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.

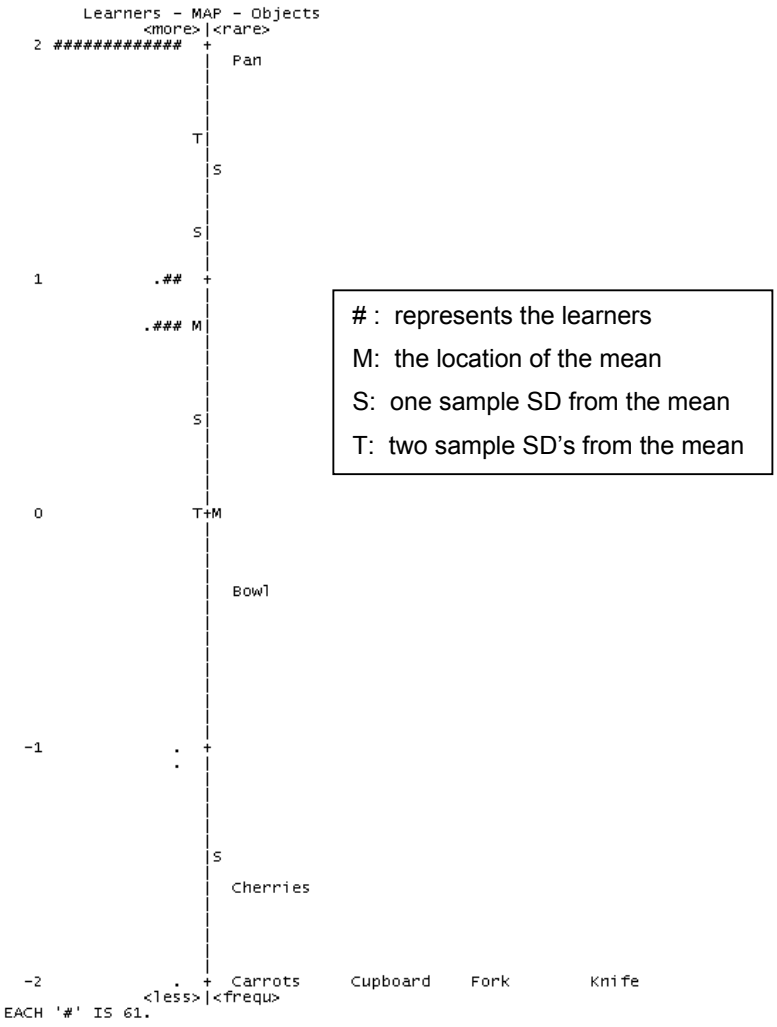

Figure 5.2: Item-Learner Map for all languages – Picture 1

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.

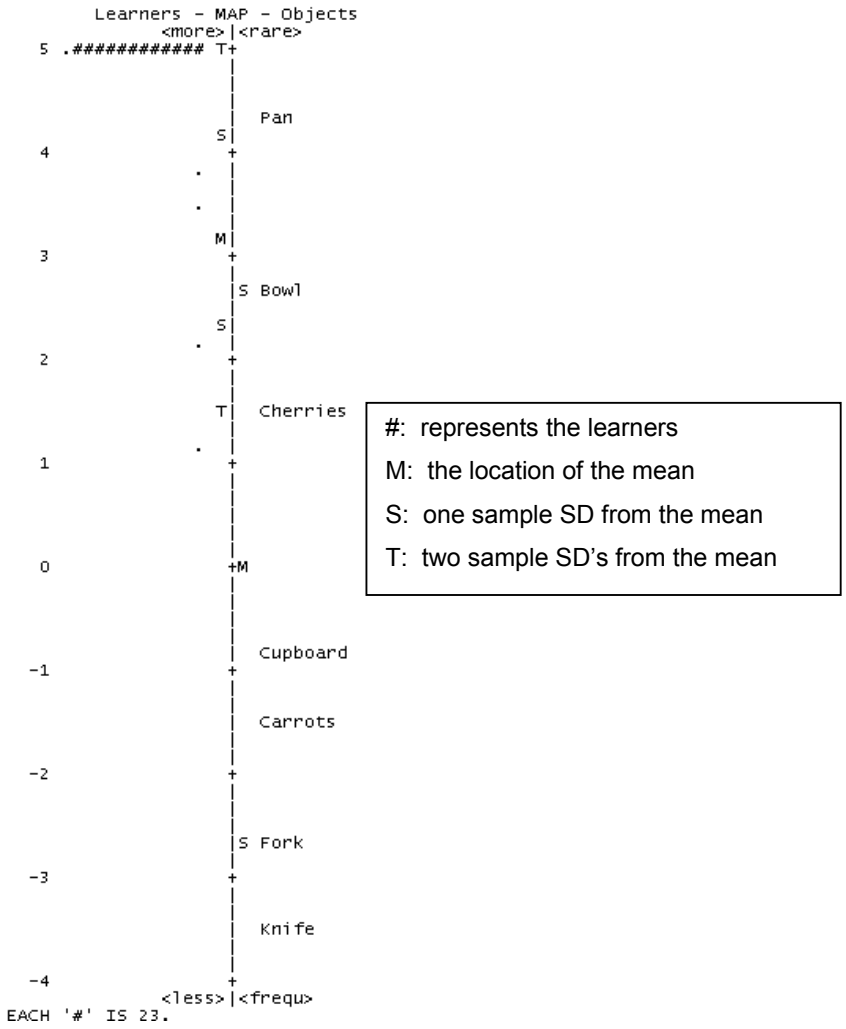

Figure 5.3: Item-Learner Map for Afrikaans – Picture 1

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.

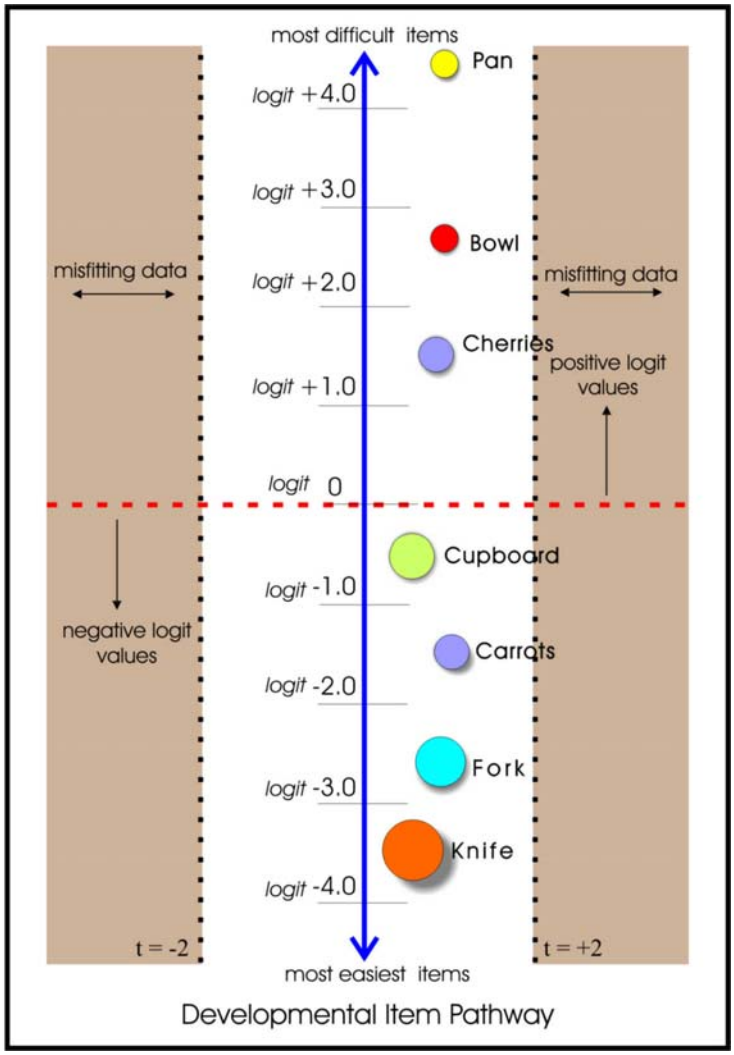

Figure 5.4: Item Development Pathway for Picture 1 – Afrikaans group

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.

