9	13010093
950	1	5	-2.34	1.39	.26	-.9	.17	-.6	q .97	.59	100.0	86.9	13010203
963	1	5	-2.34	1.39	.26	-.9	.17	-.6	p .97	.59	100.0	86.9	13010333
970	1	5	-2.34	1.39	.26	-.9	.17	-.6	o .97	.59	100.0	86.9	13010433
995	1	5	-2.34	1.39	.26	-.9	.17	-.6	n .97	.59	100.0	86.9	13010733
1003	1	5	-2.34	1.39	.26	-.9	.17	-.6	m .97	.59	100.0	86.9	14010023
1021	1	5	-2.34	1.39	.26	-.9	.17	-.6	l .97	.59	100.0	86.9	14010213
1035	1	5	-2.34	1.39	.26	-.9	.17	-.6	k .97	.59	100.0	86.9	14010373
1036	1	5	-2.34	1.39	.26	-.9	.17	-.6	j .97	.59	100.0	86.9	14010383
1054	1	5	-2.34	1.39	.26	-.9	.17	-.6	i .97	.59	100.0	86.9	14010593
1061	1	5	-2.34	1.39	.26	-.9	.17	-.6	h .97	.59	100.0	86.9	14010663
1195	1	5	-2.34	1.39	.26	-.9	.17	-.6	g .97	.59	100.0	86.9	11010093
1210	1	5	-2.34	1.39	.26	-.9	.17	-.6	f .97	.59	100.0	86.9	11010263
1219	1	5	-2.34	1.39	.26	-.9	.17	-.6	e .97	.59	100.0	86.9	11010393
1228	1	5	-2.34	1.39	.26	-.9	.17	-.6	d .97	.59	100.0	86.9	11010523
1236	1	5	-2.34	1.39	.26	-.9	.17	-.6	c .97	.59	100.0	86.9	11010623
1251	1	5	-2.34	1.39	.26	-.9	.17	-.6	b .97	.59	100.0	86.9	11010793
1240	1	4	-2.22	1.45	.26	-.9	.19	-.6	a .98	.60	100.0	84.9	11010663
MEAN	5.0	9.3	.09	.86	.93	.0	.84	.0			71.9	72.1	
S.D.	2.3	1.7	1.44	.32	.29	.6	.30	.4			15.2	8.5	

9.23 APPENDIX W – ITEM STATISTICS PICTURE 2 SEPEDI

TABLE 10.1 Baseline Picture Vocabulary Test ZOU972WS.TXT Oct 28 15:10 2009
INPUT: 1362 Learners 22 Objects MEASURED: 404 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.06 REL.: .53 ... Object: REAL SEP.: 13.32 REL.: .99

Object STATISTICS: MISFIT ORDER SEPEDI PVT2 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
10	785	1318	.50	.07	1.21	6.8	1.34	6.5	A .47	.57	66.7	72.0	Castle
13	954	1223	-.43	.08	1.14	3.3	1.26	3.5	B .36	.44	75.0	77.9	Windmill
16	693	1223	.92	.07	1.13	4.6	1.12	2.8	C .47	.54	63.0	69.8	Padlock
17	720	1220	.78	.07	1.07	2.5	1.06	1.5	D .50	.53	66.0	69.9	Toadstool
14	816	1222	.32	.07	1.04	1.2	1.02	.5	E .48	.50	71.0	71.6	Tortoise
8	1239	1320	-3.07	.15	1.02	.2	1.02	.2	e .44	.44	95.0	95.0	Butterfly
11	967	1318	-.44	.08	.97	-.7	.96	-.6	d .56	.55	79.1	78.3	Wasp
9	863	1319	.11	.07	.86	-4.6	.79	-4.4	c .63	.56	77.0	73.8	Kite
12	808	1314	.38	.07	.80	-7.0	.73	-6.2	b .66	.57	79.2	72.6	Pigeon
15	688	1222	.94	.07	.79	-8.1	.74	-7.3	a .66	.54	78.4	69.7	Violin
MEAN	853.3	1269.9	.00	.08	1.00	-.2	1.00	-.3			75.1	75.1	
S.D.	158.4	47.9	1.12	.02	.14	4.7	.20	4.2			8.7	7.3	

9.24 APPENDIX X – ITEM ORDER PICTURE 2 SEPEDI

TABLE 13.1 Baseline Picture Vocabulary Test ZOU972WS.TXT Oct 28 15:10 2009
INPUT: 1362 Learners 22 Objects MEASURED: 404 Learners 10 Objects 2 CATS

Learner: REAL SEP.: 1.06 REL.: .53 ... Object: REAL SEP.: 13.32 REL.: .99

Object STATISTICS: MEASURE ORDER SEPEDI PVT2 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
15	688	1222	.94	.07	.79	-8.1	.74	-7.3	.66	.54	78.4	69.7	Violin
16	693	1223	.92	.07	1.13	4.6	1.12	2.8	.47	.54	63.0	69.8	Padlock
17	720	1220	.78	.07	1.07	2.5	1.06	1.5	.50	.53	66.0	69.9	Toadstool
10	785	1318	.50	.07	1.21	6.8	1.34	6.5	.47	.57	66.7	72.0	Castle
12	808	1314	.38	.07	.80	-7.0	.73	-6.2	.66	.57	79.2	72.6	Pigeon
14	816	1222	.32	.07	1.04	1.2	1.02	.5	.48	.50	71.0	71.6	Tortoise
9	863	1319	.11	.07	.86	-4.6	.79	-4.4	.63	.56	77.0	73.8	Kite
13	954	1223	-.43	.08	1.14	3.3	1.26	3.5	.36	.44	75.0	77.9	Windmill
11	967	1318	-.44	.08	.97	-.7	.96	-.6	.56	.55	79.1	78.3	Wasp
8	1239	1320	-3.07	.15	1.02	.2	1.02	.2	.44	.44	95.0	95.0	Butterfly
MEAN	853.3	1269.9	.00	.08	1.00	-.2	1.00	-.3			75.1	75.1	
S.D.	158.4	47.9	1.12	.02	.14	4.7	.20	4.2			8.7	7.3	

9.25 APPENDIX Y – PICTURE 3 ITEMS

- Show the picture on the opposite page to the child.
- Ask him/her to point to the **yacht** and wait for them to do that.
- Continue this procedure, asking for:

  some cash,
  the microscope,
  some jewellery,
  the saxophone.

- On the Pupil Record Sheet put a line through the bracket beneath each picture correctly identified.

9.26 APPENDIX Z – LEARNER – ITEM STATISTICS ALL LANGUAGES PIC 3

TABLE 6.1 Baseline Picture Vocabulary Test ZOU462WS.TXT Aug 5 12:18 2010
INPUT: 1362 Learners 22 Objects MEASURED: 1130 Learners 5 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 15.66 REL.: 1.00

