EXAMINING THE EQUIVALENCE OF THE PIRLS 2016 RELEASED TEXTS IN SOUTH AFRICA ACROSS THREE LANGUAGES

by

KAREN ROUX

A thesis submitted in fulfilment of the requirements of the degree

PHILOSOPHIAE DOCTORIS

in

Assessment and Quality Assurance

FACULTY OF EDUCATION

UNIVERSITY OF PRETORIA

Supervisor: Dr S. van Staden

Co-Supervisor: Prof E. J. Pretorius

2020

DECLARATION

I declare that the thesis, which I hereby submit for the degree PhD Assessment and Quality Assurance in Education and Training at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

.....

Karen Roux

28 September 2020

ETHICAL CLEARANCE CERTIFICATE

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Faculty of Education

RESEARCH ETHICS COMMITTEE

CLEARANCE CERTIFICATE	CLEARANCE NUMBER: **SM 17/07/04**
DEGREE AND PROJECT	PhD Examining the equivalence of the PIRLS 2016 released texts in South Africa across three languages
INVESTIGATOR	Ms Karen Roux
DEPARTMENT	Science, Mathematics and Technology Education
APPROVAL TO COMMENCE STUDY	17 August 2017
DATE OF CLEARANCE CERTIFICATE	28 September 2020

CHAIRPERSON OF ETHICS COMMITTEE: Prof Funke Omidire

CC

Ms Bronwynne Swarts
Dr Surette van Staden
Prof Elizabeth Pretorius

This Ethics Clearance Certificate should be read in conjunction with the Integrated Declaration Form (D08) which specifies details regarding:

- Compliance with approved research protocol,
- No significant changes,
- Informed consent/assent,
- Adverse experience or undue risk,
- Registered title, and
- Data storage requirements.

ETHICS STATEMENT

The author, whose name appears on the title page of this thesis, has obtained, for the research described in this work, the applicable research ethics approval. The author declares that he/she has observed the ethical standards required in terms of the University of Pretoria's Code of ethics for researchers and the Policy guidelines for responsible research.

-⌘-

DEDICATION

This doctoral study is dedicated to all the staff at the Centre for Evaluation and Assessment who, for the last three cycles, have tirelessly conducted the Progress in International Reading Literacy Study in South Africa.

-⌘-

ACKNOWLEDGEMENTS

When I first started my academic career and dreamt about one day obtaining my PhD, I thought to myself that it would be an incredible challenge which would require commitment and a solid support structure, but little did I realise just how difficult it would be in practice. It took many late nights, with a few early mornings sprinkled in between, often fuelled by caffeine and sheer grit, but at long last I managed to reach the finishing line. I can honestly say that this document, into which I poured my heart and soul, would not have been possible without the assistance of those people who always been there for me. Each and every one of you means the world to me.

First and foremost, to my loving husband, Edrick, my soulmate and inspiration who has always been there for me. I would not be who I am without you. Every moment and every day you were there, telling me that I can do this and that there is nothing standing in my way, taking care of my every need and even satisfying my unknown need to obtain a crystal unicorn. Shade of my heart, I cannot imagine having done this without you by my side. With all my heart and soul, I love you.

To my mother Susan, for always believing in me no matter how much I doubted myself. No matter how many times I felt as though I would not be able to take the mounting pressure, I knew that I could always count on you. I love you so much.

To Dricky, my mother-in-law, thank you so much for all the coffee and support. The simple act of asking how things were going meant far more than you can imagine.

To Surette, my friend and mentor, I cannot express in words how grateful I am to have met you all those years ago. Your unwavering support and guidance were, and undoubtedly will continue to be, incredibly important to me. You have been my inspiration on so many levels and I hope that you know just how amazing you are.

To Lilli, I can honestly say that this study would not have gone nearly as well without your extensive knowledge of languages and our long discussions over coffee. Your wisdom, insight and sincere belief in me helped me to find my way whenever I stumbled.

To Cilla, my friend and editor, thank you so much for your assistance with the language editing. I cannot imagine where I would be right now if our paths had not crossed during my Master's degree.

To everybody at the CEA, each of you has been a huge part of my life and in many ways, this study would never have realised had I never met you. Thank you all for the chats and the support you have always shown me.

To the various panellists who assisted me with the study, your patience and comments helped shaped this study into what it is today, I am truly grateful for all the assistance which you have given me.

Last but not by any means the least, to Marcus, a friend who has been a part of my team since I can remember, who has always supported me unconditionally and given me gifts ranging from unicorns to Space Cows to help keep me positive as time wore on, thank you so much for always being there for me.

In closing, I wish to leave you all with the poem “Oh Captain! My Captain!” by Walt Whitman, a poem that is very dear to me and Surette.

-⌘-

O Captain! My Captain!

By Walt Whitman

O Captain! my Captain! our fearful trip is done,
The ship has weather'd every rack, the prize we sought is won,
The port is near, the bells I hear, the people all exulting,
While follow eyes the steady keel, the vessel grim and daring;
But O heart! heart! heart!
O the bleeding drops of red,
Where on the deck my Captain lies,
Fallen cold and dead.

O Captain! my Captain! rise up and hear the bells;
Rise up—for you the flag is flung—for you the bugle trills,
For you bouquets and ribbon'd wreaths—for you the shores a-crowding,
For you they call, the swaying mass, their eager faces turning;
Here Captain! dear father!
This arm beneath your head!
It is some dream that on the deck,
You've fallen cold and dead.

My Captain does not answer, his lips are pale and still,
My father does not feel my arm, he has no pulse nor will,
The ship is anchor'd safe and sound, its voyage closed and done,
From fearful trip the victor ship comes in with object won;
Exult O shores, and ring O bells!
But I with mournful tread,
Walk the deck my Captain lies,
Fallen cold and dead.

Whitman, W. (1891). *Leaves of Grass: O Captain My Captain (pp. 262-263).* Philadelphia: David McKay.

ABSTRACT

The Progress in International Reading Literacy Study (PIRLS) is a large-scale reading comprehension assessment, which assesses Grade 4 learners' reading literacy achievement. The findings from the last cycle of PIRLS 2016 indicated that South African Grade 4 and 5 learners performed poorly in reading comprehension. This finding confirms the previous cycles' results where South African learners achieved the lowest results across the participating countries. Approximately eight out of ten Grade 4 learners cannot read for meaning in any of the tested languages. Due to the poor results in PIRLS, the President of South Africa stated that every ten-year old child should be able to read for meaning, thus cementing reading literacy as a national aim. The aim of this mixed methods research was to determine whether the PIRLS Literacy 2016 and PIRLS 2016 limited release texts are equivalent across languages, specifically English, Afrikaans and isiZulu.

Four research sub-questions were explored to assist in addressing the main research question posed by this study: To what extent are the PIRLS 2016 released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent? As this study took the form of a sequential explanatory mixed methods approach, the first phase investigated the South African Grade 4 and 5 results by firstly looking at descriptive statistics, such as percentages and means. After the initial exploration of the data, I conducted Rasch analyses to determine whether the items from the limited release texts showed measurement invariance – in other words, whether the items behaved differently for different groups of learners. As part of the Rasch analyses, individual item-fit statistics and differential item functioning (DIF) were conducted using RUMM2030. In phase two, the limited release texts were analysed by experts who attended workshops and completed open-ended questionnaires regarding the equivalence of the identified texts. The qualitative phase was conducted in order to complement and extend on the quantitative findings of phase one.

The findings revealed that the limited release texts, with their accompanying items, were not equivalent across the different languages. However, by looking at the items that displayed DIF, there is not a clear pattern as the items did not universally favour one language nor did the texts discriminate universally against a particular language. An in-depth look at the texts and items themselves revealed that the *Flowers on the Roof* text is considered the poorest translation into Afrikaans and isiZulu. Overall, all the texts were considered to be appropriate for South African

learners as the texts made use of rich vocabulary and introduced the learners to new ideas and concepts. Thus, this study offers new insights into the equivalence of the PIRLS assessments as well as possible reasons for the non-equivalence for each of the limited release texts. Based on the findings of this study, recommendations and further research are provided.

-⌘-

KEY WORDS

- Cultural Equivalence
- Differential Item Functioning
- Education
- Functional Equivalence
- Linguistic Equivalence
- Metric Equivalence
- PIRLS 2016
- PIRLS Literacy 2016
- Rasch Analysis
- Rasch Measurement Theory
- Reading Comprehension
- Reading Literacy
- Secondary Analysis
- Translation

-⌘-

LANGUAGE EDITOR CERTIFICATE

EDITING SERVICES

To whom it may concern

This letter serves to confirm that editing and proofreading was done for:

Karen Roux

Faculty of Education
University of Pretoria

Doctoral Study:

Examining the Equivalence of the PIRLS 2016 Released Texts in South Africa across Three Languages

Supervisor: Dr S. van Staden
Co-Supervisor: Prof E. J. Pretorius

Cilla Dowse
21 September 2020

Cilla Dowse
PhD in Assessment and Quality Assurance in Education and Training: University of Pretoria 2014
Programme on Editing Principles and Practices: University of Pretoria 2009
Basic Editing and Proofreading: McGillivray Linnegar Associates 2008

Rosedale Farm
P.O. Box 48
Van Reenen
Free State
cilla.dowse@gmail.com
Cell: 084 900 7837

-⌘-

LIST OF ACRONYMS

AERA	American Educational Research Association
ANOVA	Analysis of Variance
BCM	Black Consciousness Movement
BICS	Basic Interpersonal Communication Skills
BNC	British National Corpus
C2005	Curriculum 2005
CALP	Cognitive Academic Language Proficiency
CAPS	Curriculum and Assessment Policy Statement
CEA	Centre for Evaluation and Assessment
COCA	Corpus of Contemporary American English
CR	Constructed Response
DBE	Department of Basic Education
DIF	Differential Item Functioning
DoE	Department of Education
DPC	Data Processing Centre
ETS	Educational Testing Service
FAL	First Additional Language
GETC	General Education and Training Certificate
HL	Home Language
ICC	Item Characteristic Curve
IDB Analyzer	International Database Analyzer
IEA	International Association for the Evaluation of Educational Achievement
IIAL	Incremental Introduction of African Languages
ILSA	International Large-Scale Assessments
IQCM	International Quality Control Monitor
IRT	Item Response Theory
ISC	International Study Centre
ISCED	International Standard of Classification of Education
ITC	International Test Commission
LiEP	Language in Education Policy
LoLT	Language of Learning and Teaching
MC	Multiple Choice

NAF	National Adaptation Forms
NCS	National Curriculum Statement
NEPA	National Education Policy Act
NGO	Non-Government Organisation
NP	National Party
NQF	National Qualifications Framework
NRC	National Research Co-ordinator
NSC	National Senior Certificate
OBE	Outcomes-Based Education
PanSALB	Pan South African Language Board
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
QDG	Questionnaire Development Group
QL	Qualitative
QN	Quantitative
RA	Rasch Analysis
RDG	Reading Development Group
RE	Reading Ease
RNCS	Revised National Curriculum Statement
RSA	Republic of South Africa
RUMM	Rasch Unidimensional Measurement Model
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SAL	Second Additional Language
SASA	South African Schools Act
SE	Standard Error
SES	Socio-Economic Status
SGB	School Governing Body
SL	Source Language
SOP	Survey Operations Procedures
ST	Source Text
TIMSS	Trends in Mathematics and Science Study
TL	Target Language
TT	Target Text
UK	United Kingdom

UNESCO	United Nations Educational, Scientific and Cultural Organisation
US	United States

-⌘-

GLOSSARY

Agglutinative Language	The isiZulu language is part of the Nguni language group and has an agglutinative structure as the orthography includes long, complex word units, that is, a root word may include several morphemes
Coherence	Examines the semantic connection of a text and whether it is accessible and relevant
DIF	Used to determine whether test items function differently for different test-takers, that is, whether there is bias in person measurement
IDB-Analyzer	A plug-in for the Statistical Package for the Social Sciences, used to combine and analyse IEA data
PIRLS	International comparative (cross-national) study assessing Grade 4 learners' reading comprehension

-⌘-

TABLE OF CONTENTS

CHAPTER 2

PROGRESS IN INTERNATIONAL READING LITERACY STUDY

CHAPTER 3

A REVIEW OF THE LITERATURE AND CONCEPTUAL FRAMEWORK

CHAPTER 5

PHASE ONE: PIRLS LITERACY 2016 GRADE 4 QUANTITATIVE DATA PRESENTATION AND RESULTS

CHAPTER 6

PHASE ONE: PIRLS 2016 GRADE 5 QUANTITATIVE DATA PRESENTATION AND RESULTS

CHAPTER 7

PHASE TWO: LITERARY TEXTS QUALITATIVE DATA PRESENTATION AND RESULTS

CHAPTER 8

PHASE TWO: INFORMATIONAL TEXTS QUALITATIVE DATA PRESENTATION AND RESULTS

-⌘-

LIST OF TABLES

-⌘-

LIST OF FIGURES

-⌘-

LIST OF APPENDICES

The appendices, due to their size, can be located on the CD.

- Appendix A: PIRLS Literacy 2016 and PIRLS 2016 Overall Results
- Appendix B: PIRLS Literacy 2016 and PIRLS 2016 Overall Results by Language
- Appendix C: PIRLS Literacy 2016 and PIRLS 2016 Released Texts Scoring Guide (available at http://timssandpirls.bc.edu/pirls2016/international-results/pirls/appendices/)
- Appendix D: The Pearl (Afrikaans and isiZulu versions)
- Appendix E: Flowers on the Roof (Afrikaans and isiZulu versions)
- Appendix F: Macy and the Red Hen (Afrikaans and isiZulu versions)
- Appendix G: African Rhinos and the Oxpecker Birds (Afrikaans and isiZulu versions)
- Appendix H: The Green Sea Turtle's Journey of a Lifetime (Afrikaans and isiZulu versions)
- Appendix I: Permission letters from the PIRLS national research centre and the Centre for Evaluation and Assessment (CEA), to use the PIRLS Literacy 2016 and PIRLS 2016 data and released texts for secondary analysis

-⌘-

CHAPTER 1

Introduction and Overview of the Study

1.1 ORIENTATION

The Progress in International Reading Study (PIRLS) is an international comparative study begun in 2001 with the aim of assessing learners' reading achievement. The study consists of assessment instruments comprised of literary and informational texts as well as national curriculum policies in reading and background questionnaires, which give an indication of home environment for learning, school climate and resources, and how instruction occurs in classrooms (Mullis & Martin, 2015).

The aim of this study is to examine the equivalence of the identified released PIRLS Literacy 2016 and PIRLS 2016 assessment instruments in the South African context, specifically the English, Afrikaans and isiZulu texts and items. This chapter presents the introduction to the foundation of the study as well as the structure of the thesis. To begin with, the background to the study (Section 1.2), and then the rationale and problem statement of the study (Section 1.3) are presented. Thereafter, the study's research questions are outlined (Section 1.4) followed by an overview of the research methodology used in this study (Section 1.5). The operational definitions of key terminology used in the study can be found in Section 1.6. Finally, the content outline for the remaining chapters is presented in Section 1.7.

1.2 BACKGROUND TO THE STUDY

The purpose of this study is to examine the equivalence of the PIRLS 2016 released assessment instruments used in the South African study. Several reading achievement tests have been conducted in South Africa and the results often indicate large disparities in reading achievement between language groups. South Africa has participated in three cycles of the Progress in International Reading Literacy Study (PIRLS) which assessed learner reading literacy across the 11 official languages. PIRLS is aimed at assessing learner reading ability through both literary (narrative or fiction) and informational passages (factual or information-based articles). Previous cycles of PIRLS (2006 and 2011) have indicated that South African learners lag behind their peers, as illustrated on the international scale, including that of other developing countries. It appears that there is a lack of growth in South African learners' literacy

levels, and other large-scale assessments, such as the Trends in Mathematics and Science Study (TIMSS), also show significant disparities in achievement, which are indicative of the effect of South Africa's socio- and political history on education (Howie et al., 2012; Reddy, 2006; Spaull, 2015a; van der Berg, 2008; van der Berg, et al., 2016).