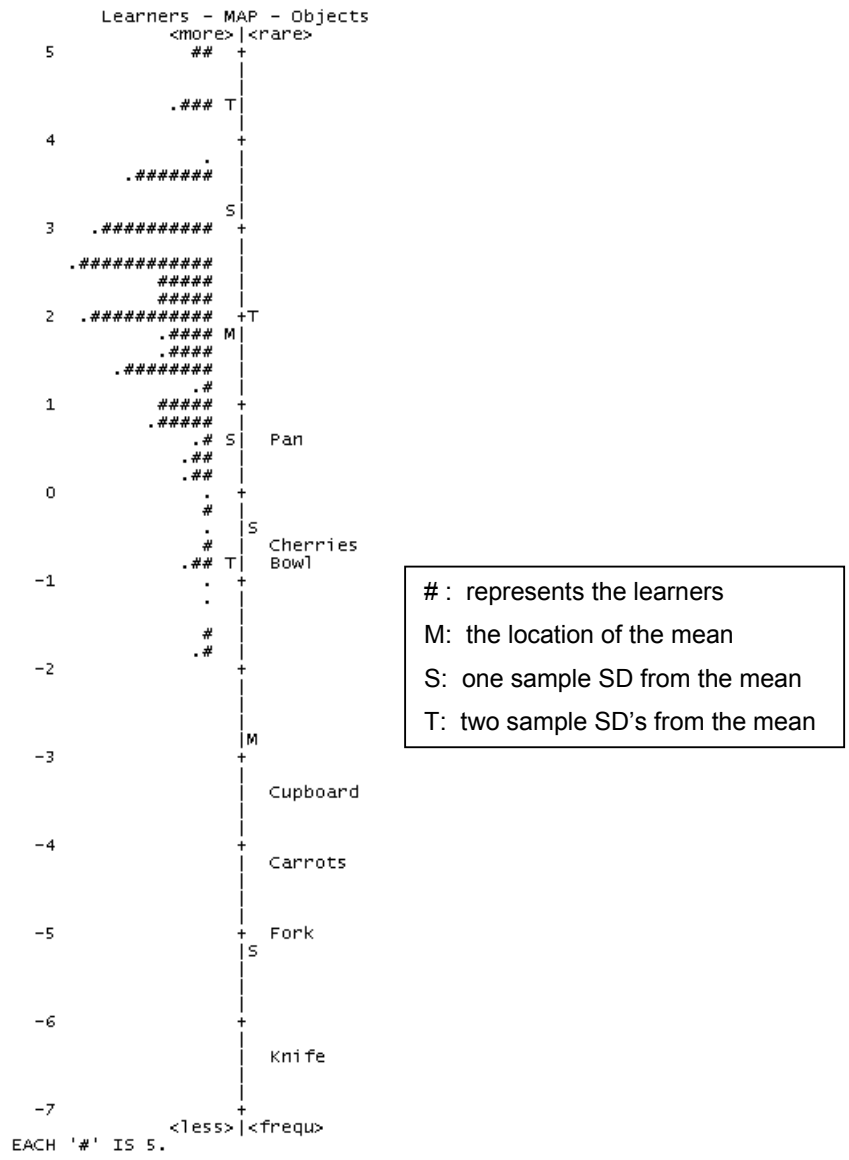

Figure 5.5: Item-Learner Map for English – Picture 1

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.

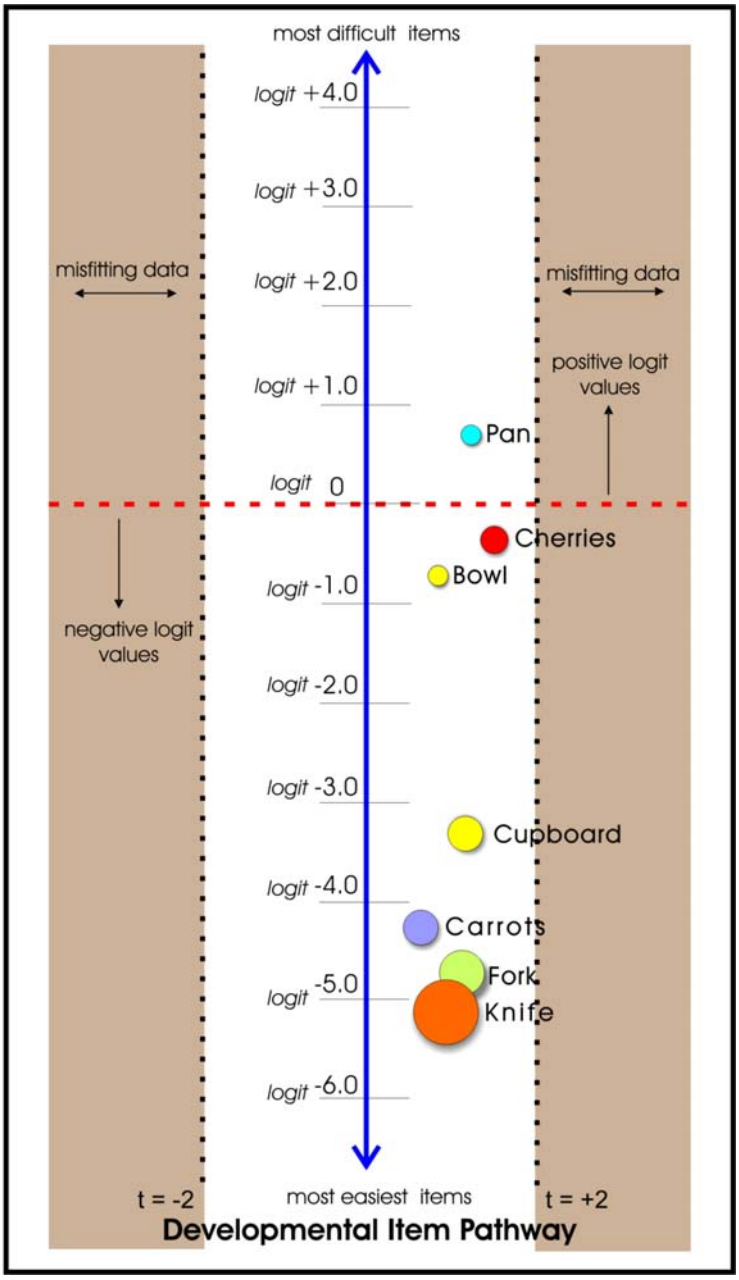

Figure 5.6: IDP for English group – Picture 1

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.

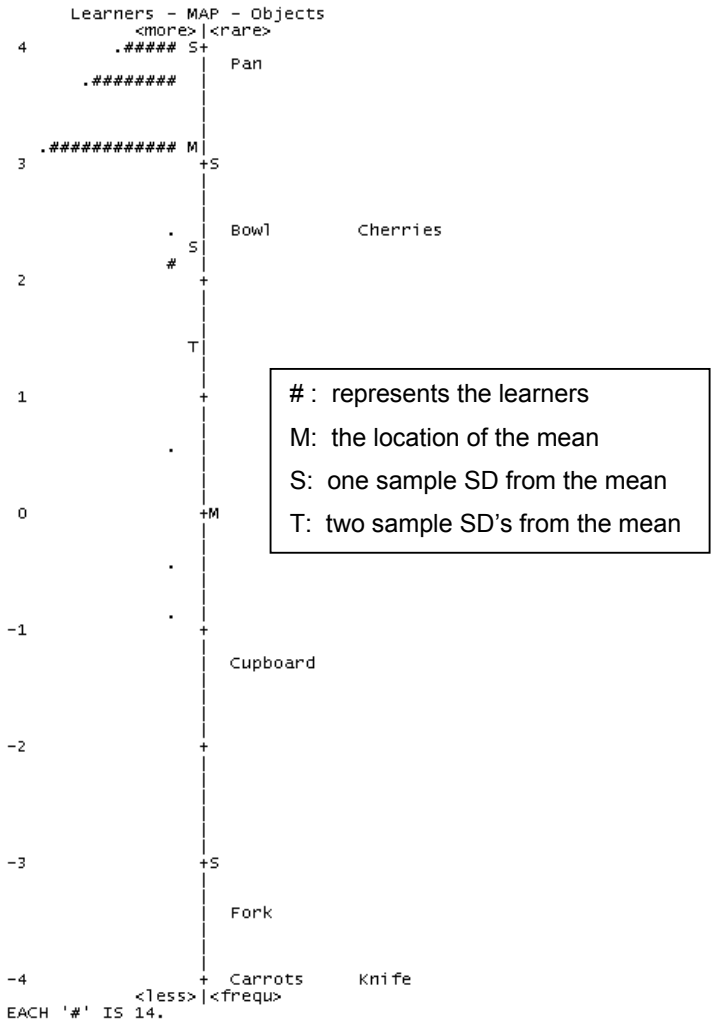

Figure 5.7: Item-Learner Map for Sepedi – Picture 1

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.