Learner STATISTICS: MISFIT ORDER ALL 3 LANGUAGES PVT 3

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
601	3	5	.68	1.04	2.26	2.6	5.51	2.7	A-.58	.47	20.0	71.2	23010711
706	3	5	.68	1.04	2.26	2.6	5.51	2.7	B-.58	.47	20.0	71.2	3010171
1024	3	5	.68	1.04	2.26	2.6	5.51	2.7	C-.58	.47	20.0	71.2	14010243
90	1	5	-1.90	1.35	2.49	1.6	5.31	2.1	D-.47	.54	60.0	85.2	5010083
123	1	5	-1.90	1.35	2.49	1.6	5.31	2.1	E-.47	.54	60.0	85.2	5010413
637	1	5	-1.90	1.35	2.49	1.6	5.31	2.1	F-.47	.54	60.0	85.2	24010253
1067	1	5	-1.90	1.35	2.49	1.6	5.31	2.1	G-.47	.54	60.0	85.2	14010723
1246	1	5	-1.90	1.35	2.49	1.6	5.31	2.1	H-.47	.54	60.0	85.2	11010733
1187	3	5	.68	1.04	1.87	2.0	5.21	2.6	I-.36	.47	60.0	71.2	22010523
1217	3	5	.68	1.04	1.87	2.0	5.21	2.6	J-.36	.47	60.0	71.2	11010363
348	3	5	.68	1.04	1.83	1.9	5.17	2.6	K-.34	.47	60.0	71.2	18010331
1269	3	5	.68	1.04	1.83	1.9	5.17	2.6	L-.34	.47	60.0	71.2	20010181
95	1	5	-1.90	1.35	2.47	1.6	4.92	2.0	M-.44	.54	60.0	85.2	5010133
930	1	5	-1.90	1.35	2.47	1.6	4.92	2.0	N-.44	.54	60.0	85.2	3010413
1009	1	5	-1.90	1.35	2.47	1.6	4.92	2.0	O-.44	.54	60.0	85.2	14010083
1078	1	5	-1.90	1.35	2.47	1.6	4.92	2.0	P-.44	.54	60.0	85.2	21010101
1154	2	5	-.44	1.10	3.12	2.6	3.95	2.6	Q-.74	.56	20.0	76.2	22010123
628	2	5	-.44	1.10	2.82	2.3	3.37	2.3	R-.52	.56	20.0	76.2	24010143
1230	2	5	-.44	1.10	2.82	2.3	3.37	2.3	S-.52	.56	20.0	76.2	11010543
351	2	5	-.44	1.10	2.79	2.3	3.27	2.2	T-.50	.56	20.0	76.2	18010361
847	2	5	-.44	1.10	2.79	2.3	3.27	2.2	U-.50	.56	20.0	76.2	16010661
1114	2	5	-.44	1.10	2.79	2.3	3.27	2.2	V-.50	.56	20.0	76.2	21010491
1127	2	5	-.44	1.10	2.79	2.3	3.27	2.2	W-.50	.56	20.0	76.2	21010631

1143	2	5	-.44	1.10	2.79	2.3	3.27	2.2	X-.50	.56	20.0	76.2	21010791
1312	2	5	-.44	1.10	2.79	2.3	3.27	2.2	Y-.50	.56	20.0	76.2	20010631
41	1	4	-1.83	1.38	2.47	1.6	3.06	1.5	Z-.34	.54	50.0	82.4	4010431
BETTER FITTING OMITTED													
1086	1	5	-1.90	1.35	.34	-.9	.20	-.4	z .88	.54	100.0	85.2	21010181
1121	1	5	-1.90	1.35	.34	-.9	.20	-.4	y .88	.54	100.0	85.2	21010571
1147	1	5	-1.90	1.35	.34	-.9	.20	-.4	x .88	.54	100.0	85.2	22010043
1157	1	5	-1.90	1.35	.34	-.9	.20	-.4	w .88	.54	100.0	85.2	22010183
1158	1	5	-1.90	1.35	.34	-.9	.20	-.4	v .88	.54	100.0	85.2	22010193
1159	1	5	-1.90	1.35	.34	-.9	.20	-.4	u .88	.54	100.0	85.2	22010203
1168	1	5	-1.90	1.35	.34	-.9	.20	-.4	t .88	.54	100.0	85.2	22010313
1170	1	5	-1.90	1.35	.34	-.9	.20	-.4	s .88	.54	100.0	85.2	22010333
1179	1	5	-1.90	1.35	.34	-.9	.20	-.4	r .88	.54	100.0	85.2	22010423
1188	1	5	-1.90	1.35	.34	-.9	.20	-.4	q .88	.54	100.0	85.2	22010573
1206	1	5	-1.90	1.35	.34	-.9	.20	-.4	p .88	.54	100.0	85.2	11010223
1212	1	5	-1.90	1.35	.34	-.9	.20	-.4	o .88	.54	100.0	85.2	11010293
1227	1	5	-1.90	1.35	.34	-.9	.20	-.4	n .88	.54	100.0	85.2	11010513
1231	1	5	-1.90	1.35	.34	-.9	.20	-.4	m .88	.54	100.0	85.2	11010583
1233	1	5	-1.90	1.35	.34	-.9	.20	-.4	l .88	.54	100.0	85.2	11010603
1254	1	5	-1.90	1.35	.34	-.9	.20	-.4	k .88	.54	100.0	85.2	11010823
1309	1	5	-1.90	1.35	.34	-.9	.20	-.4	j .88	.54	100.0	85.2	20010601
1321	1	5	-1.90	1.35	.34	-.9	.20	-.4	i .88	.54	100.0	85.2	17010051
1326	1	5	-1.90	1.35	.34	-.9	.20	-.4	h .88	.54	100.0	85.2	17010101
1327	1	5	-1.90	1.35	.34	-.9	.20	-.4	g .88	.54	100.0	85.2	17010111
1339	1	5	-1.90	1.35	.34	-.9	.20	-.4	f .88	.54	100.0	85.2	17010241
1272	1	4	-1.73	1.42	.34	-.9	.23	-.4	e .88	.57	100.0	83.5	20010221
1225	1	3	-1.24	1.62	.22	-.8	.19	-.6	d .98	.65	100.0	84.6	11010473
1226	1	3	-1.24	1.62	.22	-.8	.19	-.6	c .98	.65	100.0	84.6	11010503
1346	1	3	-1.24	1.62	.22	-.8	.19	-.6	b .98	.65	100.0	84.6	17010351
1360	1	3	-1.24	1.62	.22	-.8	.19	-.6	a .98	.65	100.0	84.6	17010541
MEAN	2.7	4.9	.38	1.30	1.01	.1	.91	.1			76.1	77.1	
S.D.	1.4	.4	1.79	.33	.55	.9	.83	.7			20.1	5.0	

9.27 APPENDIX AA – ITEM STATISTICS ALL LANGUAGES PIC 3

```
TABLE 10.1 Baseline Picture Vocabulary Test       ZOU462WS.TXT Aug  5 12:18 2010
INPUT: 1362 Learners  22 Objects  MEASURED: 1130 Learners  5 Objects  2 CATS
-------------------------------------------------------------------------------
Learner: REAL SEP.: .00  REL.: .00 ... Object: REAL SEP.: 15.66  REL.: 1.00

         Object STATISTICS:  MISFIT ORDER ALL 3 LANGUAGES PVT3
```

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
21	722	1111	-.55	.08	1.14	3.7	1.11	2.0	A .54	.59	69.0	73.5	Jewellery
20	549	1115	.48	.08	1.05	1.6	1.01	.3	B .60	.61	69.8	71.9	Microscope
22	417	1105	1.29	.08	1.04	1.2	1.00	.0	C .61	.62	73.5	76.3	Saxophone
19	959	1119	-2.44	.11	.93	-1.1	.67	-2.3	b .55	.52	88.1	88.0	Cash
18	432	1129	1.21	.08	.86	-4.0	.79	-3.1	a .67	.62	79.8	75.8	Yacht
MEAN	615.8	1115.8	.00	.09	1.01	.3	.92	-.6			76.0	77.1	
S.D.	203.4	8.1	1.39	.01	.10	2.6	.16	1.8			7.1	5.7	

9.28 APPENDIX BB – LEARNER-OBJECT STATISTICS AFRIKAANS – PICTURE 3

TABLE 6.1 Baseline Picture Vocabulary Test ZOU342WS.TXT Aug 5 15:29 2010
INPUT: 1362 Learners 22 Objects MEASURED: 1195 Learners 5 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 15.22 REL.: 1.00