Many factors contribute to disparities in literacy performance; one important factor could be the lack of parity in the tools used to assess literacy performance. Learner performance is dependent on many aspects, for example, the familiarity to different types of texts, vocabulary, cultural experiences as well as their own behaviours and attitudes. When international assessments are conducted, a key aspect is to ensure that the source text (ST) and the translated text (TT) are equivalent, meaning that they measure the same construct (Arffman, 2013). As PIRLS is a large international and cross-cultural study, it requires translation from a source language (SL) to a target language (TL), usually from English to the TL of participating countries.

The translation process focuses on linguistic equivalence (Peña, 2007); however, it is necessary to look at other aspects such as cultural, functional and metric equivalence (*cf*. Chapter Three). This study focuses on the equivalence of the PIRLS Literacy 2016 and PIRLS 2016 limited released literary and informational texts used in the South African study. The PIRLS results give an indication of where different countries sit on the same measurement bar. By analysing the selected PIRLS texts, the study might provide some clarification as to why South African learners are struggling to attain good literacy levels even when assessed in a home language.

1.2.1 South African Language and Education Context

The social landscape of South Africa includes multilingual citizens in a multicultural society. During the 1990s, South Africa underwent political reform and established the current Constitution, in 1996, that recognises eleven official languages[1]. The 2011 Census data indicate that there are 57,1 million South African citizens. The largest group speaks isiZulu (22%) followed by isiXhosa (16%) and Afrikaans (13%). Less than 10% of the population speak the remaining eight languages, respectively (StatsSA, 2012) with English only being spoken by 9%.

[1] Afrikaans, English, isiNdebele, isiXhosa, isiZulu, Sepedi, Sesotho, Setswana, siSwati, Tshivenda and Xitsonga (languages are alphabetically ordered).

The South African Constitution of 1996 explicitly states that all children in South Africa have the right to be educated in the language of their choosing. In line with this, the Department of Basic Education (DBE) drafted a new policy for language education, namely the Language-in-Education Policy (LiEP), which is guided by the South African Schools Act (act 84 of 1996). LiEP underpins the language of learning in schools in order to promote home language instruction, specifically from Grades 1 to 3. Even though LiEP stipulates that home language instruction be the norm, there are various difficulties in implementing the policy. As Pretorius (2008) explains, the Language of Learning and Teaching (LoLT) is not necessarily the home language or first language spoken at home and this could help explain the reading literacy problem, as not all learners receive adequate teaching in their home language.

As a consequence of implementing a new Constitution, the educational curriculum has undergone several reforms to address the inequalities of the past (Howie et al., 2017). The first curriculum introduced was Curriculum 2005 which was separated into three policy documents, for the Foundation Phase, Intermediate Phase and the Senior Phase. This curriculum was revised in 2001 and released in the public domain as the Revised National Curriculum Statement (RNCS) in 2002 (DBE, 2012). As a result of shortcomings in the RNCS, the Curriculum and Assessment Policy Statement (CAPS) was created, not as a new curriculum replacing the old, but rather as an amendment to the National Curriculum Statement (NCS). The CAPS was implemented in 2012 to learners in the Foundation Phase (Grades R – 3) and Grade 10 learners. The CAPS was rolled out to the Intermediate Phase and Grade 11 in 2013, and for the Senior Phase and Grade 12, in 2014.

Taking into account the multilingual nature of South Africa, the CAPS encourages learners to become proficient in at least two languages and adopts an additive approach to multilingualism, whereby “the home language is strengthened and affirmed while any further language learning is seen as adding value” (DBE, 2011, p. 105). This policy directive means that all learners may learn in their home language and then learn at least one additional language of their choosing. As such, depending on the location of the school, different languages are offered as the LoLT. However, in practice, learners who enter schools with a specific LoLT, such as one of the African languages, change to English or Afrikaans as the LoLT from Grade 4 onwards (Pretorius, 2014). The reason for either switching to English or Afrikaans as a LoLT in the Intermediate Phase, is that Grade 12 learners write their final school leaving examination in either of the aforementioned languages (Taylor & von Fintel, 2016). Moreover, these are two

languages have well-developed academic registers with infrastructure and resources that enables these languages to be used in education.

1.2.1.1 Disparity in Literacy Levels in South Africa

South Africa is a multicultural and multilingual country entrenched in high poverty with deep inequalities due to apartheid, which creates a volatile educational environment and perpetuates literacy inequalities (Spaull, 2015b). This education environment has led to uneven literacy levels within the schooling context as well as unequal educational opportunities for learners coming from diverse SES backgrounds (Graven, 2014). The inequality in education has been investigated by various scholars (*cf.* Spaull & Kotze, 2015; van Dyk & White, 2019; van der Berg et al., 2016; van der Berg & Gustafsson, 2019), particularly focusing on access to quality education, learner performance as well as availability of school and classroom resources. Since 1994, the South African government has tried to uplift poorer public schools by categorising schools according to different quintile levels and adjusting the funding from government to mostly assist poorer schools. The quintile system is based on the location of the school by looking at the school's socio-economic status (SES) (Ogbonnaya & Awuah, 2019). Quintile one schools are the poorest quintile while quintile five is the least poor. Quintile one to three schools receive the most funding because they experience high poverty contexts where many learners attend schools with pit toilets, temporary classrooms and extremely large class sizes as result of apartheid, where most of the funding was allocated to schools attended by white learners (Dass & Rinquest, 2017). The quintile system seeks to rectify the inequalities by introducing no-fee schools, and more funding to quintile one to three schools. As such, it is important to ensure that large-scale assessments, used for heterogeneous groups of learners are equivalent, valid and fair.

1.2.2 Overview of PIRLS Literacy 2016 and PIRLS 2016 Results

The Progress in International Reading literacy Study (PIRLS) is an international comparative evaluation of Grade 4 reading literacy which specifically focuses on reading comprehension. PIRLS is conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA) which undertakes several large-scale comparative studies that benchmark performance of learners in reading, mathematics, science, civic education as well as in information, communication and technology. The IEA conducts the aforementioned studies to obtain a greater understanding of the effects of policies and practices of educational

systems, providing participating countries with linkages between the intended, implemented and attained curricula by means of collecting background information through questionnaires (Mullis & Martin, 2015).

The PIRLS Reading Assessment Framework, focusing on the processes of comprehension, purposes for reading, and reading behaviours and attitudes towards reading, underpins the assessment of learners' reading achievement (Mullis & Martin, 2015). By utilising assessment instruments and questionnaires, PIRLS provides multi-level data of reading literacy skills and aims to describe trends in reading literacy across the globe of learner reading achievement (Mullis & Martin, 2015).

1.2.3 Threats to the Validity of Test Instruments

Validity, viewed as a key aspect of assessments (AERA et al., 2014), is defined as the extent to which a test score is appropriate for intended interpretation and use of the test (Kane, 2013). To evaluate the validity of test scores, validity evidence should be gathered to show appropriateness of the interpretation and use of the test instruments (Wools, 2015). Taylor (2013) describes validity in assessment as "evaluating logical arguments and empirical evidence to determine whether they support proposed inferences from, as well interpretations and uses of, assessment results" (p. 2). Table 1.1 depicts the three forms of validity which the study explored:

Table 1.1: Types of Validity

Type of Validity	Description
Face	This is understood as a "subjective and cursory judgement of a concept, assessment instrument or any other conceptualisation to ascertain whether, on its face, it appears valid" (Mostert, 2007, p. 380). It is also a form of internal validity (Gaber, 2010).
Content	This "…is used as a theoretical concept that focuses on the extent to which an assessment instrument shows evidence of fair and comprehensive coverage of the domain of items that it supposed to cover" (Babbie, 2007, p. 2). Content validity looks at the following: • Relevance to the curriculum;

Type of Validity	Description
	• Focus on what was taught; • Comprehensive content coverage; • The proportion of the scope of learning; and • Sampled potential content (Babbie, 2007, p. 2).
Construct	This refers to the notion that "the appropriateness, meaningfulness, and usefulness of score-based inferences are inseparable and that the unifying force behind this integration is the trustworthiness of empirically grounded score interpretation" (Messick, 1989, p. 8). Construct validity can be established in several ways: • the measure is able to reflect clients' developmental changes; • the measure correlates with old measures that have proven construct validity; • the underlying dimensions or traits of the construct have been identified using factor analysis; • the measure's internal consistency has been demonstrated using statistical techniques such as biserial correlations; • the measure has convergent and discriminant validity; and • the construct is tested experimentally to see whether supposed interventions alter the construct in the hypothesized direction (Anastasi, 1988, pp. 147-150).

An examination of the translation procedures followed in PIRLS Literacy 2016 and PIRLS 2016 may shed some light on the reasons for poor performance, facilitate a deeper understanding of the quality of the translation process in a large-scale assessment as well as identify possible reasons why learners performed poorly on certain texts or items. In order to examine the possible threats to validity of the PIRLS instruments, this study considers face validity and Peña's four concerns for cross-cultural assessments, namely linguistic, functional, cultural and metric equivalence (discussed in Chapter Three).

Equivalence of cross-cultural assessments refer to the similarity and comparability of an assessment across different language and culture groups and ensuring that learners have equal opportunities to demonstrate their abilities (Peña, 2007). This study focuses on macro, meso

and micro levels of equivalence in order to assist in determining whether the PIRLS texts are equivalent in terms of Peña's four concerns (equivalences) to safeguard against validity threats against cross-cultural assessments.

During the PIRLS assessments, learners were tested in the language of learning and teaching used during the Foundation Phase, i.e. Grades 1-3 (Howie et al., 2017); for example, Grade 4 learners were assessed in all official languages for PIRLS Literacy 2016, whereas Grade 5 learners were assessed only in English, Afrikaans and isiZulu in PIRLS 2016 (Howie et al., 2017a; Howie et al., 2017b). For some learners, the home language (HL) and language of learning and teaching (LoLT) are not necessarily the same. Consequently, these learners' educational experiences are affected by current educational policy which states that HL education be given in the Foundation Phase, with a First Additional Language (FAL) (usually English) introduced as a subject. A switch is made from HL to English or Afrikaans as it becomes the LoLT from the Intermediate Phase (Grade 4) onwards.

1.3 RATIONALE AND PROBLEM STATEMENT

This section reports on the study's rationale and aim in Section 1.3.1 and is followed by a discussion of the potential contribution of this study (Section 1.3.2).

1.3.1 Rationale and Research Aim

Since its first participation in PIRLS in 2006, South Africa's learners have continuously performed poorly with most of the learners in PIRLS 2006, 2011 and 2016 not even reaching the low benchmark, which describes basic reading skills and strategies. Several factors play a role in learner assessment, particularly in multilingual education systems, and include possible construct and content validity issues as well as translation validity. The intention of this study is to establish the equivalence of PIRLS Literacy 2016 and PIRLS 2016 assessment instruments used in the South Africa study. This study examined the English, Afrikaans and isiZulu results on PIRLS Literacy 2016 and PIRLS 2016. The reason for selecting these three languages are twofold: PIRLS 2016 only tested these three languages due to poor performance during previous rounds of PIRLS (*cf*. Howie et al., 2008; 2012); isiZulu is part of the Nguni language group, the largest spoken language in South Africa, and there is an interest to see whether there has been improvement over the last ten years.

By focusing on instrument validity, this study plans to determine whether there is equivalence

across languages in terms of language translations at macro, meso and micro levels of equivalence. It is possible that the texts, which were translated from the source language (SL) to the target language (TL), are not equivalent. As a result, there might be differential functioning on items indicating possible non-equivalence between the SL and TL. It then begs the question whether PIRLS Literacy 2016 and PIRLS 2016 are equivalent for South African learners.

1.3.2 Potential Contribution of the Research

Given South Africa's legacy of inequality, the results release after each PIRLS cycle is controversial and provides opportunity for speculation about the equivalence of achievement tests between languages. Such speculation not only comes from different role players in the South African education sector, but is also often intuitive or based on small scale evidence from other studies and inadvertently becomes the accepted truth as it applies to PIRLS data too. As such, it is important to examine ILSAs such as PIRLS on the issues of equivalence.

This study takes the form of a secondary analysis of the PIRLS Literacy 2016 and PIRLS 2016 English, Afrikaans and isiZulu results. The potential contribution of this study is that there are various factors associated with reading literacy achievement and this study aims to ascertain if the PIRLS Literacy 2016 and PIRLS 2016 released texts and items were equivalent across English, Afrikaans and isiZulu, using both quantitative and qualitative methods.

By examining these assessments instruments, this study could add to knowledge and understanding of test development, specifically focusing on their equivalence, and informing national policy and curriculum by expanding on the current body of knowledge of cross-cultural research. It may give insights into how reading comprehension assessments are developed and how equivalence is established across different languages in South Africa. Taking the above into account, an in-depth analysis of the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments was conducted to establish whether these assessments were equivalent across English, Afrikaans and isiZulu.

1.4 MAIN RESEARCH QUESTION AND SUB-QUESTIONS

This study intends to determine the extent to which the PIRLS assessment instruments are equivalent for South African Grade 4 and 5 English, Afrikaans and isiZulu learners by

examining the limited[2] released PIRLS Literacy 2016 and PIRLS 2016 texts. The PIRLS Literacy 2016 released texts include *The Pearl* and *African Rhinos and the Oxpecker Birds*. The PIRLS 2016 released texts comprise *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime*. One text, namely *Flowers on the Roof*, was included in both PIRLS assessments. The main research question of this study asks:

To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?

Five subsidiary questions, which assist in answering the main question, were formulated as such:

1. *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?*
2. *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016?*
3. *To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?*
4. *How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?*
5. *How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?*

The findings of the abovementioned sub-questions will be discussed in Chapters Five to Eight in an attempt to address the main research question of establishing whether the PIRLS Literacy 2016 and PIRLS 2016 texts and items are equivalent across the three languages.

1.5 OVERVIEW OF THE RESEARCH METHODOLOGY

This study explored equivalence in cross-cultural testing by focusing on the South African English, Afrikaans and isiZulu learners' results of the PIRLS Literacy 2016 and PIRLS 2016 assessments. A sequential explanatory mixed methods design was utilised to explore the South African learners' results on PIRLS Literacy 2016 and PIRLS 2016. This specific mixed methods design entailed two phases: first, a quantitative data collection and analysis phase

[2] Limited released texts refer to the texts made available by the IEA after each cycle of assessment.

followed by the second, qualitative data collection and analysis phase (Creswell, 2013; Creswell & Plano Clark, 2011).

Phase One included a secondary analysis of the PIRLS Literacy 2016 and PIRLS 2016 South African results, particularly across the three languages, namely English, Afrikaans and isiZulu. This phase included descriptive analysis and Rasch analysis. The results from the first phase informed the second phase. Phase Two involved a rigorous and in-depth qualitative look at the South African Grade 4 and 5 results by conducting workshops, evaluating the texts across each language and examining learner responses to constructed response items. It should be noted that during Phase Two, only items that presented differential item functioning (DIF) were explored.

1.6 DEFINITIONS OF KEY TERMS FOR THE STUDY

As part of contextualising this study, it is important to give meaning and understanding to certain terms used within this study. These key terms are used throughout the thesis.