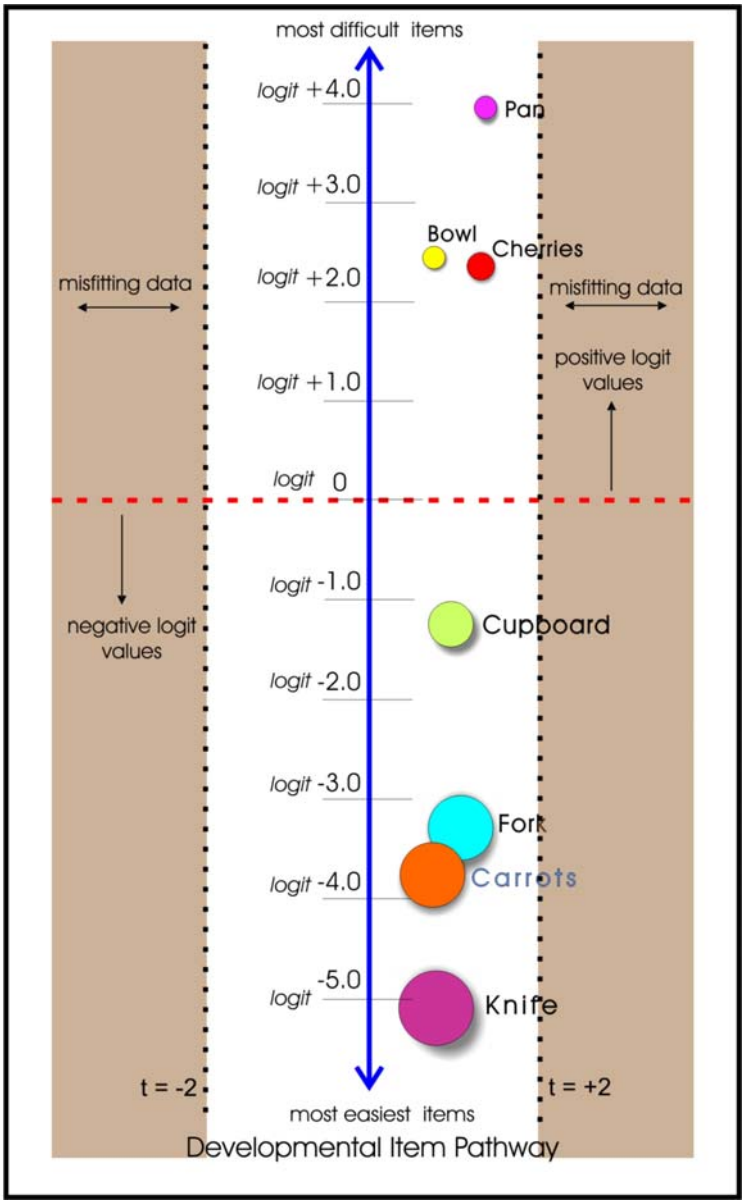

Figure 5.8: IDP for Sepedi group – Picture 1

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.

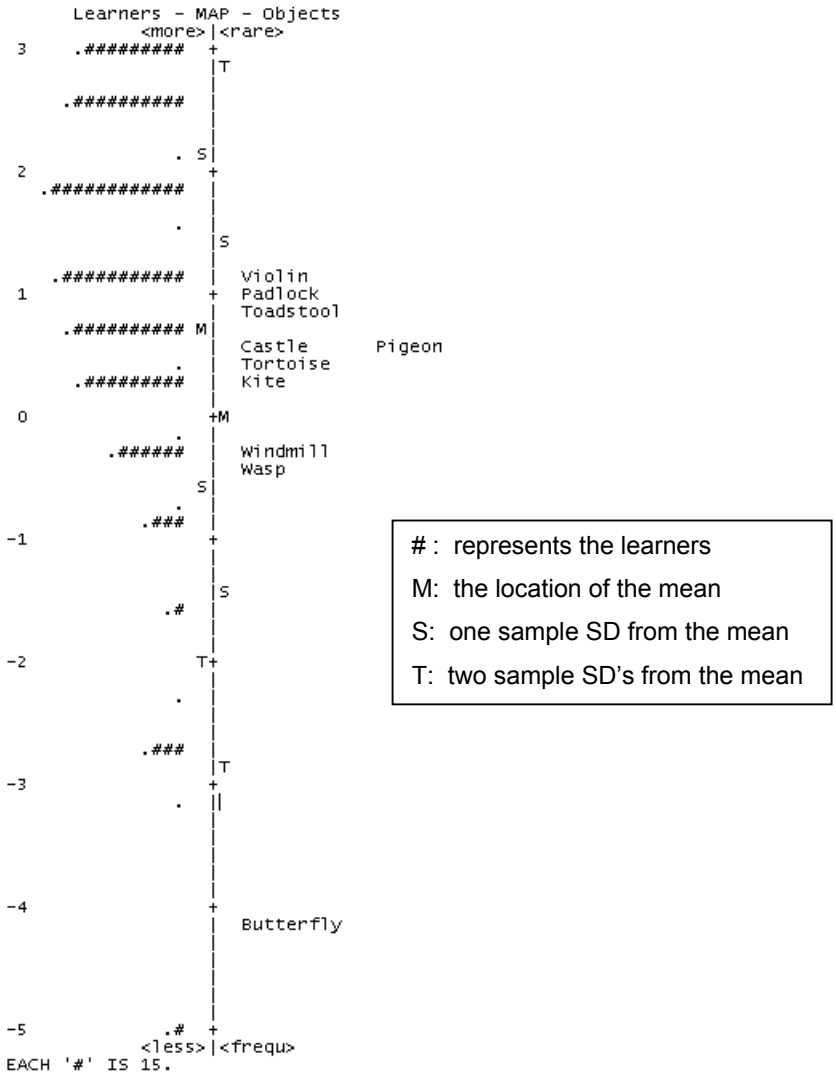

Figure 5.10: Item-Learner Map for all languages– Picture 2

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.

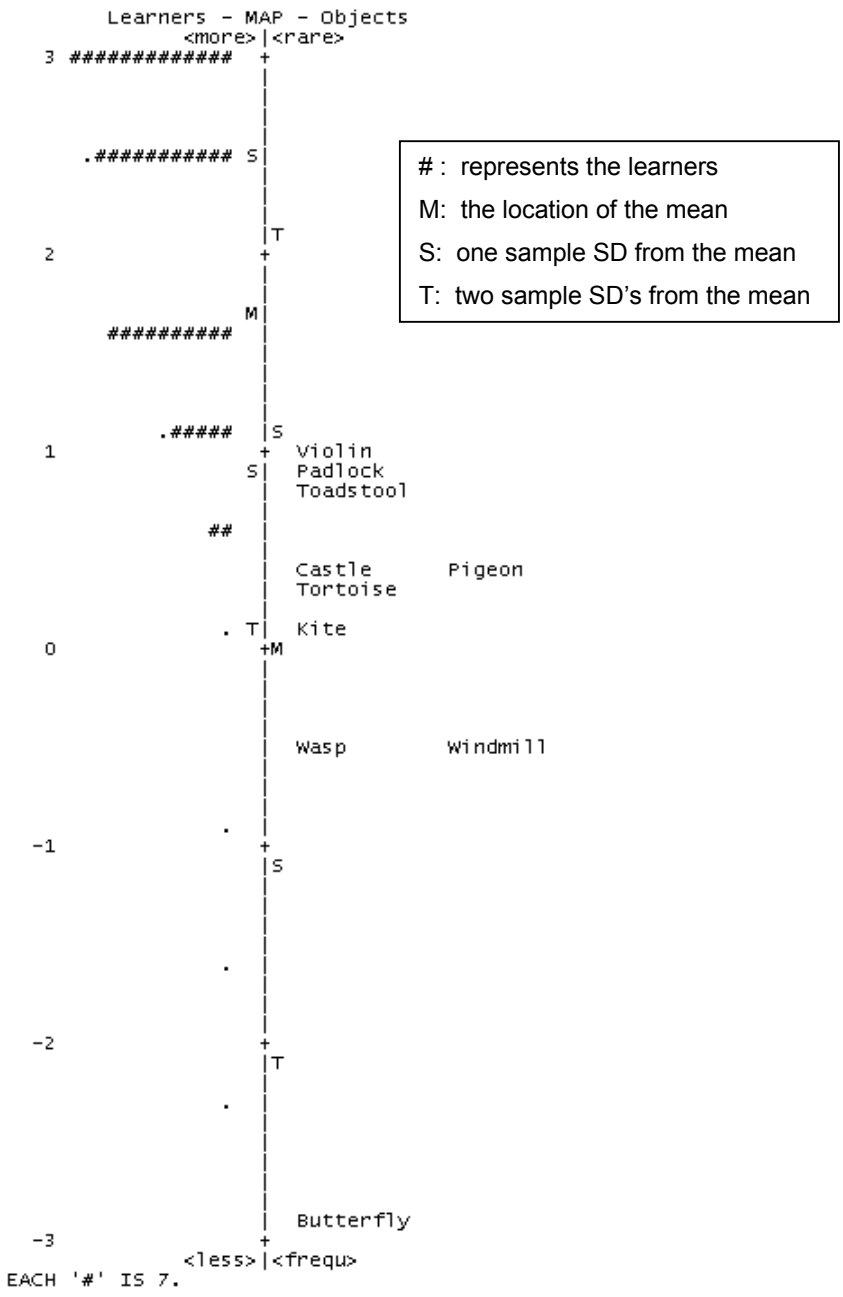

Figure 5.11: Item-Learner Map for Afrikaans– Picture 2

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.