Learner STATISTICS: MISFIT ORDER AFR PVT3

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
376	4	5	1.80	1.19	1.77	1.3	9.90	3.1	A-.86	.32	80.0	80.0	18010631
719	4	5	1.80	1.19	1.77	1.3	9.90	3.1	B-.86	.32	80.0	80.0	3010291
85	3	5	.63	1.03	2.63	3.1	4.90	2.8	C-.89	.45	20.0	70.6	5010033
90	1	5	-1.83	1.30	2.24	1.5	4.59	1.9	D-.48	.50	60.0	83.6	5010083
123	1	5	-1.83	1.30	2.24	1.5	4.59	1.9	E-.48	.50	60.0	83.6	5010413
637	1	5	-1.83	1.30	2.24	1.5	4.59	1.9	F-.48	.50	60.0	83.6	24010253
1067	1	5	-1.83	1.30	2.24	1.5	4.59	1.9	G-.48	.50	60.0	83.6	14010723
1246	1	5	-1.83	1.30	2.24	1.5	4.59	1.9	H-.48	.50	60.0	83.6	11010733
601	3	5	.63	1.03	2.17	2.5	4.38	2.6	I-.57	.45	20.0	70.6	23010711
699	3	5	.63	1.03	2.17	2.5	4.38	2.6	J-.57	.45	20.0	70.6	3010091
706	3	5	.63	1.03	2.17	2.5	4.38	2.6	K-.57	.45	20.0	70.6	3010171
1024	3	5	.63	1.03	2.17	2.5	4.38	2.6	L-.57	.45	20.0	70.6	14010243
1049	3	5	.63	1.03	2.17	2.5	4.38	2.6	M-.57	.45	20.0	70.6	14010543
95	1	5	-1.83	1.30	2.23	1.5	4.22	1.8	N-.44	.50	60.0	83.6	5010133
588	1	5	-1.83	1.30	2.23	1.5	4.22	1.8	O-.44	.50	60.0	83.6	23010581
604	1	5	-1.83	1.30	2.23	1.5	4.22	1.8	P-.44	.50	60.0	83.6	23010751
715	1	5	-1.83	1.30	2.23	1.5	4.22	1.8	Q-.44	.50	60.0	83.6	3010251
725	1	5	-1.83	1.30	2.23	1.5	4.22	1.8	R-.44	.50	60.0	83.6	3010351
930	1	5	-1.83	1.30	2.23	1.5	4.22	1.8	S-.44	.50	60.0	83.6	3010413
1009	1	5	-1.83	1.30	2.23	1.5	4.22	1.8	T-.44	.50	60.0	83.6	14010083
1078	1	5	-1.83	1.30	2.23	1.5	4.22	1.8	U-.44	.50	60.0	83.6	21010101
105	3	5	.63	1.03	1.81	1.9	4.09	2.4	V-.34	.45	60.0	70.6	5010243
169	3	5	.63	1.03	1.81	1.9	4.09	2.4	W-.34	.45	60.0	70.6	7010013
192	3	5	.63	1.03	1.81	1.9	4.09	2.4	X-.34	.45	60.0	70.6	7010263

1001	3	5	.63	1.03	1.81	1.9	4.09	2.4	Y-.34	.45	60.0	70.6	13010803
1187	3	5	.63	1.03	1.81	1.9	4.09	2.4	Z-.34	.45	60.0	70.6	22010523
1217	3	5	.63	1.03	1.81	1.9	4.09	2.4	-.34	.45	60.0	70.6	11010363
4	3	5	.63	1.03	1.76	1.8	4.05	2.4	-.32	.45	60.0	70.6	4010021
348	3	5	.63	1.03	1.76	1.8	4.05	2.4	-.32	.45	60.0	70.6	18010331
1140	3	5	.63	1.03	1.76	1.8	4.05	2.4	-.32	.45	60.0	70.6	21010761
1269	3	5	.63	1.03	1.76	1.8	4.05	2.4	-.32	.45	60.0	70.6	20010181
1154	2	5	-.46	1.08	2.90	2.5	3.49	2.5	-.75	.53	20.0	75.1	22010123
1184	2	5	-.46	1.08	2.90	2.5	3.49	2.5	-.75	.53	20.0	75.1	22010493
628	2	5	-.46	1.08	2.61	2.2	2.96	2.2	-.53	.53	20.0	75.1	24010143
1230	2	5	-.46	1.08	2.61	2.2	2.96	2.2	-.53	.53	20.0	75.1	11010543
84	2	5	-.46	1.08	2.58	2.2	2.86	2.1	-.50	.53	20.0	75.1	5010023
351	2	5	-.46	1.08	2.58	2.2	2.86	2.1	-.50	.53	20.0	75.1	18010361
847	2	5	-.46	1.08	2.58	2.2	2.86	2.1	-.50	.53	20.0	75.1	16010661
1114	2	5	-.46	1.08	2.58	2.2	2.86	2.1	-.50	.53	20.0	75.1	21010491
1127	2	5	-.46	1.08	2.58	2.2	2.86	2.1	-.50	.53	20.0	75.1	21010631
1143	2	5	-.46	1.08	2.58	2.2	2.86	2.1	-.50	.53	20.0	75.1	21010791
1312	2	5	-.46	1.08	2.58	2.2	2.86	2.1	-.50	.53	20.0	75.1	20010631
BETTER FITTING OMITTED													
1086	1	5	-1.83	1.30	.39	-.8	.24	-.4	z .86	.50	100.0	83.6	21010181
1121	1	5	-1.83	1.30	.39	-.8	.24	-.4	y .86	.50	100.0	83.6	21010571
1147	1	5	-1.83	1.30	.39	-.8	.24	-.4	x .86	.50	100.0	83.6	22010043
1157	1	5	-1.83	1.30	.39	-.8	.24	-.4	w .86	.50	100.0	83.6	22010183
1158	1	5	-1.83	1.30	.39	-.8	.24	-.4	v .86	.50	100.0	83.6	22010193
1159	1	5	-1.83	1.30	.39	-.8	.24	-.4	u .86	.50	100.0	83.6	22010203
1168	1	5	-1.83	1.30	.39	-.8	.24	-.4	t .86	.50	100.0	83.6	22010313
1170	1	5	-1.83	1.30	.39	-.8	.24	-.4	s .86	.50	100.0	83.6	22010333
1179	1	5	-1.83	1.30	.39	-.8	.24	-.4	r .86	.50	100.0	83.6	22010423
1188	1	5	-1.83	1.30	.39	-.8	.24	-.4	q .86	.50	100.0	83.6	22010573
1206	1	5	-1.83	1.30	.39	-.8	.24	-.4	p .86	.50	100.0	83.6	11010223
1212	1	5	-1.83	1.30	.39	-.8	.24	-.4	o .86	.50	100.0	83.6	11010293
1227	1	5	-1.83	1.30	.39	-.8	.24	-.4	n .86	.50	100.0	83.6	11010513
1231	1	5	-1.83	1.30	.39	-.8	.24	-.4	m .86	.50	100.0	83.6	11010583
1233	1	5	-1.83	1.30	.39	-.8	.24	-.4	l .86	.50	100.0	83.6	11010603
1254	1	5	-1.83	1.30	.39	-.8	.24	-.4	k .86	.50	100.0	83.6	11010823
1309	1	5	-1.83	1.30	.39	-.8	.24	-.4	j .86	.50	100.0	83.6	20010601
1321	1	5	-1.83	1.30	.39	-.8	.24	-.4	i .86	.50	100.0	83.6	17010051
1326	1	5	-1.83	1.30	.39	-.8	.24	-.4	h .86	.50	100.0	83.6	17010101
1327	1	5	-1.83	1.30	.39	-.8	.24	-.4	g .86	.50	100.0	83.6	17010111
1339	1	5	-1.83	1.30	.39	-.8	.24	-.4	f .86	.50	100.0	83.6	17010241

1272	1	4	-1.65	1.38	.39	-.9	.26	-.3	e .86	.54	100.0	81.7	20010221
1225	1	3	-1.16	1.55	.25	-.8	.22	-.6	d .98	.61	100.0	82.5	11010473
1226	1	3	-1.16	1.55	.25	-.8	.22	-.6	c .98	.61	100.0	82.5	11010503
1346	1	3	-1.16	1.55	.25	-.8	.22	-.6	b .98	.61	100.0	82.5	17010351
1360	1	3	-1.16	1.55	.25	-.8	.22	-.6	a .98	.61	100.0	82.5	17010541
MEAN	2.7	4.9	.33	1.28	1.01	.1	.95	.1			75.5	76.3	
S.D.	1.4	.4	1.75	.33	.51	.9	.90	.8			20.2	4.9	

9.29 APPENDIX CC – OBJECT STATISTICS AFRIKAANS – PICTURE 3

TABLE 10.1 Baseline Picture Vocabulary Test ZOU518WS.TXT Oct 28 15:03 2009
INPUT: 1362 Learners 22 Objects MEASURED: 303 Learners 5 Objects 2 CATS