1.6.1 Reading Literacy

For the purposes of this study, a distinction is made between literacy and reading literacy. In general, literacy is considered as a person's ability to engage in activities in which literacy is required to function in society; and to use literacy skills to read and write. As this study draws from the international large-scale study, PIRLS, the definition of reading literacy is understood as:

> Reading literacy is the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They can read to learn, to participate in communities of readers in school and everyday life, and for enjoyment (Mullis & Martin, 2015, p.12).

The PIRLS definition of reading reflects other main theories of reading literacy as a constructive and interactive process (*cf.* Anderson & Pearson, 1984; Kintsch, 2013). Authors such as Britt et al. (2012) have explained that meaning is developed through interaction between a text and a reader, one who actively constructs meaning from a text and is able to use reading strategies (Afflerbach & Cho, 2009) and a repertoire of cognitive strategies and linguistic skills to construct meaning (Baker & Beall, 2009).

1.6.2 Face Validity

Face validity entails an examination of the measure and "whether 'on the face of it'" (Vogt, 2005, p. 117) appears to measure what it intended to measure. In determining face validity, an expert or judge is asked to evaluate the measure and make a decision whether the measure appears to be valid (Martinez, 2017; Maruyama & Ryan, 2014). In addition, face validity entails the degree to which the purpose of the measure is clear to the test-taker (Bornstein, 2004). It is also considered as a form of internal validity (Gaber, 2010).

1.6.3 Equivalence of Test Instruments

This study draws from Peña's (2007) considerations for cross-cultural research, particularly in translations of text, namely linguistic, functional, metric and cultural equivalence. Each of these equivalences are regarded as possible threats to the validity of an assessment and should be considered when conducting international large-scale studies. Within this study, the umbrella term 'translation equivalence' is considered the overarching term for the four equivalences as identified by Peña (2007). As this study pays attention to each of these equivalences, it is important for the reader to be familiar with each type of equivalence. These concepts of equivalence are discussed in detail in Chapter Three, but briefly described in the sections below.

1.6.3.1 Linguistic Equivalence

Linguistic equivalence refers to translating instructions and instruments and checking the translation using the back-translation method (Chesterman, 2016). Whenever an International Large-Scale Assessment (ILSA) is conducted, the researchers and other relevant stakeholders, such as the participating countries' representatives, should consider how the assessments will be translated, bearing in mind that the source text (ST) should be linguistically equivalent to that of the target text (TT). When ensuring linguistic equivalence, the translators employed by the researchers of the study, usually make use of translation and a back-translation technique (Beck et al., 2003). A translator translates the original text into the target language (TL) producing a target text (TT); thereafter a second translator translates the TT back into the source language (SL) (Behr, 2017). The goal of linguistic equivalence is to ensure that words, sentences and phrases used in the ST and TT, as well as the linguistic meaning thereof, are the same (Peña, 2007).

1.6.3.2 Functional Equivalence

Functional equivalence indicates that the instructions and instrument produce the same behaviour (Greenfield et al., 2006). The goal of functional equivalence is to ensure that the translation from the ST to the TT is 'natural' (Bermann & Porter, 2014) – in other words, the ST and TT should, from a measurement perspective, behave similarly. When translating a text, the translators should keep in mind the target group of the TT and ensure that there is similarity between the ST and TT. Therefore, when the text is being translated, the translator should keep in mind the receptor of the TT – meaning that the TT should be understood by the target group as if they were the original receptors reading the source text (Nida & Taber, 1969[3]). Functional equivalence can highlight possible incongruence between the ST and TT.

1.6.3.3 Cultural Equivalence

Cultural equivalence considers how respondents interpret a given direction or test item that taps into the same cultural meaning for each cultural linguistic group (Chan & So, 2017). In essence, the aim of cultural equivalence is to ensure that the meaning of the construct, whether it is a piece of text, instructions or items, remains the same across different cultures and languages (Peña, 2007). When translating a ST into a TT, the translator need to keep in mind how the receptors may perceive the TT, as their own knowledge and understanding may differ due to cultural and linguistic backgrounds. Ensuring cultural equivalence might be difficult to achieve, as each culture may perceive certain aspects differently. Therefore, when ILSAs are conducted each participating country should scrutinise the construct that they wish to measure, whether it be a test, questionnaire or other form of data collection, to carefully consider the cultural implications.

1.6.3.4 Metric Equivalence

Metric equivalence refers to the difficulty of the specific item expressed in two distinct languages (Kim et al., 2003). As such, the goal of metric equivalence is to ensure that the items used in a test or assessment are the same in terms of difficulty when conducted in multiple languages (Peña, 2007). ILSAs usually develop the test instruments in English which are then translated into the languages of the participating countries. When the test items are translated,

[3] Nida's work on functional equivalence is still upheld today (*cf.* Bermann & Porter, 2014; Chunhua, 2014; Fengling, 2017).

it is crucial that the source items and the target items' difficulty remain the same. Therefore, an item developed in English and translated into a different language, should have the same level of difficulty.

This study will utilise the abovementioned equivalences to determine whether the PIRLS Literacy 2016 and PIRLS 2016 texts are equivalent between English, Afrikaans and isiZulu learners.

1.6.4 Macro, Meso and Micro Levels of Equivalence

In order to unpack the PIRLS texts in terms of equivalence as described in Section 1.6.3, a distinction was made between macro, meso and micro levels of equivalence, and similarities and/or differences in features at these levels were examined across the three languages. The macro level of equivalence includes an overview of each PIRLS text by providing information such as the genre, themes, purpose as well as the layout of the text and text length. The macro level links with functional equivalence as the macro level focuses on the characteristics of the texts. In terms of meso level of equivalence, it includes passage mapping for each text and the different processes of comprehension tested. As such, this level of equivalence relates to functional, metric and linguistic equivalences. The micro level looks at the vocabulary used in each text, the literary or rhetorical devices used, organisational elements as well as readability of the texts. The micro level relates to linguistic and cultural equivalence.

1.6.5 PIRLS Learner Performance

The PIRLS Literacy 2016 and PIRLS 2016 studies report learner results as an average score with the standard error (Mullis & Martin, 2015). The PIRLS study made use of Item Response Theory (IRT) methods to summarise the achievement for the learners on the *international centrepoint* with a mean of 500 and standard deviation of 100. The PIRLS scale ranges from 0 to 1 000. In order to assist with understanding the PIRLS scale and to determine what the learners can do in terms of reading literacy, PIRLS utilises benchmarks. There are four benchmarks on this scale, namely: Low International Benchmark (400-474 score points); Intermediate International Benchmark (475-549 score points); High International Benchmark (550-625 score points; and the Advanced International Benchmark (625 and above score points). Within this study, reference will be made to South African Grade 4 and 5 learners' reading achievement scores in relation to the international average.

1.7 CHAPTER DELINEATION FOR THE STUDY

The first chapter, *Chapter One*, aimed to provide a general outline and summary of the reasons why this study has been undertaken. It described the background, rationale as well as the potential contribution of the study. This chapter also included the study's research questions and key terms which will guide the thesis.

Chapter Two discusses the Progress in International Reading Literacy (PIRLS) 2016 study, specifically the background and development thereof. It provides an overview of the PIRLS Literacy 2016 and PIRLS 2016 Reading Assessment Framework, which includes the purposes of reading, the contexts for learning to read as well as the comprehension processes assessed. In addition, this chapter describes the two data collection tools used, namely achievement instruments and contextual questionnaires. Furthermore, it describes the data collection, analysis procedures as well as methodological norms. More importantly, the translation procedures of PIRLS Literacy and PIRLS 2016 are discussed.

Chapter Three comprises two major sections, namely the literature review and the conceptual framework. Firstly, the chapter presents the literature that will contextualise the study within the current South African education framework. The chapter offers an overview of the South African education system in terms of its history and its drive to improve education. Educational and other national policies for development are deliberated upon. The chapter also focuses on different forms of validity as they equate with equivalence; the chapter specifically looks at the linguistic, metric, functional and cultural equivalence of the PIRLS Literacy 2016 and PIRLS 2016 texts and items.

In the subsequent chapter, *Chapter Four*, the research design and methodology for the study is discussed. At the outset, the chapter discusses the methodological underpinnings of this study which is followed by the research design, sampling, data collection and analysis. The chapter also provides information about the ethical considerations for the study.

Chapter Five involves the presentation and discussion of the study's Phase One quantitative research findings for the first sub-question of the study. The first sub-question takes the form of a secondary analysis of the PIRLS Literacy 2016 learner achievement data, by examining the three released texts and items. It is the intent of this chapter to present the South African Grade 4 learner achievement via percentages and means across English, Afrikaans and isiZulu. In order to assist in addressing the first sub-question, Rasch analysis, specifically Differential

Item Functioning (DIF) is used to determine whether the PIRLS Literacy 2016 selected items show non-equivalence between the languages. It also serves as an indicator of possible issues to examine during the study's qualitative second phase.

Chapter Six follows a similar format and structure to *Chapter Five*. This chapter presents the findings the second sub-question of the study. It focuses on the PIRLS 2016 Grade 5 English, Afrikaans and isiZulu learner achievement data by investigating the three released texts and accompanying items. This chapter presents descriptive statistics, specifically means and percentages, as well as Rasch analysis. By using both aforementioned statistical techniques, it serves as an indicator of possible issues per item that is should be further explored during the second qualitative phase of this study.

Chapter Seven presents the qualitative findings for the third and fourth sub-questions by reporting on and examining workshop notes, questionnaires, learner responses, text analysis and readability analysis to determine whether the selected PIRLS Literacy 2016 and PIRLS 2016 literary (narrative) texts are equivalent across the languages. Content analysis was employed to determine the texts' face validity.

Chapter Eight is similar in structure and reporting to *Chapter Seven*; it presents the qualitative findings for the third and fifth sub-questions. However, this chapter focuses on the two informational (expository) released texts. The goal of this chapter was to determine whether the PIRLS Literacy 2016 and PIRLS 2016 informational texts are the same for English, Afrikaans and isiZulu learners.

The final chapter, *Chapter Nine*, consists of several sections that aim to give a global view and understanding of the findings of the study. It summarises the findings of the main research question for the study followed by reflections on the specific research methodology used in the study. Further reflections are offered on the conceptual framework used in this study. The chapter also draws conclusions based on the results of the study and discusses possible implications for policy and practice. Lastly, the chapter includes recommendations for future research.

-⌘-

CHAPTER 2

Progress in International Reading Literacy Study

2.1 ORIENTATION

As already stated, the Progress in International Reading Literacy Study (PIRLS) is an international comparative study first inaugurated in 2001. The study is conducted in five-year cycles which enables participating countries to monitor trends across the years. PIRLS assesses learner reading literacy achievement in the fourth year of schooling, which is a pivotal point as learners transition from '*learning to read*' to '*reading to learn*' (Mullis & Martin, 2015). South Africa, one of 61 participating systems[4], participated in PIRLS 2016 (*cf.* Table 2.1).

Table 2.1: Countries Participating in PIRLS 2016

Participating Countries			**Benchmarking Participants** [a]
Australia	Germany	Norway (5)	Buenos Aires, Argentina
Austria	Hong Kong SAR	Oman	Ontario, Canada
Azerbaijan	Hungary	Poland	Quebec, Canada
Bahrain	Iran, Islamic Rep. of	Portugal	Denmark (3)
Belgium (Flemish)	Ireland	Qatar	Norway (4)
Belgium (French)	Israel	Russian Federation	Moscow City, Russian
Bulgaria	Italy	Saudi Arabia	Federation
Canada	Kazakhstan	Singapore	Eng/Afr/Zulu - RSA (5)[5]
Chile	Kuwait	Slovak Republic	Andalusia, Spain
Chinese Taipei	Latvia	Slovenia	Madrid, Spain
Czech Republic	Lithuania	South Africa	Abu Dhabi, UAE
Denmark	Macao SAR	Spain	Dubai, UAE
Egypt	Malta	Sweden	
England	Morocco	Trinidad and Tobago	
Finland	Netherlands	United Arab Emirates	
France	New Zealand	United States	
Georgia	Northern Ireland		

[a] Benchmarking participants include education systems or sub-population groups within countries.

[4] Of the 61 participating systems, 50 were participating countries and 11 benchmarking entities.
[5] The Republic of South Africa (RSA) benchmarked at the fifth grade with schools where learners have instruction in English, Afrikaans, or isiZulu.

This chapter presents the details of PIRLS 2016, beginning with an overview of the International Association for the Evaluation of Educational Achievement (IEA) (2.2) outlining the history of the IEA and PIRLS. Section 2.3 presents a discussion on the Reading Framework that underpins the PIRLS 2016 study looking as aspects of reading. In the subsequent section (2.4), the contexts for learning to read is presented as they form the basis of the questionnaire framework. The PIRLS assessment instruments are described in Section 2.5. The translation processes followed during the PIRLS 2016 study is explored in Section 2.6 and this is followed by a description of data collection, monitoring and scoring (2.7). The data processing of PIRLS 2016 is discussed in Section 2.8. The ethical considerations of PIRLS 2016 can be found in Section 2.9 with the last section (2.10) giving concluding comments.

2.2 THE INTERNATIONAL ASSOCIATION FOR THE EVALUATION OF EDUCATIONAL ACHIEVEMENT AND PIRLS

The International Association for the Evaluation of Education Achievement (IEA) is an international non-government organisation (NGO) that has been undertaking large-scale international studies since 1958 (IEA, 2018). Over the years, the IEA membership has grown from 12 to 62 educational research institutes which represent different countries. The IEA's permanent secretariat is based in the Netherlands whereas the Data Processing Centre (DPC) is located in Germany, the International Study Center (ISC) is located at Boston College, USA and the sampling unit, StatsCan, is located in Canada (Mullis et al., 2017). The IEA undertakes several large-scale comparative studies that benchmark performance of learners in reading, mathematics, science, civic education as well as in information, communication and technology. It conducts the aforementioned studies to obtain a greater understanding of the effects of policies and practices of educational systems (Mullis et al., 2007). Moreover, the IEA provides participating countries with information about linkages between the intended, implemented and attained curricula by means of collecting background information through questionnaires (IEA, 2011).

The Progress in International Reading Literacy Study (PIRLS) is one of the large-scale assessments conducted by the IEA. PIRLS provides multi-level data of learners' reading literacy skills by utilising assessment instruments and questionnaires. It aims to describe global trends in reading literacy of learners' reading achievement (Mullis & Martin, 2015).

PIRLS was first undertaken in South Africa in 2006 and the country has since participated in two further cycles (2011, 2016). PIRLS 2006 tested South African Grade 4 and 5 learners in all 11 official languages and they achieved poor results, which indicated that the learners struggled to read for comprehension (*cf.* Howie et al., 2008). Grade 4 learners achieved an average reading literacy score of 253 (SE=4.6) whilst the Grade 5 learners obtained 302 (SE=4.6) – both scores are lower than the lowest PIRLS benchmark. During the second cycle of PIRLS, conducted in South Africa in 2011, a decision was made to test Grade 5 learners on the PIRLS assessment and the Grade 4 learners on an easier version of the assessment called prePIRLS. The Grade 5 learners obtained an average of 421 (SE=7.3) while Grade 4 learners achieved an average of 323 (SE=4.3). The most recent cycle of PIRLS (2016) revealed that the Grade 5 learners reached an average achievement score of 406 (SE=6.0) and the Grade 4 learners who participated in PIRLS Literacy[6], achieved an average score of 320 (SE=4.4).

The PIRLS 2016 reading framework underpins the development of the PIRLS assessment instruments, and is presented in the sections below.

2.3 PIRLS 2016 READING FRAMEWORK

The PIRLS Reading Framework is based on two aspects, namely reading purposes and comprehension processes which underpin the PIRLS and PIRLS Literacy assessments. It is designed to reflect the IEA's vision to incorporate current and relevant approaches to measuring learners' competence in reading literacy (Mullis & Martin, 2015). As part of the reading framework, it is important to explain reading literacy as the PIRLS assessments measure learners' reading literacy competence. Section 2.3.1 deals with what comprises reading literacy and is followed by the PIRLS reading aspects (2.3.2).