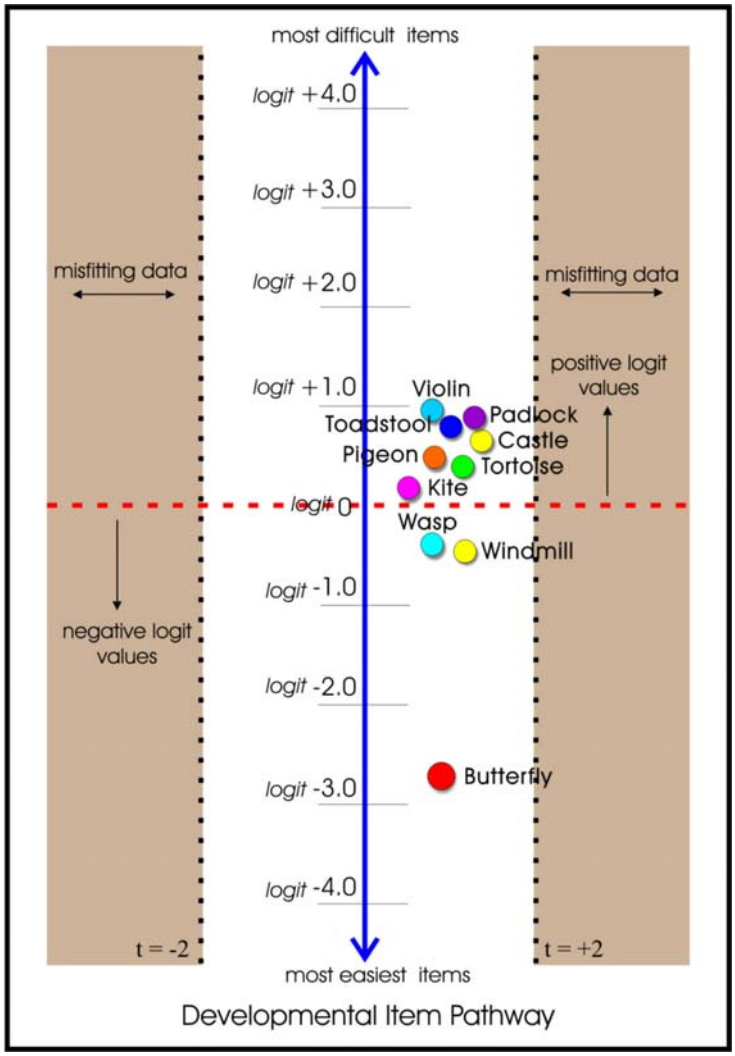

Figure 5.12: IDP for Afrikaans group – Picture 2

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.

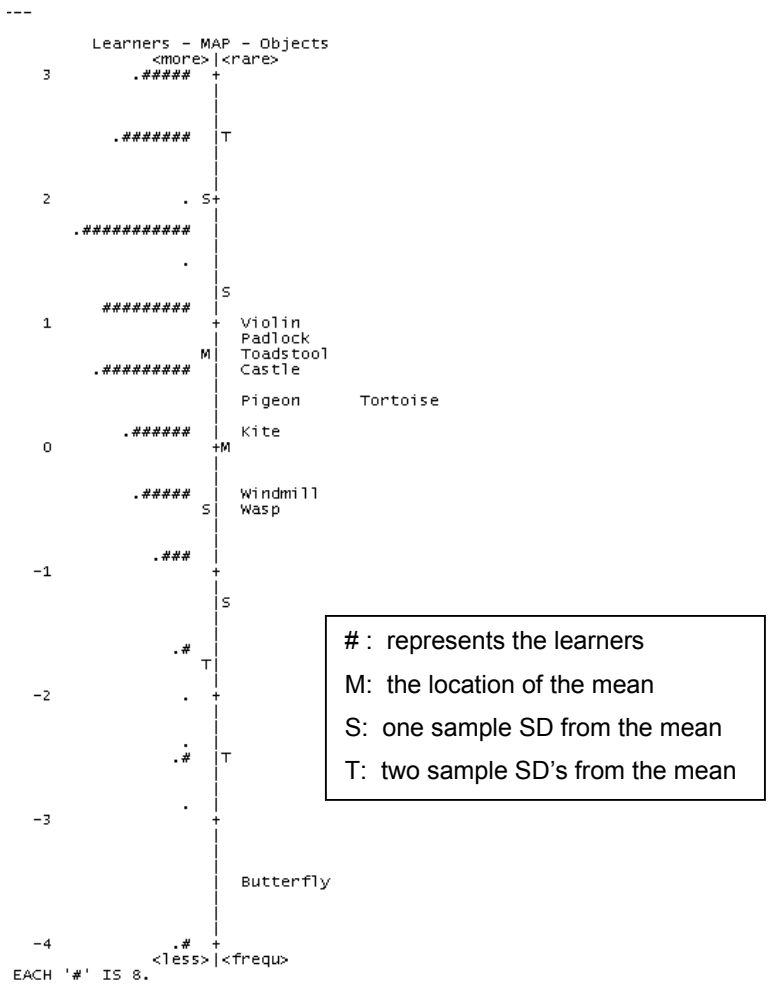

Figure 5.13: Item-Learner Map for Picture 2 - English group

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

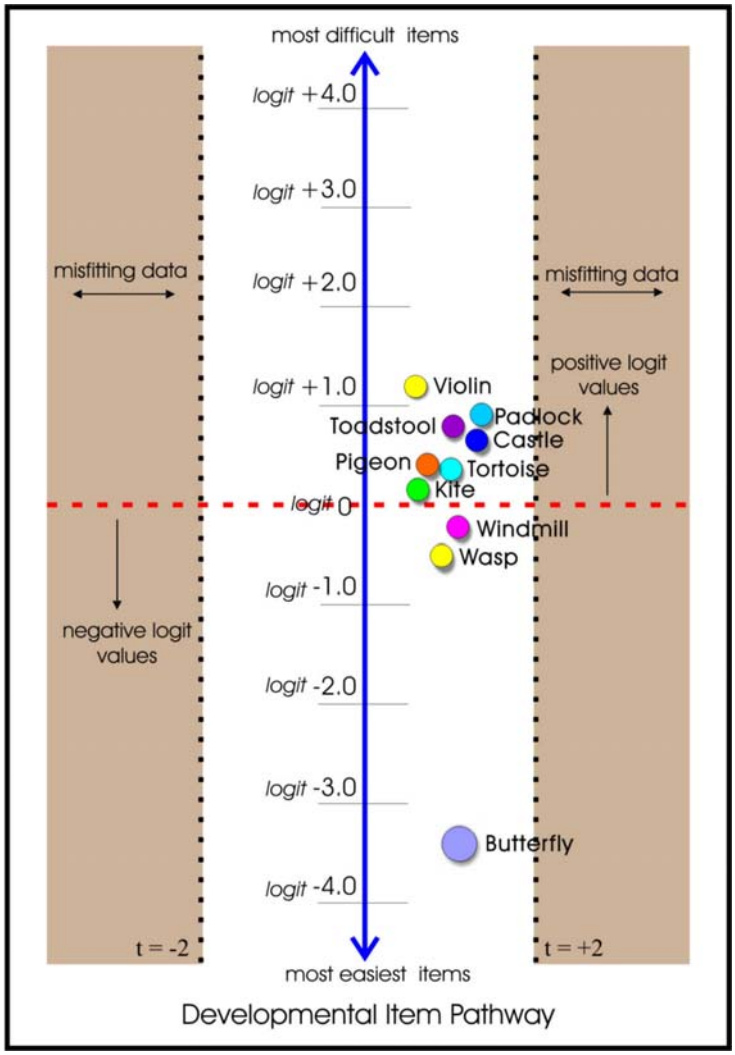

Figure 5.14: IDP for English group – Picture 2

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.