Learner: REAL SEP.: 1.73 REL.: .75 ... Object: REAL SEP.: 8.73 REL.: .99

Object STATISTICS: MISFIT ORDER AFRIKAANS PVT3 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
21	758	1175	1.47	.07	1.13	4.3	1.13	2.4	A .40	.47	65.7	71.8	Jewellery
22	441	1168	2.99	.07	1.10	2.8	1.07	1.3	B .49	.54	71.4	74.4	Saxophone
20	580	1179	2.31	.07	.97	-1.0	.94	-1.2	C .54	.52	71.4	71.1	Microscope
19	988	1183	.17	.09	.89	-2.2	.75	-2.5	b .43	.36	84.6	83.6	Cash
18	458	1193	2.93	.07	.82	-5.8	.87	-2.5	a .64	.54	80.2	74.2	Yacht
MEAN	645.0	1179.6	1.97	.07	.98	-.4	.95	-.5			74.7	75.0	
S.D.	205.5	8.3	1.06	.01	.12	3.6	.14	2.0			6.8	4.5	

9.30 APPENDIX DD – OBJECT ORDER AFRIKAANS – PICTURE 3

TABLE 13.1 Baseline Picture Vocabulary Test ZOU518WS.TXT Oct 28 15:03 2009
INPUT: 1362 Learners 22 Objects MEASURED: 303 Learners 5 Objects 2 CATS

Learner: REAL SEP.: 1.73 REL.: .75 ... Object: REAL SEP.: 8.73 REL.: .99

Object STATISTICS: MEASURE ORDER AFRIKAANS PVT3 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
22	441	1168	2.99	.07	1.10	2.8	1.07	1.3	.49	.54	71.4	74.4	Saxophone
18	458	1193	2.93	.07	.82	-5.8	.87	-2.5	.64	.54	80.2	74.2	Yacht
20	580	1179	2.31	.07	.97	-1.0	.94	-1.2	.54	.52	71.4	71.1	Microscope
21	758	1175	1.47	.07	1.13	4.3	1.13	2.4	.40	.47	65.7	71.8	Jewellery
19	988	1183	.17	.09	.89	-2.2	.75	-2.5	.43	.36	84.6	83.6	Cash
MEAN	645.0	1179.6	1.97	.07	.98	-.4	.95	-.5			74.7	75.0	
S.D.	205.5	8.3	1.06	.01	.12	3.6	.14	2.0			6.8	4.5	

9.31 APPENDIX EE – LEARNER-OBJECT STATISTICS ENGLISH – PICTURE 3

TABLE 6.1 Baseline Picture Vocabulary Test ZOU454WS.TXT Aug 5 15:32 2010
INPUT: 1362 Learners 22 Objects MEASURED: 482 Learners 5 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 15.44 REL.: 1.00

"???????1" Learner STATISTICS: MISFIT ORDER ENG pvt3

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
601	3	5	.65	1.03	2.20	2.5	4.79	2.6	A	-.56	.46	20.0	70.7	23010711
706	3	5	.65	1.03	2.20	2.5	4.79	2.6	B	-.56	.46	20.0	70.7	3010171
1078	1	5	-1.86	1.32	2.33	1.5	4.66	1.9	C	-.45	.52	60.0	84.2	21010101
348	3	5	.65	1.03	1.82	1.9	4.49	2.5	D	-.34	.46	60.0	70.7	18010331
1269	3	5	.65	1.03	1.82	1.9	4.49	2.5	E	-.34	.46	60.0	70.7	20010181
351	2	5	-.45	1.09	2.70	2.3	3.07	2.2	F	-.52	.55	20.0	75.8	18010361
847	2	5	-.45	1.09	2.70	2.3	3.07	2.2	G	-.52	.55	20.0	75.8	16010661
1114	2	5	-.45	1.09	2.70	2.3	3.07	2.2	H	-.52	.55	20.0	75.8	21010491
1127	2	5	-.45	1.09	2.70	2.3	3.07	2.2	I	-.52	.55	20.0	75.8	21010631
1143	2	5	-.45	1.09	2.70	2.3	3.07	2.2	J	-.52	.55	20.0	75.8	21010791
1312	2	5	-.45	1.09	2.70	2.3	3.07	2.2	K	-.52	.55	20.0	75.8	20010631
41	1	4	-1.78	1.36	2.33	1.6	2.88	1.4	L	-.34	.51	50.0	81.3	4010431
383	2	5	-.45	1.09	2.11	1.7	2.60	1.8	M	-.19	.55	60.0	75.8	18010711
368	2	5	-.45	1.09	2.11	1.7	2.60	1.8	N	-.19	.55	60.0	75.8	18010541
694	2	5	-.45	1.09	2.11	1.7	2.60	1.8	O	-.19	.55	60.0	75.8	3010041
69	4	5	1.83	1.19	1.59	1.1	2.47	1.2	P	-.22	.32	80.0	80.0	4010711
816	4	5	1.83	1.19	1.59	1.1	2.47	1.2	Q	-.22	.32	80.0	80.0	16010541
543	1	5	-1.86	1.32	2.18	1.4	2.45	1.2	R	-.18	.52	60.0	84.2	23010071
1090	1	5	-1.86	1.32	2.18	1.4	2.45	1.2	S	-.18	.52	60.0	84.2	21010231
1097	1	5	-1.86	1.32	2.18	1.4	2.45	1.2	T	-.18	.52	60.0	84.2	21010301
1352	1	5	-1.86	1.32	2.18	1.4	2.45	1.2	U	-.18	.52	60.0	84.2	17010451
19	2	5	-.45	1.09	1.82	1.4	2.02	1.4	V	.03	.55	60.0	75.8	4010181
691	2	5	-.45	1.09	1.82	1.4	2.02	1.4	W	.03	.55	60.0	75.8	3010011
1132	2	5	-.45	1.09	1.82	1.4	2.02	1.4	X	.03	.55	60.0	75.8	21010681

1293	2	5	-.45	1.09	1.82	1.4	2.02	1.4	Y .03	.55	60.0	75.8	20010441
553	3	4	1.20	1.27	1.75	1.3	1.85	1.0	Z-.12	.40	50.0	75.0	23010171
BETTER FITTING OMITTED													
758	1	5	-1.86	1.32	.37	-.8	.22	-.4	z .87	.52	100.0	84.2	3010691
761	1	5	-1.86	1.32	.37	-.8	.22	-.4	y .87	.52	100.0	84.2	3010721
765	1	5	-1.86	1.32	.37	-.8	.22	-.4	x .87	.52	100.0	84.2	3010771
766	1	5	-1.86	1.32	.37	-.8	.22	-.4	w .87	.52	100.0	84.2	3010781
768	1	5	-1.86	1.32	.37	-.8	.22	-.4	v .87	.52	100.0	84.2	3010801
792	1	5	-1.86	1.32	.37	-.8	.22	-.4	u .87	.52	100.0	84.2	16010281
797	1	5	-1.86	1.32	.37	-.8	.22	-.4	t .87	.52	100.0	84.2	16010351
802	1	5	-1.86	1.32	.37	-.8	.22	-.4	s .87	.52	100.0	84.2	16010401
812	1	5	-1.86	1.32	.37	-.8	.22	-.4	r .87	.52	100.0	84.2	16010501
813	1	5	-1.86	1.32	.37	-.8	.22	-.4	q .87	.52	100.0	84.2	16010511
814	1	5	-1.86	1.32	.37	-.8	.22	-.4	p .87	.52	100.0	84.2	16010521
817	1	5	-1.86	1.32	.37	-.8	.22	-.4	o .87	.52	100.0	84.2	16010551
818	1	5	-1.86	1.32	.37	-.8	.22	-.4	n .87	.52	100.0	84.2	16010561
823	1	5	-1.86	1.32	.37	-.8	.22	-.4	m .87	.52	100.0	84.2	16010621
837	1	5	-1.86	1.32	.37	-.8	.22	-.4	l .87	.52	100.0	84.2	16010811
843	1	5	-1.86	1.32	.37	-.8	.22	-.4	k .87	.52	100.0	84.2	16010891
1086	1	5	-1.86	1.32	.37	-.8	.22	-.4	j .87	.52	100.0	84.2	21010181
1121	1	5	-1.86	1.32	.37	-.8	.22	-.4	i .87	.52	100.0	84.2	21010571
1309	1	5	-1.86	1.32	.37	-.8	.22	-.4	h .87	.52	100.0	84.2	20010601
1321	1	5	-1.86	1.32	.37	-.8	.22	-.4	g .87	.52	100.0	84.2	17010051
1326	1	5	-1.86	1.32	.37	-.8	.22	-.4	f .87	.52	100.0	84.2	17010101
1327	1	5	-1.86	1.32	.37	-.8	.22	-.4	e .87	.52	100.0	84.2	17010111
1339	1	5	-1.86	1.32	.37	-.8	.22	-.4	d .87	.52	100.0	84.2	17010241
1272	1	4	-1.69	1.39	.37	-.9	.25	-.3	c .87	.55	100.0	82.4	20010221
1346	1	3	-1.18	1.58	.24	-.8	.20	-.6	b .98	.63	100.0	83.6	17010351
1360	1	3	-1.18	1.58	.24	-.8	.20	-.6	a .98	.63	100.0	83.6	17010541
MEAN	2.5	4.9	.08	1.23	.96	.0	.85	.1			76.4	76.1	
S.D.	1.2	.5	1.56	.29	.49	.9	.69	.6			20.1	4.9	