2.3.1 What is Reading Literacy?

During the IEA's Reading Literacy Study in 1991, which served as a foundation for PIRLS, reading literacy was defined as "the ability to understand and use those written language forms required by society and/or valued by the individual" (Elley, 1992, p.17). The PIRLS and PIRLS Literacy studies have over the years refined their definition of reading literacy. In the latest PIRLS assessment framework, reading literacy is defined as

[6] prePIRLS was renamed as PIRLS Literacy.

> … the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment (Mullis & Martin, 2015, p.12).

2.3.2 Aspects of Reading

The two aspects of reading described in the PIRLS reading framework include the *Purposes for Reading* (discussed in sub-section 2.3.2.1) and the *Processes of Comprehension* (discussed in sub-section 2.3.2.2).

2.3.2.1 Purposes for Reading

In any setting, people read for various reasons, which could range from reading for fun to reading for learning. There are two predominant purposes for reading that account for the reading done by learners in general; these are for *literary experience* and to *acquire and use information* (Mullis et al., 2017).

Young readers usually read for fun or about a specific interest they have and these types of reading usually focus on a narrative type text that has a story, or an informational text that includes facts about a specific topic. When learners engage with a literary passage or text, they become involved in several aspects of that particular text, such as the setting, characters, atmosphere and plot. The main source of literary text for the PIRLS and PIRLS Literacy assessments include narrative fiction, as it is applicable in an international context (Mullis et al., 2017). Poetry is excluded from the PIRLS and PIRLS Literacy assessments given the multitude of languages and the difficulty of translation. When poetry is translated from the source language to another language, it becomes difficult to keep the wording, rhyme and the essence of the poem may get lost during translation.

Learners also get involved in reading informational texts that inform readers about things in their environment, or more broadly, for example, about the life cycle of a butterfly. With time, learners start to sharpen their literacy skills and read more often to acquire information about things from books or online sources (Duke & Carlisle, 2011; Wharton-McDonald & Swiger, 2009).

The PIRLS Literacy and PIRLS assessments are designed to assess the learners' reading literacy skills by using texts that are fun to read and which comply with the two main purposes for reading. Table 2.2 shows the *Purposes for Reading* and the weight devoted to each of the purposes by study.

Table 2.2: The Reading Purposes by Study

Purposes for Reading	PIRLS Literacy	PIRLS
Literary Experience	50%	50%
Acquire and Use Information	50%	50%

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p.14), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

The *Purposes for Reading* for the PIRLS and PIRLS Literacy assessments are equally divided between reading for literary experience and to acquire and use information. Each literary text has a theme followed by unique items that focus on the overall theme, events that have occurred throughout the story, characters and setting. The informational texts have items that are formed around the information that is contained in the text (Mullis & Martin, 2015).

2.3.2.2 Processes of Comprehension

Within the PIRLS reading framework, it is understood that readers construct meaning in various ways and as such, the framework has four broad *Processes of Comprehension*: focus on and retrieve explicitly stated information; make straightforward inferences; interpret and integrate ideas and information; and evaluate and critique content and textual elements (Mullis & Martin, 2015). During a reading exercise, learners make use of their metacognitive processes and strategies which enable them to examine their understanding of the text and adjust their reading approach accordingly (Baker & Beall, 2009; Kintsch & Kintsch, 2005). Learners are also able to draw on their own background experience which equips them with the necessary skills to engage with the different texts (Kintsch, 2013).

The four *Comprehension Processes* form the foundation for the development of reading comprehension items. Each text in the PIRLS and PIRLS Literacy assessments has between 13 and 15 items that assess each comprehension process. By using an array of different items, it is possible to gauge the learners' reading abilities and skills by examining their answers. Items

are specifically designed to keep in mind the comprehension processes as well as the length and complexity of the text.

There are slight differences in the emphasis across the *Processes of Comprehension* for the PIRLS Literacy and PIRLS studies as shown in Table 2.3 below.

Table 2.3: The Comprehension Processes by Study

Processes of Comprehension	PIRLS Literacy	PIRLS
Focus on and Retrieve Explicitly Stated Information	50%	20%
Make Straightforward Inferences	25%	30%
Interpret and Integrate Ideas and Information	25%	30%
Evaluate and Critique Content and Textual Elements		20%

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p.14), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

As the PIRLS Literacy assessment is an easier assessment, the emphasis on the processes is somewhat different to the PIRLS assessment. Half of the PIRLS Literacy assessment covers the *Focus on and Retrieve Explicitly Stated Information* process whereas this process constitutes one-fifth of the PIRLS assessment. A description of each of the *Processes of Comprehension* is given below.

Focus on and Retrieve Explicitly Stated Information

When readers focus on and retrieve explicitly stated information, they use different ways to locate the information and understand the content in order to answer a question. The question that tests this process requires the reader to focus on the particular text, specifically paying attention to words, expressions, phrases or full sentences for the purpose of constructing meaning. Additionally, this process may also require the reader to use different pieces of a text or gather information across the text. This process has very little or no inferring or interpreting from the text (Mullis & Martin, 2015). Based on the PIRLS 2016 assessment framework, the following are different types of focus on and retrieve processes:

- Identifying information that is relevant to the specific goal of reading;
- Looking for specific ideas;
- Searching for definitions of words or phrases;

- Identifying the setting of a story (for example, time and place); and
- Finding the topic sentence or main idea (when explicitly stated) (Mullis & Martin, 2015, p. 19).

Make Straightforward Inferences

For this process, readers make inferences based on what they have read in the text, such as making straightforward inferences about information that is not explicitly stated in the text. This process enables readers to move beyond the surface of the text in order to solve the 'gaps' in meaning. Straightforward inferences are usually based on information contained in the text and requires the reader to connect two or more pieces of information. It should be noted that the ideas or information may be explicitly stated, but the connection between these ideas or information is not. As such, the reader must infer the connection between the two ideas or information. Even though the inference is not explicitly stated in the text, the meaning thereof is understood. Readers who are deemed to be skilled, are able to connect two or more pieces of information and, more importantly, recognise the relationship despite the fact that it is not stated in the text (Mullis & Martin, 2015). Below are the reading tasks that exemplify this process of comprehension:

- Inferring that one event caused another event;
- Concluding what is the main point made by a series of arguments;
- Identifying generalisations made in the text; and
- Describing the relationship between two characters (Mullis & Martin, 2015, p. 20).

Interpret and Integrate Ideas and Information

Similar to that of making straightforward inferences, readers who interpret and integrate ideas and information within a text focus on either local or global meanings. This process involves the reader making sense of what the author's intent was when writing the text as well as to develop a comprehensive understanding of the text. When readers read a text and start to interpret and integrate, it means that these readers are attempting to construct either a specific or more complete understanding of the text. Readers do this by integrating their own personal knowledge and experience with meaning that resides in the text. Therefore, these readers are able to draw from their own understanding of the world, which includes their background knowledge and experience, in order to make connections that are implicit and open to

interpretation. However, the meaning that is constructed through this specific process may vary between readers as meaning-making is dependent on the readers' knowledge and experience (Mullis & Martin, 2015). The reading tasks that exemplify this process of comprehension include:

- Discerning the overall message or theme of a text;
- Considering an alternative to actions of characters;
- Comparing and contrasting text information;
- Inferring a story's mood or tone; and
- Interpreting a real-world application of text information (Mullis & Martin, 2015, p. 21).

Evaluate and Critique Content and Textual Elements

With this process, readers are expected to evaluate and critique elements of a text, meaning that the reader moves from constructing meaning to critically evaluating the text itself. The readers, whilst reading the selected text, are able to take a step back from the text and critique it. During this process, readers evaluate and critique what they are reading either from a personal or objective viewpoint. When readers are engaged in this process, they draw from their language knowledge and skill as well as specific features of the text. Consequently, these readers make judgements drawn from their own knowledge and experience. These judgements may include the author's choices for conveying meaning, such as literary and rhetorical devices. Readers who are equipped with good reading literacy skills are able to evaluate visual as well as textual features of a text and explain their functions (Mullis & Martin, 2015). Reading tasks that exemplify this process of comprehension include:

- Judging the completeness or clarity of information in the text;
- Evaluating the likelihood that the events described could really happen;
- Evaluating how likely an author's argument would be to change what people think and do;
- Judging how well the title of the text reflects the main theme;
- Describing the effect of language features, such as metaphors or tone; and
- Determining an author's perspective on the central topic (Mullis & Martin, 2015, p. 22).

In light of the above descriptions of the four *Processes of Comprehension*, Mullis and Martin (2015) provide the following explanation to keep in mind when investigating the different processes:

> Initially, it may seem that locating and extracting explicitly stated information would be less difficult than, for example, making interpretations across an entire text and integrating those with external ideas and experiences. However, all texts are not equal and can vary with regard to length, syntactic complexity, abstractness of ideas, and organizational structure. Thus, the nature of the text can impact the difficulty of the question asked, across and within the four types of comprehension processes (p.18).

To develop background information on participating countries' context, questionnaires, pertaining to the school, home, class as well as the national context, underpinned by a questionnaire framework are administered to participants. The PIRLS questionnaire framework is described in the next section.

2.4 PIRLS 2016 CONTEXT QUESTIONNAIRE FRAMEWORK

The PIRLS Literacy and PIRLS studies also collect extensive background information about the learners' home, school and classroom contexts, as well as learner characteristics and their attitudes towards reading (Hooper, Mullis & Martin, 2015), which provides valuable information on the educational systems across different countries and how these may be improved. Background information is collected through contextual questionnaires as well as from the PIRLS 2016 Encyclopedia (*cf.* Howie et al., 2017). Even though the current study does not focus on the contextual questionnaires, it is nevertheless important to describe these questionnaires in order to provide a comprehensive overview of PIRLS.

The participating countries were able to collect PIRLS background information as the learners, their parents/guardians, teachers and school principals completed questionnaires which provide data about the learners' home and school contexts. Each questionnaire covers an array of questions and is an indispensable component of the PIRLS and PIRLS Literacy studies. The questionnaires may provide policy-relevant information about each participating country's reading contexts (Mullis & Martin, 2015). The PIRLS and PIRLS Literacy studies consider the learners' reading literacy achievement in conjunction with the contexts for learning to read by

collecting data on the home support for learning, educational structures, school organisation, curricula, teacher education as well as the teachers' classroom practices (Mullis & Martin, 2015). Figure 2.1 illustrates the PIRLS contexts for learning framework.

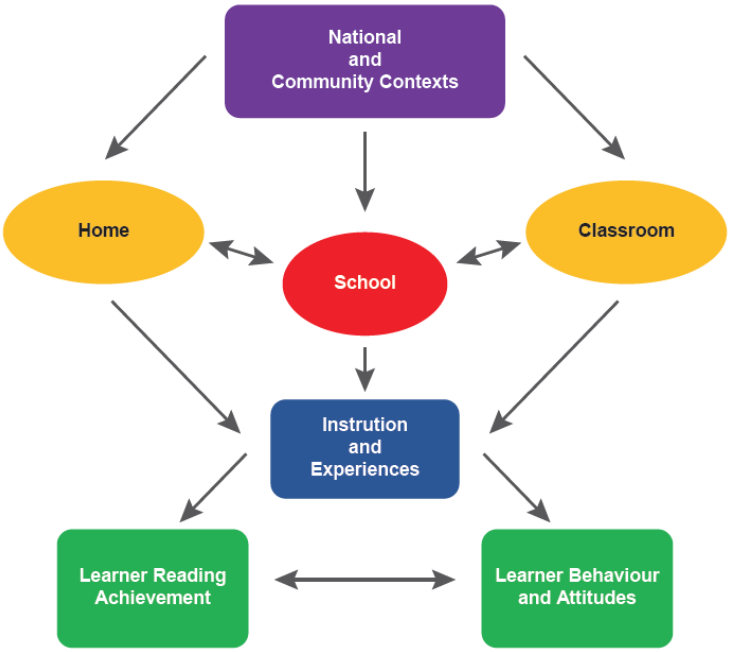

Figure 2.1: PIRLS 2016 Conceptual Framework (Mullis et al., 2009 in Howie et al., 2017)

In the following sub-sections, the PIRLS contexts for learning to read are unpacked: National and Community Context (2.4.1); Home Context (2.4.2); School Context (2.4.3); Classroom Context (2.4.4) and Learner Characteristics and Attitudes toward Learning (2.4.5).

2.4.1 The National and Community Context

Mullis and Martin (2015) explain that there are several factors that contribute to learners' reading literacy acquisition, such as cultural, social, political and economic factors. Important policy decisions about education, are made at the national and community level, on how to efficiently implement the curriculum in a country, given the abovementioned factors. There are several national characteristics and decisions that should be taken into consideration when discussing effective reading literacy (*cf*. Table 2.4).

Table 2.4: Factors within the National and Community Contexts

Factors	Description
Language and emphasis on literacy	History of language and literacy in a country;

Factors	Description
	multilingual approaches; oral literacies
Economic resources, population demographics, and geographic characteristics	The country's levels of wealth and distribution thereof; population size and diversity; geographic differences
Organisation and structure of the educational system	Centralised or decentralised educational system; policy decisions
Learner flow	Learners' educational progress through school; pre-primary education; age of entry; grade retention; grouping for reading instruction
Reading curriculum in the primary grades	Rigour of the reading curriculum; policies about reading literacy; materials and methods used for reading instruction
Teachers and teacher education	Implementation of curriculum; teacher training, specifically in content and pedagogical approaches; professional development
Monitoring curriculum implementation	Monitoring and evaluating systems in place; standardised testing; school inspections; audits; teacher and learner feedback

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p. 32-36), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

The information on contexts is drawn from the learners' home context, the school context and the classroom context, each of which is described below.

2.4.2 The Home Context

Nationally and internationally much research has been conducted to provide insight into the learners' home environment and how important it is for reading literacy development (*cf.* Ackermann, 2017; Roux 2014; UNESCO, 2017). The home context is made up of various factors that may have an impact on the learners' reading literacy development, and as such, the

PIRLS study collected contextual data by means of the *Learning to Read Survey*[7] as well as the *Learner Questionnaire* administered to participants in the PIRLS Literacy and PIRLS studies. Table 2.5 indicates the five main factors, with a description of each, offering information on the home context.

Table 2.5: Factors within the Home Context

Factors	Description
Home resources for learning	Important proxy for socio-economic status; parent education; parent occupation; home resources
Language(s) spoken in the home	Learning to read is reliant on a learner's language experience and language spoken at home
Parental educational expectations and academic socialisation	Parents express their educational expectations for their children; parents and children engage in academic socialisation as it may have long term effects on learners' performance
Early literacy activities and early numeracy activities	Parental involvement in early literacy activities with child; reading aloud with young children; oral language important for literacy acquisition
Home reading support	Intervention in reading; tutoring; other supplementary schooling; excel on specific exams or subjects

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p. 37-40), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

Moving with the times, there are other factors, such as the availability of eBooks and other new media devices, which might have an effect on the learners' reading literacy development. In a previous cycle of PIRLS, Mullis et al. (2009) explained that the home environment should be the focus of activities that encourage the acquisition of reading literacy skills. Leu et al. (2009)

[7] Also known as the Home Questionnaire.

raised the concern that learners who do not have access to e-resources and media devices might be disadvantaged, which would then have an effect on them in the classroom and school context.