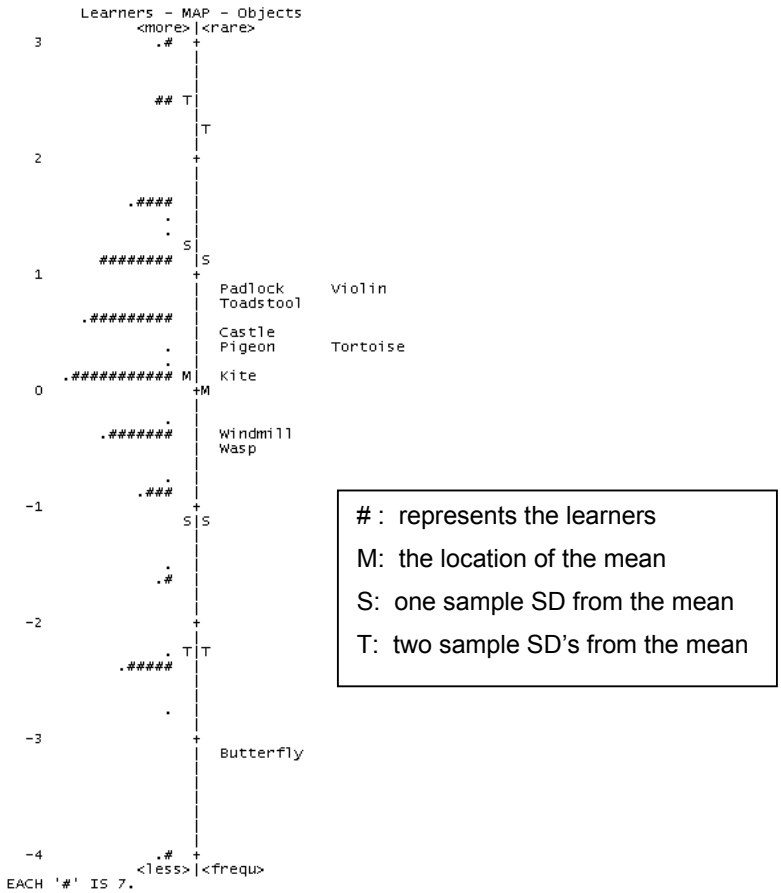

Figure 5.15: Item-Learner Map for Picture 2 - Sepedi group

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

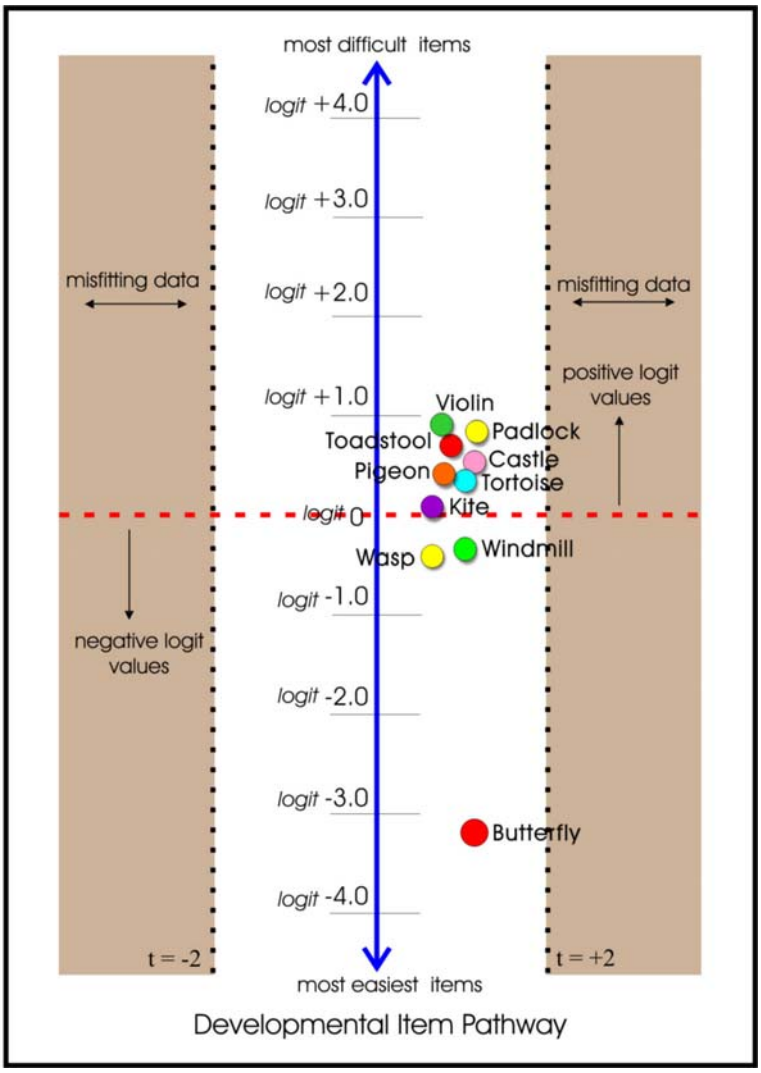

Figure 5.16: IDP for Sepedi group – Picture 2

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.

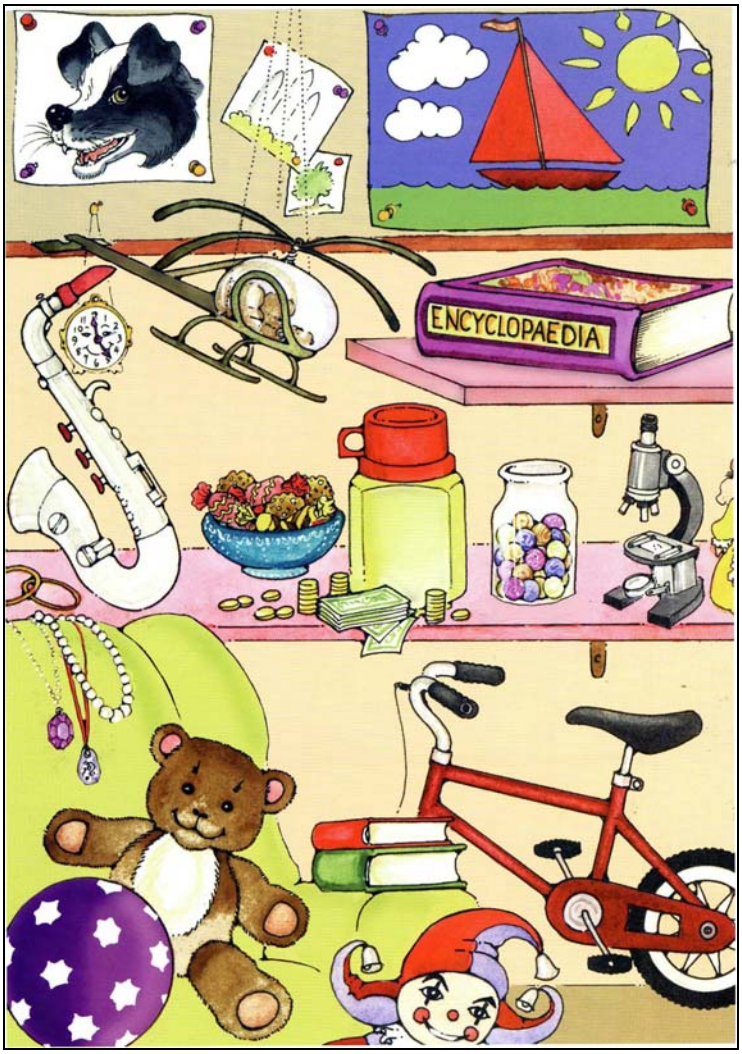

Figure 5.17: Picture 3

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.

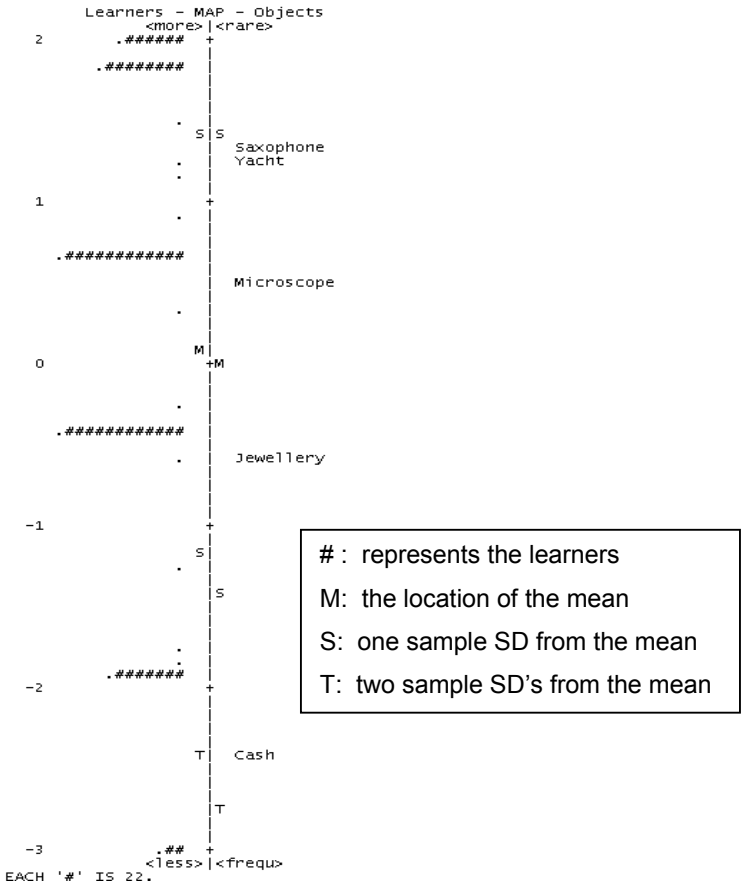

Figure 5.18: Item-Learner Map for all languages– Picture 2

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.