9.32 APPENDIX FF – OBJECT STATISTICS ENGLISH – PICTURE 3

TABLE 10.1 Baseline Picture Vocabulary Test ZOU450WS.TXT Oct 28 12:59 2009
INPUT: 1362 Learners 22 Objects MEASURED: 514 Learners 5 Objects 2 CATS

Learner: REAL SEP.: 1.74 REL.: .75 ... Object: REAL SEP.: 8.93 REL.: .99

Object STATISTICS: MISFIT ORDER ENGLISH PVT3 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
21	791	1191	1.39	.07	1.12	3.9	1.11	1.9	A .40	.47	66.6	72.5	Jewellery
22	482	1184	2.85	.07	1.06	1.8	1.04	.7	B .51	.54	71.9	73.5	Saxophone
20	607	1195	2.26	.07	.94	-2.0	.90	-2.3	C .56	.52	72.5	71.1	Microscope
19	1021	1199	.04	.09	.90	-2.0	.70	-2.9	b .42	.35	85.7	85.0	Cash
18	484	1208	2.86	.07	.79	-7.2	.70	-6.5	a .67	.54	80.9	73.7	Yacht
MEAN	677.0	1195.4	1.88	.07	.96	-1.1	.89	-1.8			75.5	75.2	
S.D.	205.7	8.0	1.06	.01	.12	3.8	.17	2.9			6.8	5.0	

9.33 APPENDIX GG - OBJECT ORDER ENGLISH – PICTURE 3

TABLE 13.1 Baseline Picture Vocabulary Test ZOU450WS.TXT Oct 28 12:59 2009
INPUT: 1362 Learners 22 Objects MEASURED: 514 Learners 5 Objects 2 CATS

Learner: REAL SEP.: 1.74 REL.: .75 ... Object: REAL SEP.: 8.93 REL.: .99

Object STATISTICS: MEASURE ORDER ENGLISH PVT3 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
18	484	1208	2.86	.07	.79	-7.2	.70	-6.5	.67	.54	80.9	73.7	Yacht
22	482	1184	2.85	.07	1.06	1.8	1.04	.7	.51	.54	71.9	73.5	Saxophone
20	607	1195	2.26	.07	.94	-2.0	.90	-2.3	.56	.52	72.5	71.1	Microscope
21	791	1191	1.39	.07	1.12	3.9	1.11	1.9	.40	.47	66.6	72.5	Jewellery
19	1021	1199	.04	.09	.90	-2.0	.70	-2.9	.42	.35	85.7	85.0	Cash
MEAN	677.0	1195.4	1.88	.07	.96	-1.1	.89	-1.8			75.5	75.2	
S.D.	205.7	8.0	1.06	.01	.12	3.8	.17	2.9			6.8	5.0	

9.34 APPENDIX HH – LEARNER-OBJECT STATISTICS SEPEDI – PICTURE 3

TABLE 6.1 Baseline Picture Vocabulary Test ZOU046WS.TXT Aug 5 15:37 2010
INPUT: 1362 Learners 22 Objects MEASURED: 346 Learners 5 Objects 2 CATS

Learner: REAL SEP.: .00 REL.: .00 ... Object: REAL SEP.: 15.56 REL.: 1.00

"????????3" Learner STATISTICS: MISFIT ORDER SEPEDI PVT3

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
90	1	5	-1.87	1.33	2.38	1.6	5.03	2.0	A-.48	.52	60.0	84.6	5010083
123	1	5	-1.87	1.33	2.38	1.6	5.03	2.0	B-.48	.52	60.0	84.6	5010413
637	1	5	-1.87	1.33	2.38	1.6	5.03	2.0	C-.48	.52	60.0	84.6	24010253
1067	1	5	-1.87	1.33	2.38	1.6	5.03	2.0	D-.48	.52	60.0	84.6	14010723
1246	1	5	-1.87	1.33	2.38	1.6	5.03	2.0	E-.48	.52	60.0	84.6	11010733
1024	3	5	.65	1.03	2.21	2.5	4.96	2.7	F-.58	.46	20.0	70.7	14010243
1187	3	5	.65	1.03	1.85	1.9	4.67	2.6	G-.37	.46	60.0	70.7	22010523
1217	3	5	.65	1.03	1.85	1.9	4.67	2.6	H-.37	.46	60.0	70.7	11010363
95	1	5	-1.87	1.33	2.35	1.5	4.40	1.8	I-.43	.52	60.0	84.6	5010133
930	1	5	-1.87	1.33	2.35	1.5	4.40	1.8	J-.43	.52	60.0	84.6	3010413
1009	1	5	-1.87	1.33	2.35	1.5	4.40	1.8	K-.43	.52	60.0	84.6	14010083
1154	2	5	-.45	1.09	2.99	2.5	3.70	2.6	L-.74	.55	20.0	75.5	22010123
628	2	5	-.45	1.09	2.71	2.3	3.18	2.2	M-.53	.55	20.0	75.5	24010143
1230	2	5	-.45	1.09	2.71	2.3	3.18	2.2	N-.53	.55	20.0	75.5	11010543
223	2	5	-.45	1.09	2.16	1.8	2.75	1.9	O-.23	.55	60.0	75.5	7010593
231	2	5	-.45	1.09	2.16	1.8	2.75	1.9	P-.23	.55	60.0	75.5	7010683
236	2	5	-.45	1.09	2.16	1.8	2.75	1.9	Q-.23	.55	60.0	75.5	7010733
650	2	5	-.45	1.09	2.16	1.8	2.75	1.9	R-.23	.55	60.0	75.5	24010393
937	2	5	-.45	1.09	2.16	1.8	2.75	1.9	S-.23	.55	60.0	75.5	13010043
953	2	5	-.45	1.09	2.16	1.8	2.75	1.9	T-.23	.55	60.0	75.5	13010233
1151	2	5	-.45	1.09	2.16	1.8	2.75	1.9	U-.23	.55	60.0	75.5	22010083
1181	2	5	-.45	1.09	2.16	1.8	2.75	1.9	V-.23	.55	60.0	75.5	22010453
210	1	5	-1.87	1.33	2.20	1.4	2.41	1.2	W-.17	.52	60.0	84.6	7010443
648	1	5	-1.87	1.33	2.20	1.4	2.41	1.2	X-.17	.52	60.0	84.6	24010373