2.4.3 The School Context

Valuable information regarding the school context is gathered during the PIRLS study as most formal education occurs within the school environment. Research has found that the school environment can influence not only achieving the school's curricular aims and objectives but also learner performance (*cf*. Mullis et al., 2017; Zimmerman, 2017). Each school can be seen as an integrated system where each aspect or action taken may affect a part of the whole school. The following table depicts the specific school aspects explored during the PIRLS study.

Table 2.6: Factors within the School Context

Factors	Description
School location (urban, peri-urban, rural)	Some schools in urban areas appear to have more resources and support
School composition by learner socio-economic background	Socio-economic status (SES) of school can influence learner achievement; other factors may also explain the correlation between SES and learner achievement
Instruction affected by resource shortages	Quality school resources are essential for quality teaching; general and subject-specific resources; school and classroom libraries convert to multi-media centres to include e-resources
Teacher working conditions and job satisfaction	Better teacher working conditions are associated with higher learner achievement; aspects include workload, facilities and instructional materials
Principal leadership	A school principal can affect learners' academic performance; successful school principals focus on mission of school, seeking opportunities and monitoring curriculum implementation
School emphasis on academic success	School emphasis on academic success has an

Factors	Description
	effect on learner achievement; factors include teachers' expectations, parental support and curriculum implementation
Safe, orderly and disciplined school	School characteristics associated with learner achievement; safe, orderly and disciplined schools have a positive school climate; bullying amongst learners leads to poor academic results

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p. 40-45), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

An environment conducive to teaching and learning occurs where both teachers and learners have respect for each other as well as participating in constructive discussions (Cohen et al., 2009). If the school no longer has a positive climate, it can have a major effect on the teachers and learners. The lack of discipline in a school could lead to bullying among the learners which in turn results in poor academic performance (Rothon et al., 2011), particularly in the context of the classroom.

2.4.4 The Classroom Context

As part of formal education, learners receive their schooling in a classroom. Therefore, it is likely that there is an association between the teacher, the classroom environment and learner achievement. The PIRLS Literacy and PIRLS 2016 studies specifically looked at factors that enhance teaching and learning (*cf*. Table 2.7).

Table 2.7: Factors within the Classroom Context

Factors	Description
Teacher preparation and experience	Teacher content knowledge is critical; teacher professional development to incorporate new methods or strategies in the classroom; pedagogy in teaching reading
Classroom resources	Variety of resources in the classroom; classroom resources to be used during teaching; introduction of digital media into teaching and

Factors	Description
	learning
Instructional time	Emphasis on instructional time may affect learners' opportunities to learn; instructional time to be used effectively; homework given extends instruction
Instructional engagement	Importance of learner engagement with teacher instruction; instructional engagement can take place where learners listen and have discussions with their teachers and peers; teachers should manage their classrooms effectively to ensure engagement
Instruction for online reading	Online reading is an emerging aspect for reading instruction; learners use different methods to engage with online text in contrast to paper text; online reading assumes computer literacy
Classroom assessment	Teachers should monitor learners' progress and achievement throughout the school year; teachers provide feedback on assessment for learning

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p. 45-51), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

Each of the abovementioned factors may have an influence on learners' academic achievement. In a previous cycle of PIRLS, it was stated that even if learners attended class regularly, it does not mean that they will obtain higher achievement scores as the time spent in the classroom was not spent effectively (Mullis et al., 2007).

2.4.5 Learner Characteristics and Attitudes toward Reading

In addition to the abovementioned four contexts for learning, the PIRLS study also explores learner characteristics and their attitudes toward reading and learning. Reading literacy encapsulates more than just a learner's ability to construct meaning from a text, it also includes

the learner's attitudes and behaviour (Mullis et al., 2007). The PIRLS study looked at four different aspects that envelop learner characteristics and their attitude toward reading (*cf.* Table 2.8).

Table 2.8: Learner Reading Literacy Behaviours and Attitudes

Factors	Description
Learner readiness to learn	Learners should have prerequisite knowledge to engage with text; hunger and sleep deprivation lead to poor concentration in class
Learner motivation	Motivation is an important aspect to success in reading; motivation should be facilitated at home, classroom and at school; motivation could be intrinsic and/or extrinsic
Learner self-concept	Learners' own perceived confidence; self-concept is linked with learners' motivation; learners' self-concept relative to their peers' or their own experiences
Learner reading literacy behaviours	Learners who have a strong reading self-concept and are motivated tend to read more often than those who have a weak self-concept or who are not motivated; supportive home environment also plays a role in nurturing learners' reading habits

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p. 51-53), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

Research has shown that when learners experience a positive attitude towards reading, they are often better readers (Mullis et al., 2017). Learners who have more positive attitudes and have strong motivation for reading tend to perform higher than those who do not like reading (*cf.* Fives et al., 2014; Sainsbury & Schagen, 2004).

Having provided an overview of the various frameworks that underpin the development of PIRLS, the PIRLS Literacy 2016 and PIRLS 2016 instruments are explored in the subsequent section.

2.5 PIRLS LITERACY 2016 AND PIRLS 2016 INSTRUMENTS

PIRLS made use of two types of instruments, namely the achievement booklets and the contextual questionnaires. The PIRLS assessment framework (*cf*. Mullis & Martin, 2015) states that PIRLS consisted of 16 different booklets, one of which is the Reader. PIRLS booklets are compiled using a matrix design in which the passages and booklets are divided into blocks (Mullis & Martin, 2015). These instruments were developed in English by two international committees working with the International Study Centre (ISC) and the participating countries' National Research Coordinators (NRCs). In terms of the contextual questionnaires, the Questionnaire Development Group (QDG) refined, added or removed questionnaire items in line with current educational and reading trends.

This section specifically examines the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments (2.5.1) and the contextual questionnaires (2.5.2). Grade 4 and 5 learners' performance was derived from the reading achievement booklets and the contextual questionnaires were used to obtain valuable background information from the learners, their parents, the teachers and the school principals.

2.5.1 PIRLS Literacy 2016 and PIRLS 2016 Achievement Booklets

The PIRLS Literacy 2016 and PIRLS 2016 achievement booklets had to take into consideration the goals and coverage of texts, as stipulated in the PIRLS framework (Mullis & Martin, 2015). The PIRLS Reading Development Group (RDG) indicated that the achievement booklets should test the two purposes for reading; *reading for literary experience* and *to acquire and use information*.

PIRLS Literacy 2016 and PIRLS 2016 made use of a rotated test design and the matrix design of the achievement booklets included 12 passages and between 13 – 15 items. Each passage and its items were assigned to a block, thereafter the different blocks were allocated to individual booklets (Mullis & Martin, 2015). Of the 12 blocks, five literary blocks were developed for PIRLS, these are labelled *PRLit1-5* and the five informational blocks are labelled *PRInf1-5*. The PIRLS blocks also included six trend passages from previous PIRLS cycles. In order to establish a link between PIRLS and PIRLS Literacy, two blocks were taken from PIRLS Literacy, these include *PLLit3* and *PLInf3* and included in the PIRLS matrix design. Table 2.9 shows the PIRLS 2016 matrix design.

Table 2.9: PIRLS 2016 Matrix Blocks

Purposes for Reading	**Block**					
Literary Experience	PRLit1	PRLit2	PRLit3	PRLit4	PRLit5	PLLit3*
Acquire and Use Information	PRInf1	PRInf2	PRInf3	PRInf4	PRInf5	PLInf3*

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p. 59), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

*PIRLS Literacy 2016 Text

The PIRLS Literacy 2016 matrix design replicated the PIRLS matrix design, with five literary blocks, labelled *PLLit1-5* and five informational blocks labelled *PLInf1-5*. The PIRLS Literacy study used two blocks from the PIRLS study and these are labelled *PRLit1* and *PRInf1*. Of the ten PIRLS Literacy blocks, four were trend passages from prePIRLS that were used in the PIRLS Literacy study. prePIRLS was subsumed into PIRLS Literacy during the 2016 cycle. The remaining blocks were newly developed. Table 2.10 shows the PIRLS Literacy matrix block design.

Table 2.10: PIRLS Literacy 2016 Matrix Blocks

Purposes for Reading	**Block**					
Literary Experience	PLLit1	PLLit2	PLLit3	PLLit4	PLLit5	PRLit1*
Acquire and Use Information	PLInf1	PLInf2	PLInf3	PLInf4	PLInf5	PRInf1*

Note. From *PIRLS 2016 Assessment Framework* (2nd ed., p. 59), by I. V. S. Mullis and M. O. Martin (Eds.), 2015, TIMSS & PIRLS International Study Center. Copyright 2015 by TIMSS & PIRLS International Study Center.

*PIRLS 2016 Text

The PIRLS Literacy 2016 and PIRLS 2016 booklet design incorporated one literary and one informational block. The learners were required to complete each block in 40 minutes. Based on the number of literary and informational blocks available, the PIRLS booklet design consisted of ten PIRLS blocks (passages and items) and two PIRLS Literacy blocks (passages and items). These 12 blocks were then spread across 16 booklets; the 16th booklet is usually referred to as the Reader. The Reader is presented in a more authentic setting where the passages are presented in a magazine-style format and the items are given in a separate booklet. Table 2.11 depicts the PIRLS 2016 booklet design (Mullis & Martin, 2015, p. 60).

Table 2.11: PIRLS 2016 Booklet Design

Booklet	Part 1	Part 2
1	PRInf2	PRLit1
2	PRLit3	PRInf2
3	PLInf3*	PRLit3
4	PLInf3*	PRLit4
5	PRLit4	PRInf1
6	PRLit2	PRInf1
7	PRInf3	PRLit2
8	PLLit3*	PRInf3
9	PLLit3*	PRInf4
10	PRInf4	PRLit1
11	PRLit3	PRInf1
12	PLInf3*	PRLit2
13	PRInf3	PRLit1
14	PLLit3*	PRInf2
15	PRInf4	PRLit4
16 (Reader)	PRLit5	PRInf5

*PIRLS Literacy Text

The PIRLS study made use of Item Response Theory (IRT) to generate a comprehensive picture of each participating country's reading achievement by pooling every learner's responses to the booklets that they were assigned (Mullis & Martin, 2015). The blocks in the Reader were not linked to any other blocks in the other booklets.

The PIRLS Literacy booklet design was similar to the PIRLS booklet design, consisting of 15 booklets and a Reader. Each booklet consisted of two blocks and learners had 40 minutes to complete each block. As with PIRLS, the PIRLS Literacy booklets contained one literary and one informational passage with accompanying items.

The achievement booklets were randomly assigned to learners in advance of testing. The achievement booklets were administered in the LoLT used from Grades 1 to 3 of the learner's formal education. Grade 5 learners in South Africa were assessed in English, Afrikaans and isiZulu whereas the Grade 4 learners were tested in all official languages. Completion of the assessment booklets, containing one literary (fictional stories) and one informational passage, was dependent on the learners' ability to comprehend text, assessed by multiple choice and constructed response items (Mullis & Martin, 2015). This study examines five released texts which consist of three literary and two informational texts. Table 2.12 shows the selected PIRLS Literacy and PIRLS texts that will be utilised during the study:

Table 2.12: Selected PIRLS Literacy and PIRLS Passages

PIRLS Literacy 2016		**PIRLS 2016**		**PIRLS Literacy & PIRLS 2016**
Literary Passage	Informational Passage	Literary Passage	Informational Passage	Literary Passage
The Pearl	African Rhinos and Oxpecker Birds	Macy and the Red Hen	The Green Sea Turtle's Journey of a Lifetime	Flowers on the Roof

The abovementioned passages are followed by a series of items that assess the purposes for reading as well as the processes of comprehension. Table 2.13 indicates the word count and number of items for each of the released passages.

Table 2.13: Word Count and Number of Items for selected PIRLS Literacy and PIRLS Texts

Title of the Text	**Word Count** [a]	**Number of Items**
The Pearl	530	15
African Rhinos and Oxpecker Birds	446	17
Macy and the Red Hen	947	16
The Green Sea Turtle's Journey of a Lifetime	948	16
Flowers on the Roof	805	13

[a] Word count based on English text.

In the design of the PIRLS Literacy and PIRLS texts and items, the learners were expected to make use of a repertoire of reading strategies, which included retrieving and focusing on main ideas, making simple and more complex inferences and evaluating text provided. Data

collection for PIRLS consisted of a one-day testing session, allowing for 40 minutes for the completion of each of two blocks in the assessment booklet.

PIRLS also collected extensive information through the use of contextual questionnaires to provide policy makers with insights into improvement of reading achievement (Mullis & Martin, 2015). These include the school, teacher, parent and learner questionnaires. The questionnaires are a vital part of the data collection process and covered a wide array of policy-relevant information. Information gained from the questionnaires could provide valuable insight into the effectiveness of educational strategies in the participating countries (Mullis & Martin, 2015). However, although they form a valuable aspect of PIRLS, for the purpose of this study, the questionnaires are only mentioned and will not form part of the current study.

2.5.2 PIRLS Contextual Questionnaires

An important reason for conducting PIRLS is to examine possible school, community, home and learner factors related to learners' reading literacy (Mullis & Martin, 2015). In order to study these factors, the PIRLS study utilised a variety of questionnaires to collect data about each of these contexts for learning. These questionnaires were completed by the school principals, teachers, parents and learners. Additionally, the curriculum questionnaire, which deals with national and community contexts, was completed by the participating countries' National Research Co-ordinator (NRC) and is reported in the participating countries' PIRLS Encyclopedia (Howie et al., 2017).

2.5.2.1 Learner Questionnaire

The Learner Questionnaire was completed by all learners who participated in the PIRLS 2016 and PIRLS Literacy 2016 studies. The learners were requested to complete the questionnaire after they had completed the PIRLS assessment. This questionnaire gathered background information about the learners' home and school environments. In addition, the Learner Questionnaire also gathered information about learners' self-concept, and their behaviours and attitudes toward reading (Mullis & Martin, 2015). The questionnaire was in the test language[8] and took the learners approximately 15-30 minutes to complete.

[8] The test language was the same as the Language of Learning and Teaching (LoLT) in which the learners were taught from Grades 1 to 3.

2.5.2.2 Learning to Read Survey

The Learning to Read Survey (home questionnaire) was addressed to the parents or primary caregivers of learners who participated in the PIRLS 2016 and PIRLS Literacy 2016 studies. The questionnaires were provided to the learners to give to their parents prior to the testing and were collected on the day of testing. The questionnaire specifically asked the parents about their demographics, parent-child literacy activities, reading resources in the home as well as the parents' attitude toward reading. The Learning to Read Survey was provided in both English and the test language and was designed to be completed in approximately 15 minutes (Mullis & Martin, 2015).

2.5.2.3 Teacher Questionnaire

The Teacher Questionnaire was administered on the day of testing to the reading teachers of learners who participated in PIRLS 2016 and PIRLS Literacy 2016. The questionnaire collected information about the classroom environment for reading instruction (Mullis & Martin, 2015). The Teacher Questionnaire focused on classroom and teacher characteristics. Classroom characteristics include reading instructional time and approaches whereas teacher characteristics included career satisfaction, professional development and teacher education. The Teacher Questionnaire was provided in English and required approximately 35 minutes to complete (Mullis & Martin, 2015).

2.5.2.4 School Questionnaire

The School Questionnaire was completed on the day of testing by the principal of participating schools during the PIRLS 2016 and PIRLS Literacy 2016 studies. This questionnaire collected information about the school composition and school climate. School composition included the school facilities and resources, enrolment, principal leadership and education, whereas school climate focused on academic success and school discipline and safety. The School Questionnaire took approximately 30 minutes to complete and was available in English.

2.6 PIRLS 2016 TRANSLATION OF INSTRUMENTS

Since PIRLS 2016 and PIRLS Literacy 2016 are both large-scale international comparative studies, it was necessary to translate the instruments into the languages used in each participating country. The testing materials were developed by the TIMSS & PIRLS

International Study Center in US English (Martin, Mullis & Hooper., 2017). Participating countries were required to translate the testing materials into their language of instruction. In addition to the translation, the countries also made cultural adaptations to the testing materials as necessary (Martin et al., 2017).