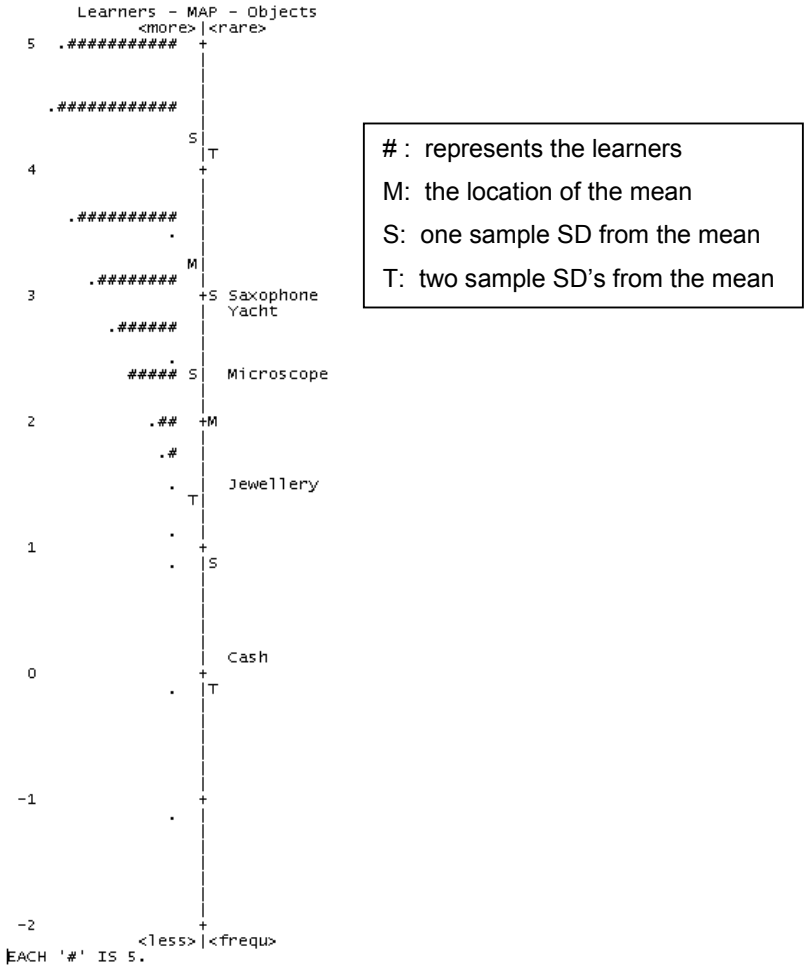

Figure 5.19: Item-Learner Map for Picture 3 - Afrikaans group

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

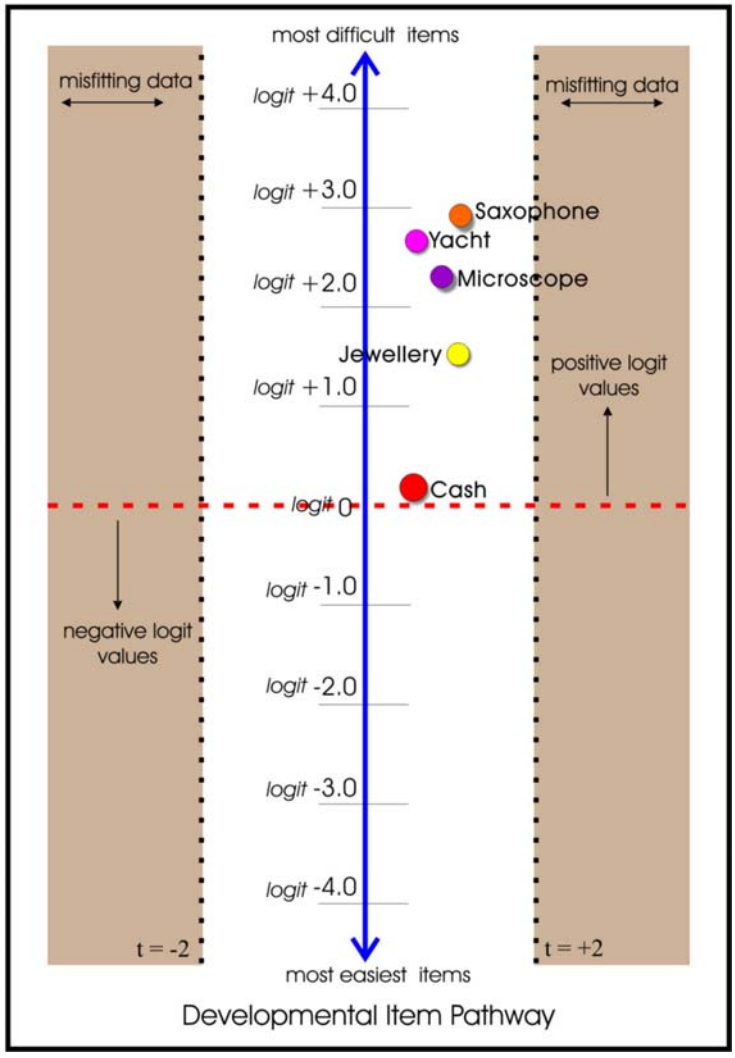

Figure 5.20: IDP for Afrikaans group – Picture 3

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.