1152	1	5	-1.87	1.33	2.20	1.4	2.41	1.2	Y	-.17	.52	60.0	84.6	22010093
190	2	5	-.45	1.09	1.83	1.4	2.07	1.4	Z	.02	.55	60.0	75.5	7010233
BETTER FITTING OMITTED														
936	1	5	-1.87	1.33	.36	-.9	.22	-.4	z	.88	.52	100.0	84.6	13010033
945	1	5	-1.87	1.33	.36	-.9	.22	-.4	y	.88	.52	100.0	84.6	13010133
962	1	5	-1.87	1.33	.36	-.9	.22	-.4	x	.88	.52	100.0	84.6	13010323
965	1	5	-1.87	1.33	.36	-.9	.22	-.4	w	.88	.52	100.0	84.6	13010363
976	1	5	-1.87	1.33	.36	-.9	.22	-.4	v	.88	.52	100.0	84.6	13010493
1000	1	5	-1.87	1.33	.36	-.9	.22	-.4	u	.88	.52	100.0	84.6	13010793
1007	1	5	-1.87	1.33	.36	-.9	.22	-.4	t	.88	.52	100.0	84.6	14010063
1015	1	5	-1.87	1.33	.36	-.9	.22	-.4	s	.88	.52	100.0	84.6	14010143
1037	1	5	-1.87	1.33	.36	-.9	.22	-.4	r	.88	.52	100.0	84.6	14010403
1053	1	5	-1.87	1.33	.36	-.9	.22	-.4	q	.88	.52	100.0	84.6	14010583
1147	1	5	-1.87	1.33	.36	-.9	.22	-.4	p	.88	.52	100.0	84.6	22010043
1157	1	5	-1.87	1.33	.36	-.9	.22	-.4	o	.88	.52	100.0	84.6	22010183
1158	1	5	-1.87	1.33	.36	-.9	.22	-.4	n	.88	.52	100.0	84.6	22010193
1159	1	5	-1.87	1.33	.36	-.9	.22	-.4	m	.88	.52	100.0	84.6	22010203
1168	1	5	-1.87	1.33	.36	-.9	.22	-.4	l	.88	.52	100.0	84.6	22010313
1170	1	5	-1.87	1.33	.36	-.9	.22	-.4	k	.88	.52	100.0	84.6	22010333
1179	1	5	-1.87	1.33	.36	-.9	.22	-.4	j	.88	.52	100.0	84.6	22010423
1188	1	5	-1.87	1.33	.36	-.9	.22	-.4	i	.88	.52	100.0	84.6	22010573
1206	1	5	-1.87	1.33	.36	-.9	.22	-.4	h	.88	.52	100.0	84.6	11010223
1212	1	5	-1.87	1.33	.36	-.9	.22	-.4	g	.88	.52	100.0	84.6	11010293
1227	1	5	-1.87	1.33	.36	-.9	.22	-.4	f	.88	.52	100.0	84.6	11010513
1231	1	5	-1.87	1.33	.36	-.9	.22	-.4	e	.88	.52	100.0	84.6	11010583
1233	1	5	-1.87	1.33	.36	-.9	.22	-.4	d	.88	.52	100.0	84.6	11010603
1254	1	5	-1.87	1.33	.36	-.9	.22	-.4	c	.88	.52	100.0	84.6	11010823
1225	1	3	-1.21	1.58	.24	-.8	.20	-.6	b	.98	.63	100.0	83.5	11010473
1226	1	3	-1.21	1.58	.24	-.8	.20	-.6	a	.98	.63	100.0	83.5	11010503
MEAN	1.9	4.9	-.61	1.22	1.02	.0	.95	.0				77.6	77.4	
S.D.	1.0	.5	1.31	.26	.62	1.0	.94	.9				20.6	5.2	

9.35 APPENDIX II – OBJECT STATISTICS SEPEDI – PICTURE 3

TABLE 10.1 Baseline Picture Vocabulary Test ZOU074WS.TXT Oct 28 15:14 2009
INPUT: 1362 Learners 22 Objects MEASURED: 404 Learners 5 Objects 2 CATS

Learner: REAL SEP.: 1.65 REL.: .73 ... Object: REAL SEP.: 7.62 REL.: .98

Object STATISTICS: MISFIT ORDER SEPEDI PVT3 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
21	787	1199	1.32	.07	1.11	3.5	1.08	1.6	A .42	.47	67.3	72.4	Jewellery
22	474	1193	2.79	.07	1.04	1.2	1.00	.1	B .53	.54	72.7	73.7	Saxophone
20	616	1203	2.12	.07	.95	-1.8	.92	-1.9	C .55	.52	72.3	71.1	Microscope
19	1030	1207	-.09	.09	.89	-2.2	.70	-2.8	b .42	.35	86.2	85.2	Cash
18	500	1218	2.69	.07	.79	-7.0	.85	-3.1	a .66	.54	80.3	73.3	Yacht
MEAN	681.4	1204.0	1.77	.07	.95	-1.3	.91	-1.2			75.8	75.2	
S.D.	206.3	8.4	1.07	.01	.11	3.6	.13	1.8			6.7	5.1	

9.36 APPENDIX JJ – OBJECT ORDER SEPEDI – PICTURE 3

TABLE 13.1 Baseline Picture Vocabulary Test ZOU074WS.TXT Oct 28 15:14 2009
INPUT: 1362 Learners 22 Objects MEASURED: 404 Learners 5 Objects 2 CATS

Learner: REAL SEP.: 1.65 REL.: .73 ... Object: REAL SEP.: 7.62 REL.: .98

Object STATISTICS: MEASURE ORDER SEPEDI PVT3 BASELINE

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
22	474	1193	2.79	.07	1.04	1.2	1.00	.1	.53	.54	72.7	73.7	Saxophone
18	500	1218	2.69	.07	.79	-7.0	.85	-3.1	.66	.54	80.3	73.3	Yacht
20	616	1203	2.12	.07	.95	-1.8	.92	-1.9	.55	.52	72.3	71.1	Microscope
21	787	1199	1.32	.07	1.11	3.5	1.08	1.6	.42	.47	67.3	72.4	Jewellery
19	1030	1207	-.09	.09	.89	-2.2	.70	-2.8	.42	.35	86.2	85.2	Cash
MEAN	681.4	1204.0	1.77	.07	.95	-1.3	.91	-1.2			75.8	75.2	
S.D.	206.3	8.4	1.07	.01	.11	3.6	.13	1.8			6.7	5.1	

9.37 APPENDIX KK – LEARNER-OBJECT STATISITICS ALL LANGUAGES & PICTURES

TABLE 6.1 Baseline Picture Vocabulary Test ZOU690WS.TXT Oct 27 11:23 2009
INPUT: 1362 Learners 22 Objects MEASURED: 1361 Learners 22 Objects 2 CATS

Learner: REAL SEP.: 1.72 REL.: .75 ... Object: REAL SEP.: 8.90 REL.: .99

All Learner STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Learner
152	17	22	2.27	.57	1.56	2.0	9.90	4.0	A .05	.38	68.2	79.6	10010272
199	9	18	1.04	.54	1.31	1.5	9.90	4.9	B .10	.46	55.6	70.6	7010333
201	10	18	1.33	.54	1.16	.9	9.90	5.0	C .13	.43	61.1	69.0	7010353
252	18	22	2.61	.61	1.62	1.8	9.90	4.8	D-.05	.34	72.7	82.6	19020112
274	17	22	2.27	.57	1.54	1.9	9.90	4.0	E .05	.38	68.2	79.6	19040332
406	18	22	2.61	.61	1.77	2.1	9.90	9.9	F-.35	.34	72.7	82.6	8010202
910	20	22	3.54	.78	1.43	.9	9.90	3.6	G-.14	.24	90.9	90.9	6010632
939	6	22	-1.38	.74	3.44	3.3	9.90	3.7	H .12	.74	68.2	89.2	13010063
980	12	22	.81	.53	2.07	3.8	9.90	5.0	I .04	.55	50.0	76.0	13010543
921	19	22	3.02	.67	1.52	1.3	6.37	2.8	J-.04	.29	86.4	86.4	6010742
121	1	8	-3.28	1.47	3.24	1.8	6.21	2.3	K-.15	.64	75.0	92.9	5010393
248	21	22	4.35	1.05	1.21	.5	6.05	2.2	L-.05	.17	95.5	95.5	19010082
258	21	22	4.35	1.05	1.21	.5	6.05	2.2	M-.05	.17	95.5	95.5	19020172
271	21	22	4.35	1.05	1.21	.5	6.05	2.2	N-.05	.17	95.5	95.5	19040302
290	21	22	4.35	1.05	1.21	.5	6.05	2.2	O-.05	.17	95.5	95.5	19060502
299	21	22	4.35	1.05	1.21	.5	6.05	2.2	P-.05	.17	95.5	95.5	19070592
301	21	22	4.35	1.05	1.21	.5	6.05	2.2	Q-.05	.17	95.5	95.5	19070602
466	21	22	4.35	1.05	1.21	.5	6.05	2.2	R-.05	.17	95.5	95.5	12010032
473	21	22	4.35	1.05	1.21	.5	6.05	2.2	S-.05	.17	95.5	95.5	12010172
530	21	22	4.35	1.05	1.21	.5	6.05	2.2	T-.05	.17	95.5	95.5	12010982
536	21	22	4.35	1.05	1.21	.5	6.05	2.2	U-.05	.17	95.5	95.5	12011082
604	4	20	-2.02	.96	1.10	.4	6.05	2.3	V .68	.79	95.0	93.4	23010751
586	1	8	-3.28	1.47	3.24	1.8	5.90	2.2	W-.14	.64	75.0	92.9	23010561
836	13	22	1.10	.53	1.62	2.6	4.71	2.2	X .25	.52	54.5	74.8	16010801