2.6.1 Translation of International Instruments

As part of the PIRLS study's mandate, the translations and adaptations of the passages and items should be comparable to other participating countries. In order to ensure that the testing materials were of high quality and comparable across the participating countries, the countries followed an internationally agreed-upon standards and procedures in their preparation of the national versions of the PIRLS instruments. The development of the national versions of the assessment instruments was to accommodate participating countries' national languages as well as their unique context while upholding comparability across the different countries (Martin et al., 2017).

All testing materials underwent stringent quality assurance. Each country's translations and adaptations were externally reviewed by linguistic and assessment experts and consisted of two stages. The first stage was that of translation verification and the second stage was layout verification. These verifications were conducted before the field test and again before the main data collection.

During the first stage of verification, the national texts were compared to the international texts and provided the countries with feedback on how to improve the accuracy and comparability of the translated texts (Martin et al., 2017). When the translation verification was completed, the NRCs of participating countries reviewed the feedback and revised the national testing materials.

After the translation verification process was completed, the national testing materials were sent to the TIMSS & PIRLS International Study Center for layout verification (Martin et al., 2017). The layout verification stage ensured that the different national testing materials conformed to the prescribed format. During this stage, the TIMSS & PIRLS International Study Center also reviewed the national adaptations made to the testing materials to ensure that these adaptations did not unduly influence the comparability of the international testing materials. The materials included the learner achievement passages and items; context questionnaires, all

covers and directions for achievement booklet and questionnaires; and online covers and directions (Martin et al., 2017).

2.6.2 Translation of Instruments in South Africa

The PIRLS 2016 and PIRLS Literacy 2016 assessment instruments underwent translation from US English to the nationally selected target languages. In order to conduct the translations, South Africa's NRC had to follow the prescribed guidelines for translation. These guidelines included:

- The translation is at an appropriate level for the target population;
- No information is omitted, added, or clarified in the translated text;
- The translated text has the same meaning as the international version and uses equivalent terminology;
- The translated text has the same register (language level and degree of formality) and level of difficulty as the international version;
- Idiomatic expressions are translated appropriately, not necessarily word for word; and
- The translated text uses correct grammar, punctuation, qualifiers, and modifiers, as appropriate for the target language (Martin et al., 2017, p. 260).

2.6.2.1 Target Language

In South Africa, the PIRLS 2016 assessment instruments were translated into Afrikaans and isiZulu whereas, the PIRLS Literacy 2016 assessment instruments were translated into all official languages (Howie et al., 2017). With regards to the contextual questionnaires, the learner and home questionnaires were translated into the official languages. The home questionnaire also included the questionnaire items in English at the back of the questionnaire – this was done to assist parents who understood English better. Only the teacher and school questionnaires were administered in English. The assessment instruments and contextual questionnaires administered to English participants were adapted from US English to UK English.

2.6.2.2 Translators and Reviewers

The participating countries were advised to appoint professional translators and reviewers to work with the PIRLS testing materials (Martin et al., 2017). The South African NRC appointed professional translators who were registered with the South African Translators Institute.

Each participating country's translators and reviewers had to meet the following criteria:

- Excellent knowledge of English;
- Excellent knowledge of the target language;
- Understanding of the country's cultural background; and
- Experience in translating texts at the target grade(s) (Martin et al., 2017).

The reviewer had to assess the readability and accuracy of the translated passages. It was expected that the reviewer would be a person with experience with learners in the target grade. In line with the guidelines for translating and reviewing of the PIRLS 2016 and PIRLS Literacy 2016 instruments, the translators and reviewers had to keep in mind the following responsibilities:

- Ensuring that the translation is at the appropriate level;
- Ensuring that no information is omitted, added or clarified;
- Using equivalent terminology and ensuring that the translated text has the same meaning as the international text version;
- Ensuring that the translated text and the international version has the same register and difficulty;
- Ensuring that idiomatic expressions are translated appropriately; and
- Ensuring correct use of grammar, punctuation, qualifiers and modifiers (Martin et al., 2017).

The TIMSS & PIRLS International Study Center provided each country's NRC with a detailed list of changes to the international version that they could refer to in preparation for the main data collection process.

2.6.2.3 Translation and Adaptation

The TIMSS & PIRLS International Study Center provided each participating country's NRC with the materials which needed to be translated and adapted. The NRCs also received National

Adaptation Forms (NAFs) to document the translation, adaptation and verification processes (Martin et al., 2017). The NAFs were completed by the NRCs for each set of PIRLS and PIRLS Literacy assessment instruments and contextual questionnaires for each language administered.

The South African NRC had an immense task of translating and adapting the UK English version of the instruments into Afrikaans and isiZulu for the PIRLS study and into the 10 other official languages for the PIRLS Literacy study. Howie et al. (2017) explained that the "translation processes were protracted due [to] the subtleties and nuances of the passages" (p.36). The instruments were also released late and subsequently caused time constraints in terms of the translation process. The South African NRC and team had six weeks to complete translation, back translation, translation verification, formatting and layout as well as layout verification before printing could begin.

2.6.2.4 International Translation Verification

Upon completion of the national translations of the international instruments, it was compulsory that the instruments were internationally verified. IEA Amsterdam spearheaded the international translation verification process with the assistance of external translation verification companies (Martin et al., 2017). The verifiers of the selected companies had to meet criteria as set out by the IEA Amsterdam. The criteria included:

- Fluency in English;
- Mother tongue proficiency in the target language;
- Formal credentials as translators working in English;
- University-level education and (if possible) familiarity with the subject area; and
- Residency in the target country, or close contact with the country and its culture (Martin et al., 2017, p. 268).

The verifiers gave feedback about the translation verification for both PIRLS and PIRLS Literacy instruments. The verifiers also made use of the NAFs and were requested to correct the text of the instruments and questionnaires by adding comments onto the NAFs that described the errors (Martin et al., 2017).

One of the main tasks of the verifiers was to evaluate whether the same meaning and difficulty level in the translations and adaptations were adequate and consistent across the national instruments. This task was completed by checking the accuracy, linguistic correctness and

comparability of the national instruments; checking whether there were any deviations between international and national versions. The verifiers also provided the NRCs with alternative translations or adaptations to improve the comparability of the national instruments (Martin et al., 2017).

Throughout the translation verification process, the verifiers identified several types of errors and these included typographical errors, grammatical errors, omissions or additions of text, mistranslations, adaptations of names, gender issues and inconsistent translations (Martin et al., 2017). These errors were documented in the NAFs and sent to the NRCs to revise and improve their national versions of the international instruments. Table 2.14 depicts the verification feedback codes for PIRLS Literacy 2016 and PIRLS 2016.

Table 2.14: Verification Feedback Codes for PIRLS Literacy and PIRLS 2016

Code Number	Code Description
Code 1	Indicates a major change or error. Examples include the omission or addition of a question or answer option; incorrect translation that changes the meaning or difficulty of the item or question; and incorrect order of questions or answer options in a multiple-choice question.
Code 2	Indicates a minor change or error, such as a spelling or grammar error that does not affect comprehension.
Code 3	Indicates that while the translation is adequate, the verifier has a suggestion for an alternative wording.
Code 4	Indicates that an adaptation is acceptable and appropriate.

Note. From *Methods and Procedures in PIRLS 2016* (p. 269), by M. O. Martin et al. (Eds.), 2017, TIMSS & PIRLS International Study Center. Copyright 2017 by TIMSS & PIRLS International Study Center.

These codes were used in the NAFs to indicate the severity of the change in the national versions of the international instruments. Each code was assigned by the verifier to indicate to the NRC which changes needed their immediate attention for review.

2.7 DATA COLLECTION, MONITORING AND SCORING

Internationally, the PIRLS Literacy 2016 and PIRLS 2016 studies' field tests occurred between March and April 2015. After the conclusion of the field test, each participating country's NRC submitted the field test achievement data for analysis and review during April and May 2015.

Thereafter the PIRLS Literacy and PIRLS Item Development Task Force reviewed the field test item statistics in June 2015.

The new PIRLS Literacy and PIRLS passages were approved by the NRCs during the 5th NRC meeting in Finland in August 2015. The Southern Hemisphere countries conducted their main data collection during October to December 2015 and received scoring training in November 2015 for the constructed response items. The finalised scoring guides and training materials, used during the scoring training, was distributed to the NRCs in November 2015. The Northern Hemisphere data collection occurred between March and June 2016 and followed similar timelines as the Southern Hemisphere countries.

2.7.1 South African Field Trial

The South African field trial was conducted between 9 and 19 March 2015 (Howie et al., 2017), which was about nine months before the main study data collection for both PIRLS Literacy and PIRLS. Prior to the data collection, the Centre for Evaluation and Assessment (CEA) at the University of Pretoria, made contact with the schools who were selected to participate in the field trial. The CEA contacted 16 schools in the Gauteng province where the LoLT of the school was English as the field trial was only scheduled to be conducted in English. Of the 16 schools, two declined to participate and were replaced (Howie et al., 2017).

2.7.2 South African Main Study Data Collection

The main study data collection in South Africa took place from October to December 2015 and again in February to April 2016 due to the "Fees Must Fall" campaign which disrupted the planning and implementation of PIRLS with the university being closed for three weeks which prevented printing, packing and distributing of the instruments to the schools (Howie et al., 2017). Even though the data collection took longer, the South African NRC ensured that the data collection was compliant with the standards and guidelines set out by the IEA.

The NRCs nominated a person to be selected as the International Quality Control Manager (IQCM). IEA Amsterdam appointed an IQCM to monitor the process of each country's data collection. The main task of the IQCM were to conduct site visits at the participating schools during the data collection process, observing the PIRLS Literacy and PIRLS administration at the schools. Prior their site visits, the IQCMs attended a compulsory training session and were introduced to the PIRLS Literacy and PIRLS survey operations procedures and background

knowledge of the studies. They were also supplied with a manual that detailed their roles and responsibilities during the quality assurance process.

An external company was contracted to carry out the task of data collection across the nine provinces. The NRC conducted significance testing to determine whether there was any significant difference between the 2015 and 2016 performance of the PIRLS Literacy 2016 study but found that there is no significant difference (Howie et al., 2017).

After the completion of the data collection of the South African PIRLS and PIRLS Literacy studies, the assessment instruments were unpacked and sorted for scoring. Each of the assessment instruments contained two passages, one literary and one informational with multiple choice and constructed response items; it was constructed response items that required scoring. The South African NRC appointed a scoring team and followed strict procedures for scoring.

In preparation, batches were created for each administered language. The scoring process consisted the following processes:

- *Recruiting of Scorers:* the scoring team recruited scorers for each language administered. The scorers were also required to have had educational training.
- *Interviewing and Assigning Scorers:* selected scorers were divided into either Team A or Team B.
- *Training of Scorers:* the scoring team, consisting of CEA researchers fluent in Afrikaans, English and some African languages, received international scoring training. Thereafter, the researchers trained the local scorers for each passage's constructed response items.
- *Quality Assurance of Scoring:* team leaders for Team A and Team B received additional training to ensure quality assurance across the languages. The team leaders were responsible for the quality of instruments for their team.
- *Cross-Country Reliability Scoring:* this was conducted at the end of the PIRLS and PIRLS Literacy studies. South African scorers scored the same instruments as scored internationally using the IEA Cross-Country Scoring software. This process involved English language scorers to score items to compare to the international level.
- *Trend Reliability Scoring:* this process was conducted by scoring IEA materials from 2011 trend passages. Only countries who had participated in previous rounds of PIRLS

were able to conduct this scoring. Trend scoring was utilised to ensure that the scoring remains consistent over the years. The scorers made use of the IEA Trend Scoring software (Howie et al., 2017).

During the course of the scoring process, the NRC ensured that rigorous quality control was adhered to. A total of 25% of the assessment instruments were randomly checked by quality controllers for any discrepancies between Scorer A and Scorer B. Scorers were assigned to Team A or Team B and only scored those booklets allocated to Team A or Team B. A percentage was scored by the other team's scorers on a separate sheet, known as the reliability sheet. Figure 2.2 shows the scoring processes followed by the PIRLS Literacy and the PIRLS scoring teams.

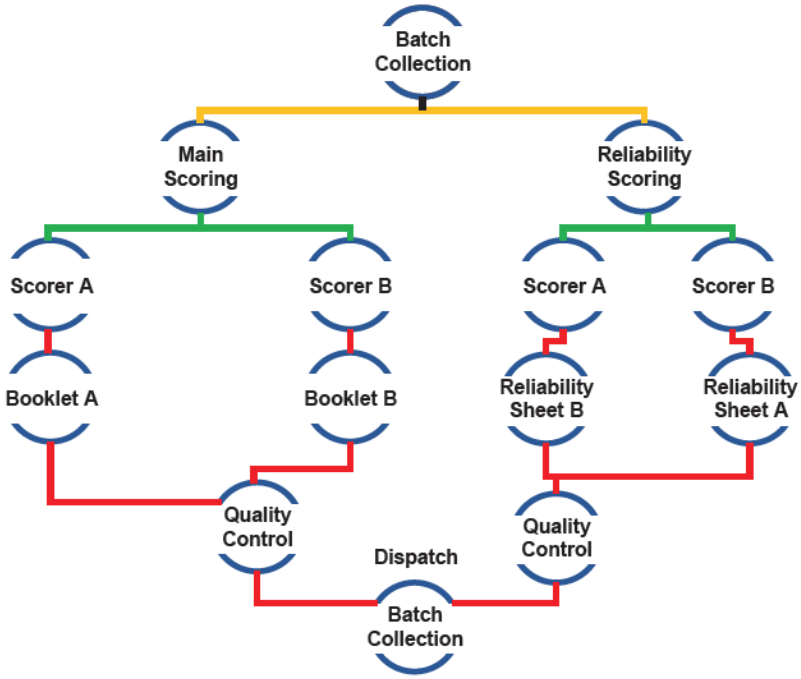

Figure 2.2: PIRLS Literacy and PIRLS Scoring Processes (Howie et al., 2017, p. 41)

Prior to the scoring process, the scoring team organised quality assurance workshops to "improve coherent understanding of the scoring process" (Howie et al., 2017, p. 41). The quality controllers concluded that the scoring process was of high quality as the reliability was above .90.

2.8 DATA CAPTURING AND PROCESSING

The IEA provided the NRCs with the Data Management Expert (DME) program for use during

the data capturing process. The IEA data centre as well as the CEA worked together to create data capturing templates for all of instruments and forms (Howie et al., 2017). The NRC appointed an external company to capture the data from the testing instruments, questionnaires and forms.

The CEA Data Manager trained about 40 data capturers on how the DME program worked, how to input the data and quality control their capturing. The CEA Data Manager was extensively involved in the capturing process as it required data cleaning, processing and sending feedback to the capturing company to correct any errors found. Data cleaning entailed identifying inaccurate records, correcting these inaccurate records and where necessary, correcting, modifying or deleting incomplete or irrelevant parts of the data (Howie et al., 2017).

The CEA Data Manager also ensured that 100% data verification was done for all the instruments. The reason for completing 100% data verification was to minimise capturing errors by using a two-person process. It should be noted that the DME program has built-in checks which assists the Data Manager with the quality control process.