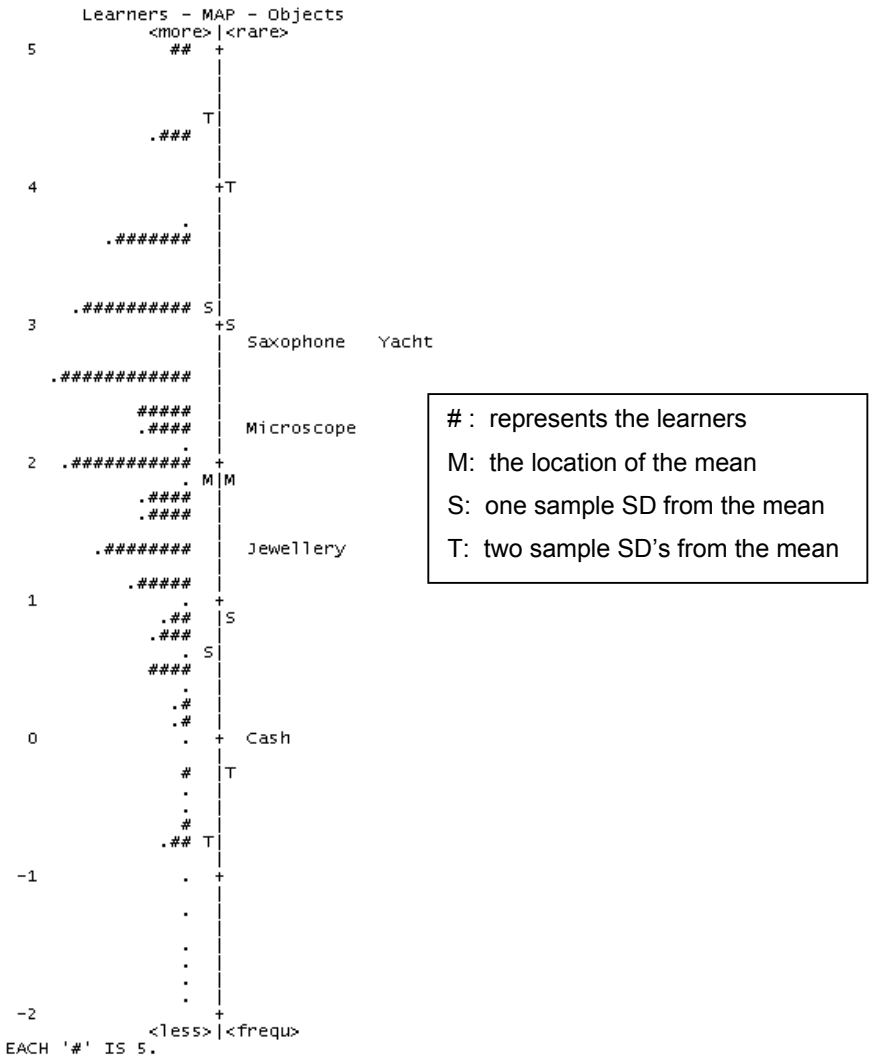

Figure 5.21: Item-Learner Map for Picture 3 - English group

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

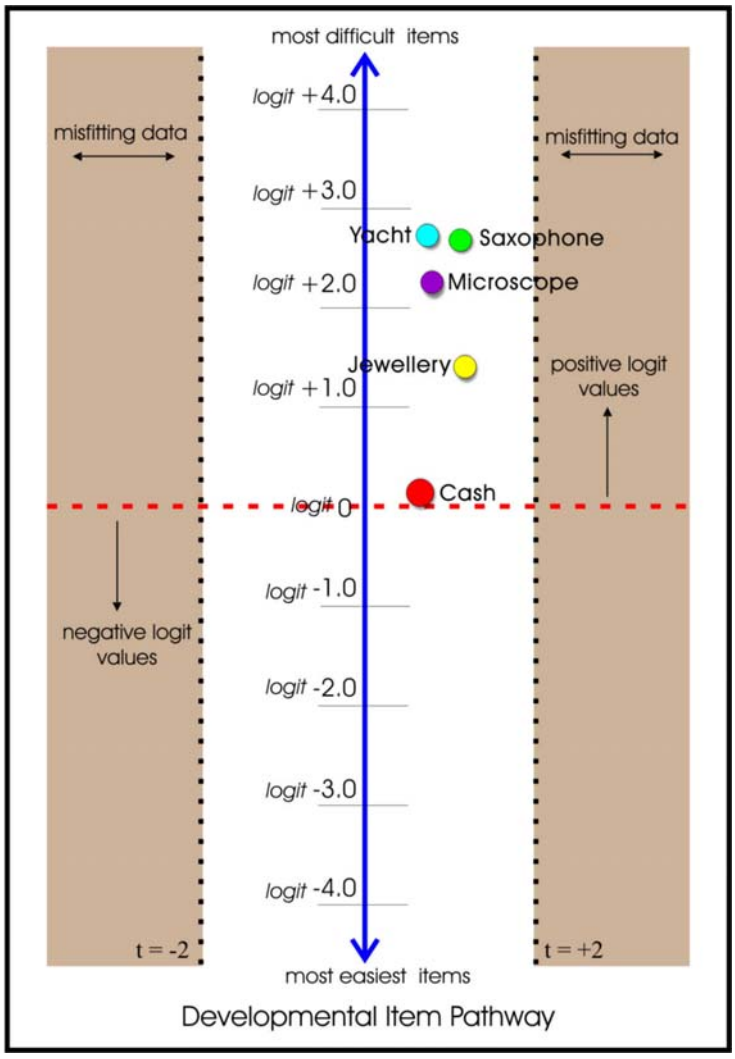

Figure 5.22: IDP for English group – Picture 3

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**.

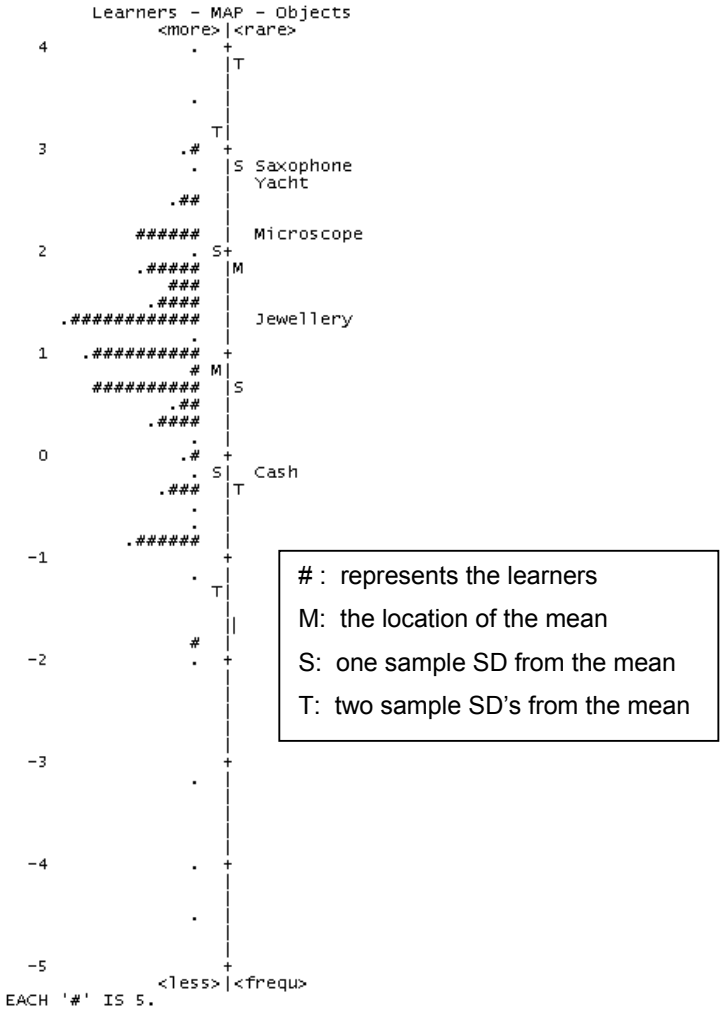

Figure 5.23: Item-Learner Map for Picture 3 - Sepedi group

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

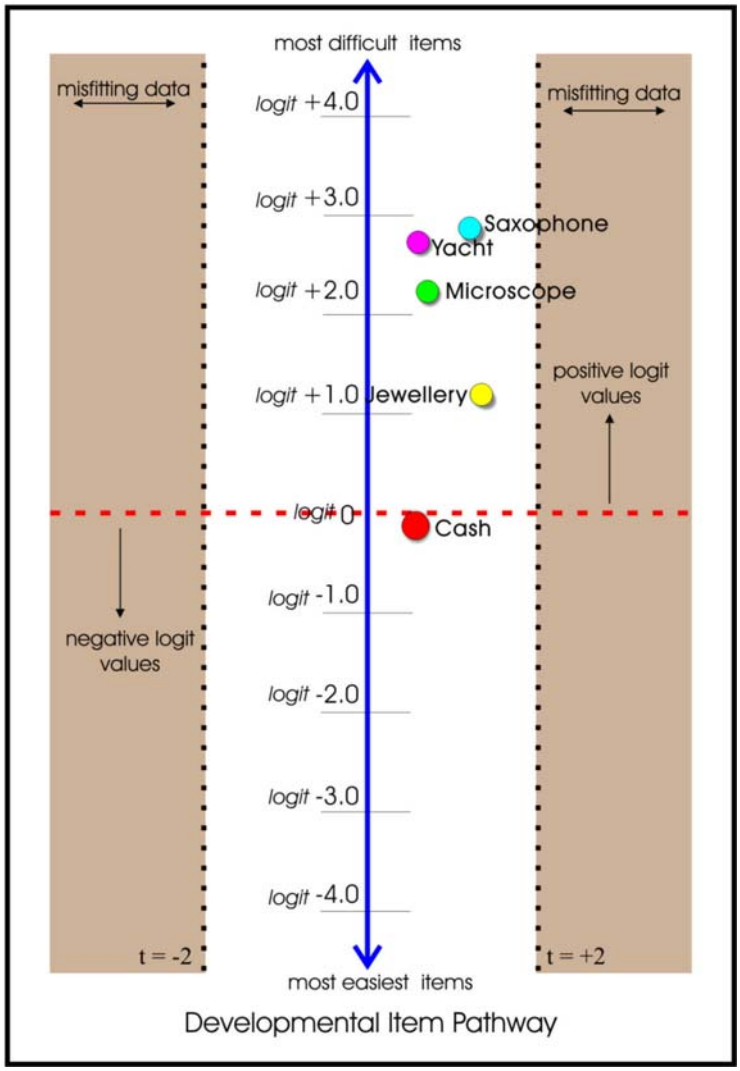

Figure 5.24: IDP for Sepedi group – Picture 3

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.