774	19	22	3.02	.67	1.09	.4	4.58	2.3	Y	.17	.29	86.4	86.4	16010071
944	12	22	.81	.53	1.71	2.7	3.94	1.9	Z	.27	.55	59.1	76.0	13010123
4	13	20	1.63	.54	1.64	2.8	2.57	1.4		.17	.44	65.0	73.0	4010021
699	14	21	1.63	.54	1.59	2.6	2.46	1.3		.21	.46	76.2	74.2	3010091
1001	9	18	1.04	.54	1.81	3.3	2.32	1.5		-.02	.46	44.4	70.6	13010803
1239	13	20	1.63	.54	1.49	2.2	2.13	1.2		.24	.44	55.0	73.0	11010653
342	13	21	1.34	.54	1.52	2.4	1.95	1.0		.30	.49	57.1	73.4	18010231
788	15	21	1.93	.55	1.61	2.5	1.65	.9		.21	.42	57.1	75.9	16010241
719	13	20	1.63	.54	1.59	2.6	1.64	.9		.23	.44	65.0	73.0	3010291
634	10	20	.75	.55	1.59	2.3	1.55	.8		.36	.54	55.0	75.0	24010223
85	10	19	1.04	.54	1.54	2.3	1.52	.8		.27	.47	57.9	72.1	5010033
105	14	22	1.37	.53	1.51	2.3	1.36	.7		.32	.48	59.1	74.0	5010243
457	13	20	1.63	.54	1.48	2.2	1.46	.7		.27	.44	55.0	73.0	8030732
84	13	21	1.34	.54	1.46	2.1	1.16	.5		.36	.49	47.6	73.4	5010023
193	10	17	1.39	.55	1.42	2.0	1.29	.6		.19	.41	47.1	68.5	7010273
BETTER FITTING OMITTED														
756	12	20	1.33	.54	.64	-2.1	.48	-.2		.59	.47	85.0	72.1	3010661
1212	12	20	1.33	.54	.63	-2.1	.48	-.2		.59	.47	95.0	72.1	11010293
611	13	21	1.34	.54	.63	-2.1	.45	-.2		.60	.49	95.2	73.4	23010831
1327	13	20	1.63	.54	.63	-2.1	.47	-.2		.55	.44	95.0	73.0	17010111
589	14	21	1.63	.54	.63	-2.1	.45	-.2		.57	.46	95.2	74.2	23010591
726	13	22	1.10	.53	.58	-2.3	.42	-.3		.64	.52	90.9	74.8	3010361
1309	12	22	.81	.53	.54	-2.4	.39	-.3		.68	.55	95.5	76.0	20010601
96	4	9	-.66	.92	.43	-1.3	.28	-.3	z	.80	.66	88.9	81.6	5010153
181	4	9	-.66	.92	.43	-1.3	.28	-.3	y	.80	.66	88.9	81.6	7010143
679	4	9	-.66	.92	.43	-1.3	.28	-.3	x	.80	.66	88.9	81.6	24010723
941	4	9	-.66	.92	.43	-1.3	.28	-.3	w	.80	.66	88.9	81.6	13010093
970	4	9	-.66	.92	.43	-1.3	.28	-.3	v	.80	.66	88.9	81.6	13010433
88	5	10	-.65	.92	.42	-1.3	.25	-.3	u	.83	.70	90.0	83.2	5010063
94	5	10	-.65	.92	.42	-1.3	.25	-.3	t	.83	.70	90.0	83.2	5010123
200	5	10	-.65	.92	.42	-1.3	.25	-.3	s	.83	.70	90.0	83.2	7010343
950	5	10	-.65	.92	.42	-1.3	.25	-.3	r	.83	.70	90.0	83.2	13010203
1021	5	10	-.65	.92	.42	-1.3	.25	-.3	q	.83	.70	90.0	83.2	14010213
1036	5	10	-.65	.92	.42	-1.3	.25	-.3	p	.83	.70	90.0	83.2	14010383
1054	5	10	-.65	.92	.42	-1.3	.25	-.3	o	.83	.70	90.0	83.2	14010593
1236	5	10	-.65	.92	.42	-1.3	.25	-.3	n	.83	.70	90.0	83.2	11010623
1240	6	10	-.50	.92	.40	-1.4	.24	-.3	m	.81	.68	100.0	83.4	11010663
89	6	11	-.59	.89	.40	-1.4	.23	-.3	l	.84	.71	100.0	84.0	5010073
106	6	11	-.59	.89	.40	-1.4	.23	-.3	k	.84	.71	100.0	84.0	5010253
226	6	11	-.59	.89	.40	-1.4	.23	-.3	j	.84	.71	100.0	84.0	7010623

1195	6	11	-.59	.89	.40	-1.4	.23	-.3	i .84	.71	100.0	84.0	11010093
1210	6	11	-.59	.89	.40	-1.4	.23	-.3	h .84	.71	100.0	84.0	11010263
1219	6	11	-.59	.89	.40	-1.4	.23	-.3	g .84	.71	100.0	84.0	11010393
1228	6	11	-.59	.89	.40	-1.4	.23	-.3	f .84	.71	100.0	84.0	11010523
202	5	11	-1.46	.97	.37	-1.2	.19	-.4	e .89	.77	100.0	86.9	7010363
212	4	11	-2.50	1.05	.25	-1.3	.13	-.6	d .92	.79	100.0	89.7	7010463
227	4	11	-2.50	1.05	.25	-1.3	.13	-.6	c .92	.79	100.0	89.7	7010633
175	2	9	-3.19	1.40	.20	-.9	.08	-.7	b .92	.79	100.0	93.1	7010073
651	3	10	-2.98	1.25	.17	-1.1	.09	-.7	a .93	.82	100.0	92.5	24010403
MEAN	14.4	20.0	2.01	.73	1.01	.0	.95	.2			80.8	80.9	
S.D.	4.8	3.2	1.65	.31	.28	.8	1.05	.6			10.0	6.9	