2.9 RESEARCH ETHICS FOR PIRLS 2016 AND PIRLS LITERACY 2016 IN SOUTH AFRICA

The PIRLS Literacy 2016 and PIRLS 2016 studies were supported from the onset by the-then Minister of Basic Education, Angie Motshekga. The Minster provided ethical clearance to the CEA at the University of Pretoria to conduct both these studies in South African schools. Thereafter, the South African NRC sought permission from the University of Pretoria's ethics committee in order to conduct the study. The ethics committee provided ethical clearance to the CEA which resulted in obtaining permission from various other stakeholders to conduct these studies. Other stakeholders included the schools, teachers as well as assent from parents for their children to participate in PIRLS. As the IEA upholds complete confidentiality and anonymity of all participants in each of their studies, school identification numbers were provided and unique identification numbers were allocated to each learner.

2.10 CHAPTER SUMMARY

This chapter provided a brief overview of the IEA's history and the reason why international large-scale assessments, such as PIRLS, were developed. In order to explain PIRLS, it was important to provide a definition of reading literacy, indicating how the IEA conceptualised it,

bearing in mind the *Purposes for Reading* as well as the four *Processes of Comprehension*. The questionnaire framework was discussed, detailing the four contextual questionnaires used during PIRLS Literacy 2016 and PIRLS 2016. The methods and procedures of PIRLS were also outlined in this chapter which included the research design and methods – that is, assessment instruments, translation of instruments, adaptations (where necessary), data collection, monitoring, scoring and data analysis. Furthermore, this chapter assisted in distinguishing between the PIRLS methodology and the methodology selected for the current study.

-⌘-

CHAPTER 3

A Review of the Literature and Conceptual Framework

3.1 ORIENTATION

The intention of this chapter is to present a comprehensive review of the current literature on the issues related to this study's research problem of whether the PIRLS Literacy 2016 and PIRLS 2016 released texts are indeed equivalent as used in the South African version of the assessment. The chapter provides a summary of the South African educational landscape and focuses on Peña's four considerations of translational equivalence when testing learners in cross-cultural research studies, such as PIRLS, PISA and SACMEQ. Often in International Large-Scale Assessments (ILSA) the participating countries are required to translate a test from a source language, usually English, to the target language. In South Africa, the PIRLS Literacy 2016 study was translated into all the official languages while the PIRLS 2016 instruments were translated into Afrikaans and isiZulu (Howie et al., 2017a; 2017b).

In order to gain a theoretical understanding of this study, this chapter considers the South African education landscape (3.2) followed in 3.3 by a brief discussion of the Language in Education Policy. In 3.4, the translation and adaptations of text are discussed. Section 3.5 explores textual coherence, and in Section 3.6, a critical analysis of translation equivalence is presented. The conceptual framework of this study is presented and argued in Section 3.7. Concluding thoughts are presented in the last section of this chapter (3.8).

3.2 THE SOUTH AFRICAN EDUCATIONAL LANDSCAPE

This section begins with a brief overview of the South African educational landscape post-1994 (3.2.1). In 3.2.2, the current curriculum policy statement is explored.

3.2.1 South African Education post-1994

Prior to the 1994 democratic elections, the education system in South Africa was divided along racial lines where language was used as a means to segregate the citizens (Kanjee & Sayed, 2013). The only two official languages recognised by the-then National Party government was English and Afrikaans (Brook Napier, 2011). None of the indigenous African languages were recognised as official languages by the government, which resulted in educational policies that did not allow for multilingualism. The African languages were supposedly recognised in the

homelands. Apartheid education was described by Mohlala as "a system of education practiced in South Africa where different population groups receive separate, unequal, and racist education based on their skin colour" (1994, p. 8). During apartheid, there were 19 different educational departments separated according to race, location and ideology (DoE, 2003). Apartheid education was meant to be different for white and non-white children (Hlatshwayo, 2000) where language in education was used as a tool of segregation (Mtsatse, 2018). Paradoxically, the apartheid government acknowledged the importance of learning in the home language but simultaneously explained that it was not always possible to provide education in home language (*cf*. UNESCO, 1953). As a result, African learners were expected to use either Afrikaans or English from Grade 4 onwards and only have an African language as their HL (*cf*. Bantu Education Act) while white learners could continue their schooling in their home language.

The Soweto uprising of 1976 is regarded today as the highpoint of resistance to Bantu Education. During 1976 school boycotts and unrest at tertiary institutions become prevalent as Black Africans no longer wanted apartheid education forced upon them. It is here where the Black Consciousness Movement (BCM), led by black students, surfaced for the first time. The BCM enjoyed strong support from the African community as Black Consciousness "stood for a rejection of white Domination in all forms – political, economic, psychological and cultural" (Christie, 1996, p. 236). Years later, in 1989, initial steps towards change was apparent with the National Party (NP) government's awareness that their apartheid policy could no longer continue. On 2 February 1990, the-then President F.W. de Klerk moved the country into a new dispensation, taking on the critical role of dismantling of apartheid.

After the first democratic elections in 1994, the new government had the enormous task of implementing a new dispensation which does not discriminate against any of its citizens (Kanjee & Sayed, 2013). One of the major tasks that was set out by the new government was to disband the 18 separated education departments and create a single national department of education. The new national education department was set out to be inclusive of all races, languages and social backgrounds (Kanjee & Sayed, 2013). With the creation of the single national education department, there was a need to review the schooling curriculum policies and implement a new curriculum which encompassed democratic values and quality education for all. The Ministry of Education, under the new dispensation, introduced three curriculum reform initiatives, of which one was to "purge the apartheid curriculum" (Jansen, 1998) of

racially biased content. The second initiative was to develop and introduce continuous assessment in schools and the last initiative was outcomes-based education. However, before these initiatives could take hold, some education policies had to be revised and new policies drafted.

One of the first new education policies was the National Education Policy Act (NEPA), Act 27 of 1996. NEPA brought into effect the policies, legislative and monitoring responsibilities of the Minister of Education. One of the objectives of this act was "to provide for the publication and implementation of national education policy" (RSA, 1996a, p. A-3). Later, the South African Schools Act (SASA), Act 84 of 1996 was promulgated to give the School Governing Body (SGB) power over the selection for LoLT, Home Language (HL) and First Additional Language (FAL). Even though the SGB has the power to make these decisions, it remains dependent on the Constitution (Joubert & Prinsloo, 2001). The act was drafted to redress the inequalities and discrimination pre-1994, to support the rights of school-going learners and their parents as well as to specify the duties and responsibilities of the state (Joubert & Bray, 2007). Even though SASA is currently in place, it has not eradicated all facets of the previous educational regime. Today there are schools still challenged by apartheid-era problems such as schools being overcrowded, insufficient class and school resources, inadequate teaching facilities and infrastructure as well as inadequately qualified and experienced teachers, to name but a few (Manyike & Lemmer, 2014). Shortly after the establishment of SASA, the Language in Education Policy (LiEP) was introduced in 1997. The LiEP was drafted with both the Constitution and the SASA serving as guidelines. The LiEP is discussed later in Section 3.3.

Since 1994, the South African education curriculum has undergone "three main stages or waves" of reform (Chisholm, 2005, p. 80). During the first stage, the curriculum was cleaned up of all derogatory elements such as race, language and gender. The second stage comprised the newly developed Curriculum 2005 (C2005) which was developed in 1997 with an international form of a standards-based curriculum called outcomes-based education (OBE). OBE was adopted as it focused on assessment *for* learning and continuous assessment rather than assessment *of* learning. OBE required that all learning should be learner-centred where the teacher should act in the capacity of a facilitator. During the second stage, the National Qualifications Framework (NQF) was used to integrate the South African education system with a common qualifications framework. In the last stage, a ministerial review committee was

appointed to review the C2005 and concluded that a major revision of the education curriculum should occur.

The ministerial review committee report indicated that the implementation of the new curriculum was hampered by several factors:

a. alignment issues between the curriculum and the assessment policy;
b. inefficient teacher training and development;
c. inadequate learning and teaching materials;
d. policy burden;
e. teaching staff and resource shortages; and
f. lack of recognition of the curriculum by educational stakeholders (Chisholm, 2005).

Thereafter, the development of the National Curriculum Statement (NCS) was prioritised and completed in 2002. It focused on the knowledge content and was intended to make the curriculum transparent and comprehensible for teachers. The NCS was a key agent of a neo-liberal and democratic South Africa as it placed emphasis on values and nation-building (DoE, 2002). It also explained the kind of learner envisioned as one who "will be imbued with the values and act in the interest of a society based on respect for democracy, equality, human dignity, life and social justice" (DoE, 2002, p.8). Section 29(2) of the Bill of Rights stipulates that all those wishing to pursue education have the right to receive the said education in the official language of their choice. However, the aforementioned section also explained that the request to receive education in a particular language should be "reasonably practicable" (RSA, 1996a).

After the implementation of the NCS, a second phase was introduced in 2006, known as the Revised National Curriculum Statement (RNCS). The RNCS brought with it new learning outcomes and assessment standards derived from the critical outcomes (DoE, 2003). It also saw a renewed focus on learner-centred education as well as introducing activity-based education. Yet another curriculum review committee was established by the Minister of Education in 2008 and the report was submitted in 2009 (DBE, 2009). The report identified four key concerns:

a. Complaints about the implementation of the NCS;
b. Teachers who were overburdened with administration;
c. Different interpretations of the curriculum requirements; and

d. Underperformance of learners (du Plessis, 2013, p.2).

Based on the above factors, the curriculum underwent further revisions. This last round of revisions resulted in amendments to the curriculum to remove redundant features and include new features with the introduction of the Curriculum Assessment Policy Statement (CAPS) in 2012.

3.2.2 Curriculum and Assessment Policy Statement

The Curriculum and Assessment Policy Statement (CAPS) was developed as an amendment to the NCS (Pinnock, 2011) and not as an entirely new curriculum. To be specific, the CAPS is an amendment to what teachers teach, not how they teach (du Plessis, 2013). Moreover, the rollout of CAPS did not imply that OBE would be removed, as OBE is only a method of teaching. What changed was the curriculum, which takes the form of a content-orientated rather than outcomes-orientated approach that inclines to more traditional forms of teaching. It should be noted that with the amendment to the NCS, the CAPS provides all subject teachers with freely accessible comprehensive National Curriculum and Assessment Policies (Maskew Miller Longman, 2012).

There are some similarities and differences between the NCS and CAPS. Both the NCS and CAPS focused on providing curricula within the ambit of the South African Constitution. This means that in both instances, the primary goal of education is to provide all learners from diverse cultural, racial and linguistic backgrounds with equal and quality education. In light of this, some major changes were made during the conceptualisation of the CAPS. Table 3.1 shows the major changes between NCS and CAPS.

Table 3.1: Major Differences between NCS and CAPS

Concept	NCS	CAPS
Qualification (structure)	Grade R to 9 General Education and Training Certificate (GETC) did not realise in practice	Grade R to 12 GETC not mentioned Exit-level is Grade 12 National Senior Certificate
Critical Outcomes	Explicitly stated	Integrated in aims, curriculum content and skills
Development Outcomes	Reflect on and explore a variety of strategies to learn more effectively	Not mentioned

Concept	NCS	CAPS
	Participate as responsible citizens in the life of local, national and global communities Be culturally and aesthetically sensitive across a range of social contexts Explore education and career opportunities Develop entrepreneurial opportunities	
Purpose	Equipping learners, irrespective of their socioeconomic background, race, gender, physical ability or intellectual ability, with the knowledge, skills and values necessary for self-fulfilment and meaningful participation in society as citizens of a free country Providing access to higher education facilitating the transition of learners from education institutions to the workplace Providing employers with a sufficient profile of a learner's competences	Not clearly indicated A list is available in the National Senior Certificate (NSC) SAQA document
Principles	Outcomes-Based Education Participatory, learner-centred and activity-based education	Encouraging an active and critical approach to learning, rather than rote and uncritical learning of given truths
Inclusivity	Only mentioned	A general aim
Outcomes	Learning outcomes	Concepts Skills Content
Assessment	Assessment standards	Content Assessment
Teacher Role	Teacher role clearly depicted Norms and standards for teachers to uphold	Teacher role not mentioned Strict guidelines about how to teach

Concept	NCS	CAPS
Learner Role	Much emphasis on group work	Emphasis on learners taking responsibility for their learning
Approach to Learning	Discovery-based	Content-driven

Note. From *Introduction to CAPS* (p. 3-4), by E. du Plessis, 2013, Curriculum and Instructional Studies, UNISA. Copyright 2013 by Curriculum and Instructional Studies, UNISA.

The above table depicts the major differences between the previous curriculum and the amended curriculum. During the overhaul of the curriculum, changes were made across the different subjects including that of Home Language (HL), First Additional Language (FAL) and Second Additional Language (SAL). Although the official languages of the country may be taken at any of the aforementioned levels, not all the languages are available at all schools. In terms of primary education, learners are supposed to be taught in their HL. Yet, when learners progress to the Intermediate Phase a switch is made from HL to English and in some instances, Afrikaans (de Kadt, 2005). This switch means that learners, who were previously taught from Grades R to 3 in an African language, now need to switch to English even though they may now have the necessary foundation in the language of learning and teaching. Although this study does not focus on the switch of languages from Grade 3 to Grade 4 and its challenges, it is nevertheless important to highlight this phenomenon.

Against extensive curricular changes, the following section provides details on how current language policy plays out, taking into consideration the Constitution and the school policies that govern the way in which schools manage their language policies.

3.3 LANGUAGE IN EDUCATION POLICY

This section deals specifically with the Language in Education Policy (LiEP). A discussion of language development and proficiency is provided in sub-section 3.3.1, which is followed by an overview of the language policy (3.3.2). Thereafter, the current language approach, namely additive bilingualism is discussed (3.3.3.) and sub-section 3.3.4 deals with the implementation of the LiEP.

3.3.1 Language Development and Proficiency

Local researchers have found that language in education is a complex issue (*cf*. Henning et al., 2001; Kros, 2010; Taylor & von Fintel, 2016). Pretorius (2014) deliberates on language and literacy skills in South Africa, and explains that learning to read in the home language helps build strong literacy skills if properly introduced and applied. Internationally, Cummins (2001) emphasises the importance of language and language acquisition from an early age and how language plays a fundamental role in a child's language development. In his work, Cummins looked at how long it takes migrant children to develop fluency in their second language (Cummins, 1979). It is here where he created the common framework for language proficiency, which has two aspects: Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP) (Cummins, 1979, 2000). The language acquisition theory, as put forward by Cummins, emphasises the phases that learners need to work through in order for them to become fluent in a language. Throughout the language development process, learners will first develop conversational fluency, or BICS, nurtured through social interaction with adults (Cummins, 2008). Thereafter, learners will, with this foundation of basic interpersonal communication skills, begin to develop CALP where they learn the language necessary to understand and discuss content in the classroom. CALP, which deals with skills essential to academic learning such as listening, reading and speaking and relies strongly on skills of written language, is most often acquired and developed in a school environment where learners need to use the academic register (Cummins, 2000) required by the different subjects. Cummins' BICS and CALP theories developed from a language acquisition framework which "makes a distinction between context embedded and context reduced situations and cognitively demanding and cognitively undemanding tasks" (Cummins, 2000, p.68).

The PIRLS assessments focus on understanding of context-reduced language, in other words CALP, where language is more abstract and fewer non-verbal and linguistic clues are used. In context-embedded language, i.e. BICS, more clues are provided such as facial expressions, hand gestures or objects used as a reference (Cummins, 2008). Moreover, context-reduced text requires a higher level of cognitive demand of the reader as 'content flexibility' is required (Cummins, 2014) as literacy demands are high. As such, context-embedded tasks are considered less difficult than context reduced tasks.