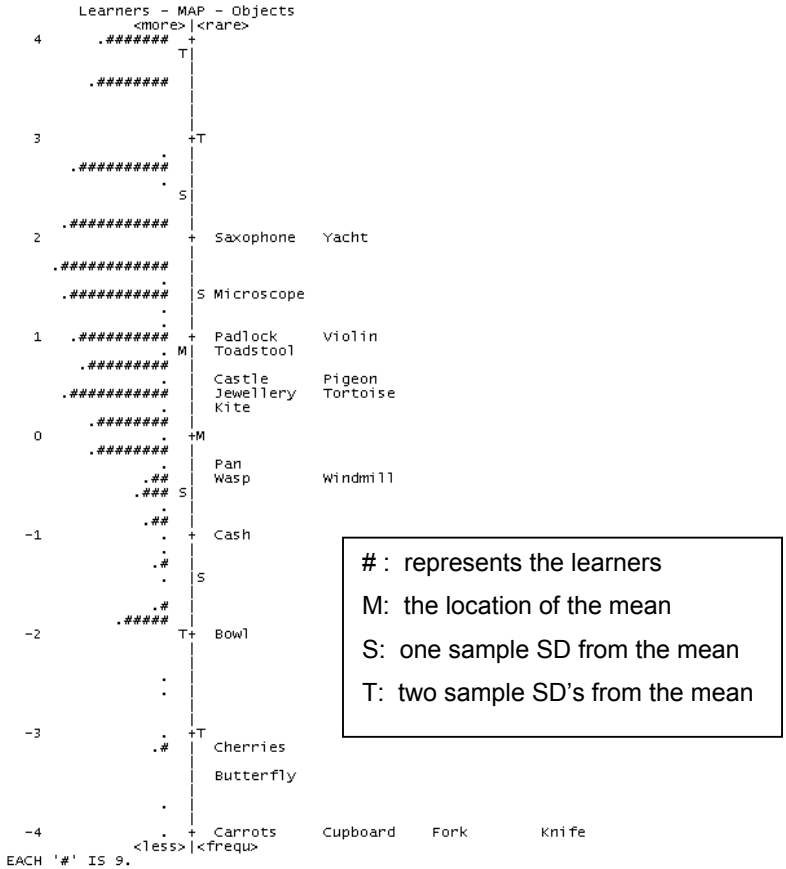

Figure 5.25: Item-Learner Map for Picture 3 for all languages

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).

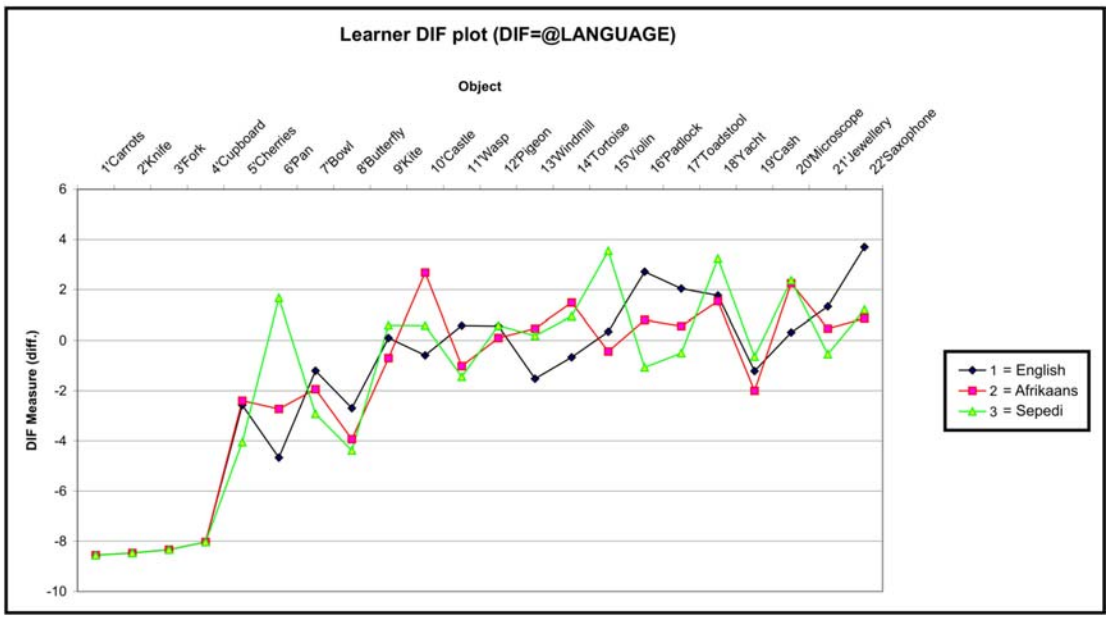

Figure 5.26: Differential Item Functioning Graph

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.

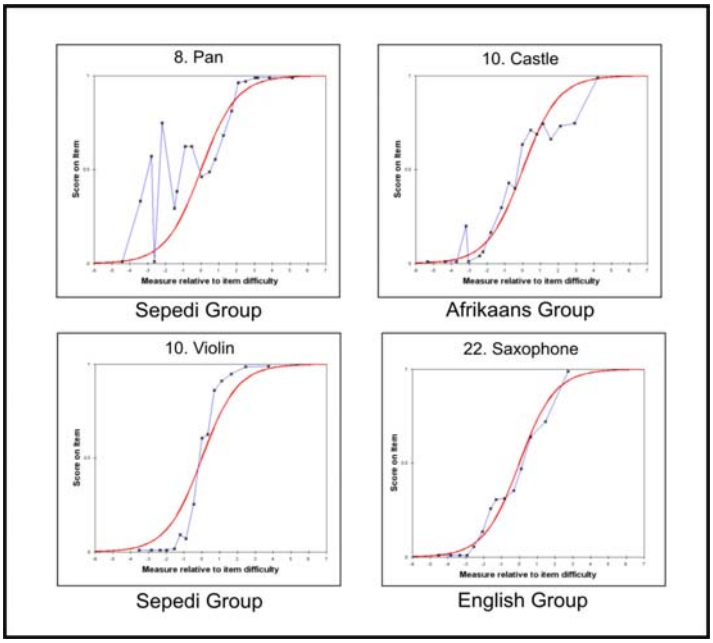

Figure 5.27: Item Characteristic Curves

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)

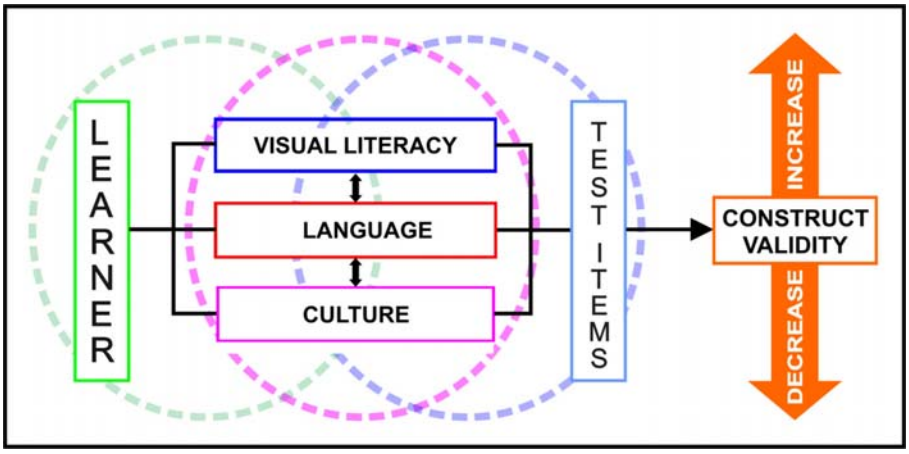

Figure 6.1: Original Conceptual Framework

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:

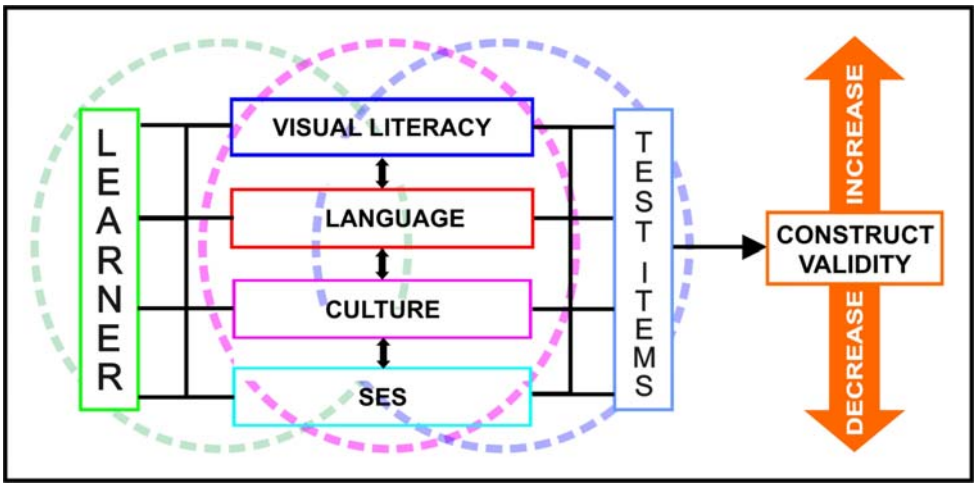

Figure 6.2: Adapted Conceptual Framework

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.