9.38 APPENDIX LL – OBJECT STATISTICS ALL LANGUAGES & PICTURES

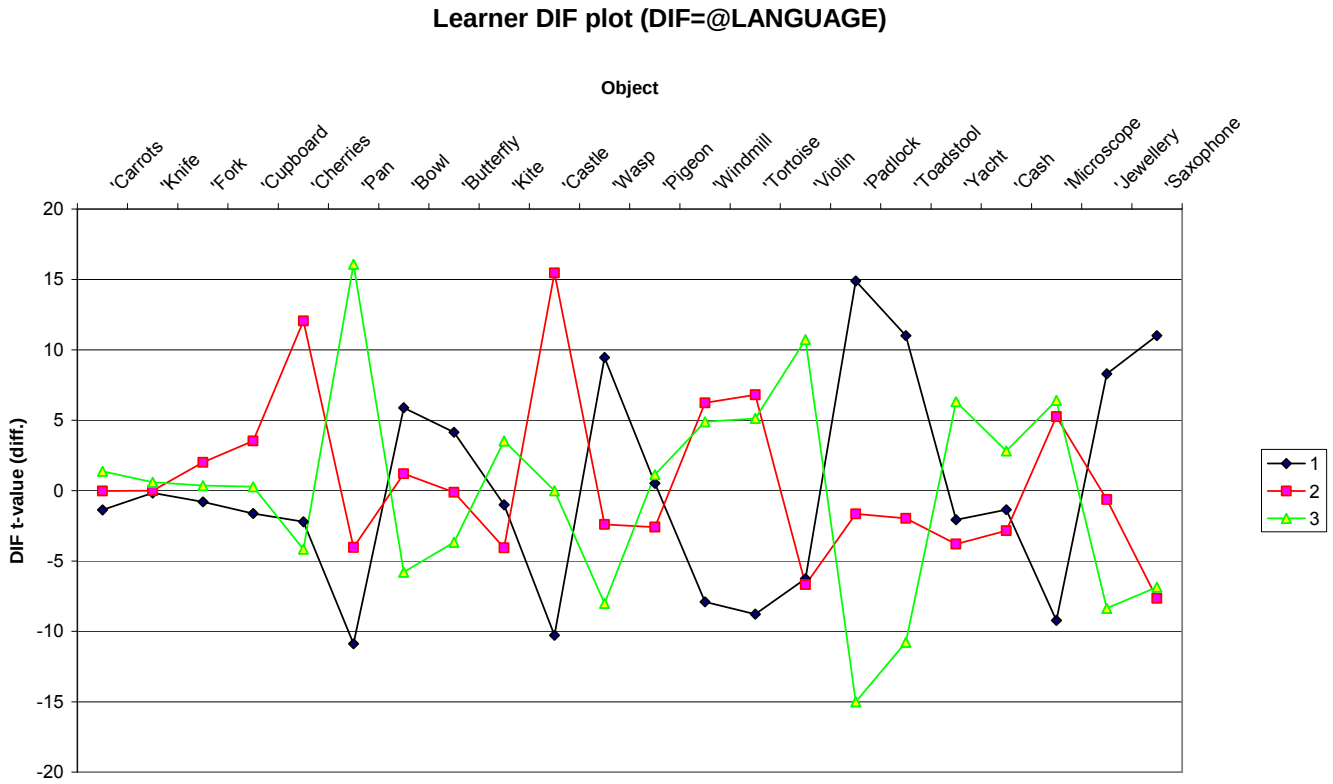
TABLE 10.1 Baseline Picture Vocabulary Test ZOU544WS.TXT Aug 6 11:13 2010
INPUT: 1362 Learners 22 Objects MEASURED: 1221 Learners 22 Objects 2 CATS

Learner: REAL SEP.: 1.70 REL.: .74 ... Object: REAL SEP.: 13.71 REL.: .99

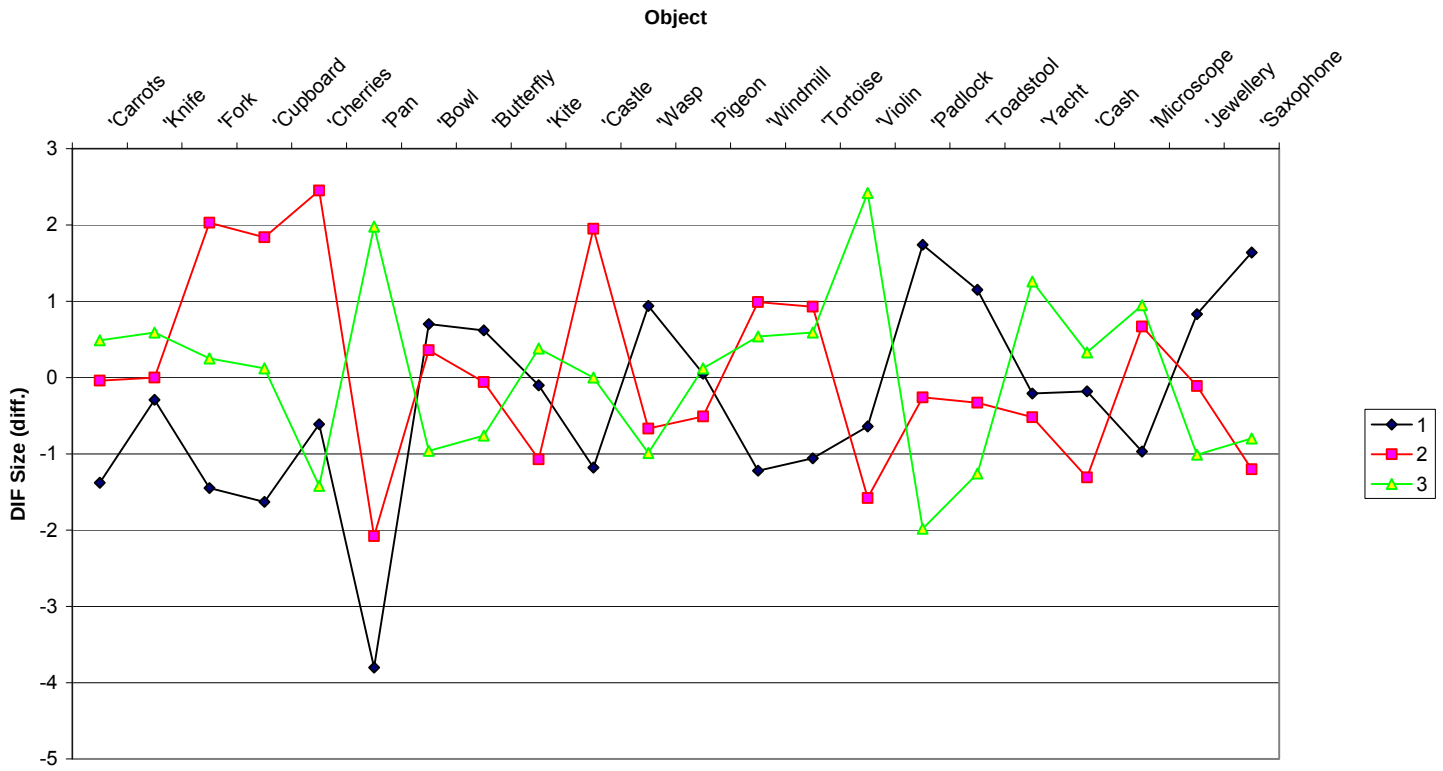
Object STATISTICS: MISFIT ORDER FOR ALL 3 GROUPS

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Object
10	717	1218	.59	.07	1.20	6.2	1.43	6.7	A .43	.54	68.5	73.2	Castle
16	635	1134	.97	.07	1.22	7.1	1.18	3.9	B .39	.51	58.7	70.9	Padlock
5	548	565	-3.11	.27	1.16	.8	.97	.1	C .20	.26	96.6	97.0	Cherries
6	892	1218	-.35	.08	1.16	3.8	1.06	.8	D .45	.51	72.6	79.0	Pan
17	658	1131	.86	.07	1.15	4.9	1.15	3.2	E .42	.50	64.2	71.0	Toadstool
21	722	1111	.50	.07	1.14	4.4	1.14	2.6	F .40	.47	65.5	72.0	Jewellery
13	885	1134	-.37	.08	1.07	1.6	1.09	1.1	G .36	.40	76.2	78.8	Windmill
22	417	1105	2.06	.07	1.09	2.5	1.07	1.1	H .50	.55	71.8	74.5	Saxophone
14	747	1133	.41	.07	1.02	.6	1.08	1.5	I .45	.47	72.9	72.3	Tortoise
11	899	1218	-.38	.08	1.04	1.1	1.04	.5	i .48	.50	78.8	79.2	Wasp
20	549	1115	1.37	.07	.97	-1.1	.94	-1.4	h .55	.53	71.5	71.0	Microscope
7	1095	1219	-1.97	.11	.90	-1.3	.56	-2.7	g .47	.41	91.5	90.6	Bowl
19	959	1119	-.98	.09	.90	-1.8	.70	-2.7	f .41	.34	85.9	85.4	Cash
9	787	1219	.23	.07	.87	-4.1	.83	-2.9	e .59	.53	78.0	74.7	Kite
12	731	1214	.51	.07	.83	-5.8	.75	-4.7	d .62	.54	79.5	73.3	Pigeon
8	1175	1220	-3.39	.17	.82	-1.4	.29	-3.4	c .39	.31	96.7	96.5	Butterfly
18	432	1129	2.00	.07	.80	-6.3	.70	-5.8	b .66	.55	80.8	74.4	Yacht
15	619	1133	1.05	.07	.71	-9.9	.64	-9.4	a .68	.51	83.0	70.8	Violin
MEAN	806.8	1119.0	-1.52	.41	1.00	.1	.92	-.7			77.4	78.0	
S.D.	227.3	141.9	3.49	.67	.15	4.4	.27	3.8			10.3	8.4	

9.39 APPENDIX MM - DIF



Learner DIF plot (DIF=@LANGUAGE)



Learner DIF plot (DIF=@LANGUAGE)