3.3.2 Overview of the Language in Education Policy

All policies relating to schooling in South Africa are based on the Constitution. Two important school policies, developed since 1994, include the South African School's Act (SASA), Act 84 of 1996, and the Language in Education Policy (LiEP) in terms of section 3(4)(m) of the National Education Policy Act, Act 27 of 1996. The LiEP aims to redress past linguistic inequalities by encouraging multilingualism and recognising 11 official languages by describing language as a subject at school as well as a medium of instruction (Nwammuo & Salawu, 2018). Due to the multiplicity of languages in South Africa, the LiEP (DoE, 1997) was introduced to support LoLT in education whereby learners can receive, where possible, education in their home language in the Foundation Phase (Grades 1 to 3). The LiEP recognises that home language is the most appropriate for teaching and learning. In addition, the policy emphasises multilingualism whereby learners are introduced to a second language in the Foundation Phase in an additive bilingual approach.

Although the LiEP was a trail-blazer in creating awareness of multilingualism in the South African education system, there are some scholars who would argue that the policy was poorly implemented and possibly undermines multilingual education (Ferreira-Meyers & Horne, 2017). A language shift occurs from Grade 4 onwards where most learners change their LoLT to English, which means that learners who have an African language as their home language, switch to English or Afrikaans as the LoLT (Nwammuo & Salawu, 2018). As South Africa has 11 official languages, it may be a positive step towards social cohesion to enforce a language policy that encourages learning and teaching in languages other than Afrikaans and English.

3.3.3 Bilingualism in South African Schools

Some scholars have also voiced their concerns about the misalignment between the language policy and the learners' contextual and linguistic experiences (*cf.* Pretorius, 2008; Probyn, 2009; Webb, 2012). Some argue that the language policy does not take into consideration learners' multilingual repertoires (*cf.* Heugh, 2013; Makalela, 2017). Currently the language policy in South Africa focuses on an additive bilingual approach which means that a language – usually English – is added in the Foundation Phase as a first language. The additive bilingual approach includes support for learners' home language while acquiring an additional language

(Robertson & Graven, 2020). The aim of this approach is to develop learners' language and literacy proficient levels in their home language as well as in a first additional language.

Some scholars express their concern with additive bilingualism; Banda (2009) stated that bilingualism makes sense when a monolingual speaker is introduced to another language but not when that speaker is already bi- or multilingual. Ferreira-Meyers and Horne (2017) argue that often access to African languages is not guaranteed at schools. Moreover, Robertson and Graven (2020) found that several African countries still use former colonial languages for educational purposes. In South African education history, the English language has enjoyed prestige and status and the use of the language has increased over the years where many South Africans use it due to its international allure. By using English as the bridge between all races and cultures, it resonates with the contemporary goals to create national cohesion (Evans, 2017; Malan, 2012). That being said, it may prove difficult for English to be universally accepted as a race and culture crossing language as its origins lie in colonialism, with many South Africans stressing the importance of having a home language, accompanied by English.

A key aspect of the current language policy was to uplift African languages and provide a bi- or multi-lingual educational environment. However, it may be difficult to incorporate a multilingual schooling system as it will incur an additional financial burden on the current education budget to appoint additional language teachers in schools, sourcing print resources in all languages as well as classrooms to accommodate the different language groupings within one school (*cf.* Heugh, 2013; Turner & Wildsmith-Cromarty, 2014). In South Africa, social cohesion is important, and as such, citizens should be able to converse in more than one language as both the home language and English are important and enable one to take part in various social and professional activities. It may then be up to the teachers themselves to attend teacher training to find ways of dealing with a classroom of children who speak a variety of languages as a home language.

3.3.4 Shortcomings of the Language in Education Policy

When looking at the language policy context in South Africa, Jansen (2002, p. 199) uses the term "political symbolism" to explain the lack of policy implementation in education. He further argues that "in most cases, however, implementation was never on the policy agenda at all" (Jansen, 2002, p. 203). Jansen was not the only scholar investigating LiEP, several scholars

also stating their discontent regarding the discrepancy between creating language policies and the lack of implementation thereof (Beukes, 2009; Coetzee-van Rooy, 2018).

There are four main themes as to why LiEP is failing in the country:

1. the difficulty of transcending colonial and apartheid influences on language policies (Bamgbose, 2011; Plüddermann, 2015);
2. the control that English has in the development of indigenous languages (Ngwenya, 2012);
3. the inability to gather enough political will to implement language policies (Masoke-Kadenge & Kadenge, 2013); and
4. the absence of language and financial support to sustain the implementation of language policies (Coetzee-van Rooy, 2018; Madiba & Mabiletja, 2008).

Nevertheless, the LiEP could be considered a good language policy in the Foundation Phase as it contributes positively to educational outcomes (*cf*. Cekiso et al., 2019). What is concerning is the number of learners who still struggle to read in their home language after completing the Foundation Phase. Even though South Africa has a good language policy, quality education is not guaranteed. A possible reason for learners' poor literacy skills in their home language could be the teachers' lack of proficiency in the language of instruction. In a study conducted by Cekiso et al. (2019) in the Eastern Cape[9], it was found that teachers were not satisfied with their teacher training as they were not trained to use African languages as the LoLT. In addition, the lack of resource materials available in the 11 official languages could also negatively contribute to learners' literacy and language skills. Based on the shortcomings of the LiEP, the Department of Education (DBE, 2013) introduced the Incremental Introduction of African Languages (IIAL) in schools. One of the main goals of IIAL is the establishment of African language teaching, and ultimately, to improve learning outcomes (DBE, 2013). As the majority of South Africans have an African language, other than Afrikaans, as a home language, it is important to enshrine multilingualism so that learners who speak an African language can participate in all spheres of society. The IIAL was drafted to also enhance social cohesion in the country. Thus, it is important that policymakers take into consideration

[9] One of the nine provinces in South Africa.

language preferences and the inclusion of African languages when drafting and implementing language policies (Gordon & Harvey, 2019).

3.4 THEORY AND ADAPTATIONS OF TRANSLATION

As this study aims to determine whether the PIRLS Literacy 2016 and PIRLS 2016 released texts, across English, Afrikaans and isiZulu, are equivalent, this section explores the theory of translation (3.4.1) and adaptations (3.4.2) and, as equivalence goes hand-in-hand with validity, an overview of validity is presented in sub-section 3.4.3.

3.4.1 Theory of Translation

Theories of translation and adaptation have progressed over the years (*cf*. Rios & Sireci, 2014; van de Vijver & Leung 1997). The changes in educational assessment throughout the years were partly due to ILSA's cross-cultural psychology and fairness of testing (ITC, 2017). Translation is a specific term that refers to linguistic discourse moving from a source language (SL) into the target language (TL). The act of translation is to transfer the linguistic features and content from the SL into their equivalent form in the TL. The source language is the language that is to be translated whereas the target language is the language into which the text is to be translated. It is crucial that the translator has substantial knowledge of both the SL and TL as well as sensitivity in conveying story features such as setting, plot and the author's intent in order to produce a version that is equivalent and comparable to the source text.

Many different definitions of translation exist that reflect factors such as cultural and historical reasons or the interest of the researcher (Long, 2013). Translators are constantly theorising as different theories occur within different paradigms. Kuhn (1962) states that achievements can be "sufficiently unprecedented to attract an enduring group of adherents away from competing modes of scientific activity"; and that achievement is "sufficiently open-ended to leave all sorts of problems for the redefined group of practitioners to solve" (p. 10). Based on Kuhn's theory, definitions and theories of translations could be understood under different paradigms. As a result, translation studies all have different kinds of definitions for the act of translation.

Jakobson (1959) notes that translation comprises two equivalent messages in two different languages, while Vottonen (2016) describes translation as an "encounter of two languages, cultures and literary traditions" (p. 207). Translation also refers to the activity in which specific linguistic discourse is provided from a SL to a TL (Chan & So, 2017). In other words, it is the

act of conveying specific linguistic entities, such as words, sentences and paragraphs, from the SL to a TL. For example, '*The rhinos and the oxpeckers help each other*' in SL is translated into Afrikaans as '*Die renosters en renostervoëls help mekaar*'. Generally, the aim of a translation is to provide an equivalent target text (TT) to the ST (*cf*. Chan & So, 2017). While various definitions of translation exist, this study focuses on Nida's definition of translation as his work focuses on equivalence of translations (*cf*. Dore, 2016; Nida, 1964).

Nida sees translation as a scientific topic and states "the transference of a message from one language to another is a valid subject for scientific description" (Nida, 1964, p. 3). He also proposes that translation should be seen as an art. Based on this belief, he defines translation as "the closest natural equivalent of the source-language message, first in terms of meaning and secondly in terms of style" (Nida, 1964, p. 12). In his view, the receptor[10] of the translated work should be considered during translation, where it was previously mostly focused on the message. As such, the reader of the TT should experience the same effect as the reader of the ST (Dore, 2016). Nida's work emphasises that the receptor is the most important part of his translation theory, resulting in dynamic equivalence.

In dynamic equivalence, the aim of the translation is to ensure that the reader of the translated version of a text responds similarly as the reader of the source text (ST) (Long, 2013), even though each language has its own structures and the target language (TL) text differs from the SL text. It is possible to produce versions of a ST as long as the target language(s) has equivalent words that could be used during the translation process. Translation is thus seen as cultural interaction since the ST and TT are, in most instances, located in different cultures. The difference in cultures and languages could possibly play a role in how the reader reads and understands the text (Long, 2013). The focus of the translation, in dynamic equivalence, is placed on the reader rather than the text producer. The term 'dynamic equivalence' was later replaced by the term 'functional equivalence' as Nida modified his view of dynamic equivalence – where the ST and TL readers understood the text in the same manner (Miao, 2000).

3.4.2 Adaptations

Van de Vijver and Leung (1997) state that there are three methods of developing tests in

[10] Receptor is the reader of the text, test or other reading material.

multiple languages, with these incorporating direct translation using the translation back-translation method, test adaptation and the development of cross-cultural and cross-linguistic tests. In most instances, the direct translation method is used as it upholds a high level of equivalence across the translated versions. Koch (2009) explains that a better way of ensuring that the ST and TT are comparable in terms of context would be to adapt the text, "that is, to translate a set of items literally and to change the words or content of other items so as to enhance their appropriateness in the new cultural context" (p. 304). After the adaptions have been made, the translated and adapted texts can then be evaluated to determine their equivalence.

Adaptation is used to ensure that high levels of equivalence are attained during the translation and adaptation processes in test adaptation. However, Geisinger (1994) has identified two concerns when using translated and adapted assessment instruments, namely content equivalence and construct equivalence between the original (source) instrument or text and the target instrument or text. Matthews-López (2003) points out that there are some difficulties with test adaptation. These include establishing construct equivalence, adapting items, piloting of the instruments and making necessary changes to ensure scale compatibility. Adaptation is usually used to adapt an existing assessment instrument from the SL into a TL in order to create equivalent instruments to be able to make comparisons between the two groups.

The process of translation and adaptation is crucial for any cross-cultural or cross-lingual assessments and as such, the International Test Commission (ITC) has developed guidelines for the development of different versions of a text or test (ITC, 2017). The ITC guidelines were developed as there was a need for quality translations and adaptations of assessments in multiple languages (Hambleton, 2005). The ITC guidelines are separated into six topics: pre-condition, test development, confirmation, administration, score scales and interpretation and documentation. Of importance to this study is the third condition, namely confirmation as it involves equivalence and validity of an assessment across different cultures and languages (ITC, 2017, p. 16).

One method to determine if there is item equivalence in the different versions, is to perform Item Response Theory (IRT) analysis or Rasch analysis (RA). In their guidelines, the ITC (2017) acknowledges that it is possible that participants who write the translated and adapted versions may score lower or higher. This occurrence could pose a threat to the validity and reliability of the test. The solution to this conundrum is to create a subsample of the source

language group to match the target language group in order to distinguish between the differences in the shape of the distributions, which could then be eliminated (ITC, 2017). The validity of assessment instruments is important for any ILSA, and as such, the next section focuses on the concept of validity, the types of validity which are of importance to this study as well as the validity of translation of the PIRLS assessments.

3.4.3 Validity

This section focuses on validity and the importance thereof to assessment instruments and translations. The concept of validity is discussed (3.4.3.1) in order to give clarity on what validity is and why it is a focal point for any assessment. Face, content and construct validity (3.4.3.2) are explored in this study as it adds to the equivalence of the PIRLS texts and items. These three validities pertain specifically to how the assessment instrument looks, whether it measured what it intended to measure and on the meaning of the construct. Sub-section 3.4.3.3 considers the validity and translation of PIRLS assessment instruments, followed in 3.4.3.4 by a discussion on validity of assessment instruments.

3.4.3.1 The Concept of Validity

Validity is viewed as a key aspect of assessments (AERA et al., 2014). It is defined as the "extent to which a test score is appropriate for intended interpretation and use of the test" (Wools, 2015, p. 10). Traditionally, both validity and reliability have been regarded as the criteria to which a language test should comply. Van der Walt and Steyn (2008) posit that reliability has received more attention than validity as the reliability of a test can be statistically determined. The authors continue by saying that in more recent times, validity has enjoyed more attention as it is crucial in developing, as well as evaluating, language tests. Nevertheless, in order to evaluate the validity of a test's scores, the researcher should find and collect validity evidence that shows the appropriateness of the interpretation. This process is called validation.

The concept of validity is regarded as whether a test measures what it initially set out to measure. However, the notion of validity is under debate and is viewed somewhat differently by different groups: liberals, moderates, traditionalists and conservatives (Newton & Shaw, 2014, pp. 176-178). The liberals consist of authors such as Moss (2007) and Kane (2013) who state that validity should be extended to the whole evaluation of testing policy and not just to a single test entity. The next group, the moderates, believe that validity is an evaluation of the adequacy of a testing policy. This group includes AREA (1999). The traditionalists of validity

include Messick (1989) and Shepard (1997), who maintain that the test score meaning and test score use are undividable and as a result, restrict the very definition of validity to "the technical evaluation of measurement-based decision-making procedures" (Wools, 2015, p. 10). The last group, namely the conservatives, think that validity should involve the technical quality of measurement procedures and explain that validity only concerns test scores. Authors in the conservatives group include Borsboom and Mellenbergh (2007) and Cizek (2012).

Davies and Elder (2005) emphasise the difference between the concept of validity and the act of validation. They explain that validity is an abstract concept and "validity is only as good as its validation procedures" (Davies & Elder, 2005, p. 795). Validation is also explained as "an activity: the collection of all possible test-related activities from multiple sources" (van der Walt & Steyn, 2007, p. 141). The multiple sources referred to here, include aspects which are traditionally conceived as face, construct and content validity. It also includes other types of validity such as predictive validity and consequential validity.

> The validation process involves the development of a coherent validity argument for and against proposed test score interpretation and uses. It takes the form of claims or hypothesis (with implied counter claims) plus relevant evidence (van der Walt & Steyn, 2007, p. 142).

In terms of validity, this study focuses on face, content, construct and translation validity as part of determining the equivalence of the PIRLS 2016 assessments.

3.4.3.2 Face, Content and Construct Validity

As already indicated, there are many different types of validity, such as face validity, content validity, construct validity, predictive validity, internal validity, external validity, concurrent validity and consequential validity, to name but a few (Cohen, Manion & Morrison, 2008). However, for the purposes of this study, only face, content and construct validity are relevant. *Face validity* "requires an examination of a measure and the items of which it is comprised as sufficient and suitable 'on its face' for capturing a concept" (Martinez, 2017, p. 1823). In other words, face validity occurs when a judge evaluates a measurement technique and makes a decision of whether the measurement technique appears to measure what it intended to measure (Maruyama & Ryan, 2014). It is also a form of internal validity (Gaber, 2010). For example, the PIRLS test, taken at face value, looks similar to other literacy tests for Grade 4 learners. The texts used in the PIRLS assessments are similar to a passage or story that learners can find

in the classroom, libraries and at home, and the question formats (constructed response and multiple choice) are familiar formats for assessing reading comprehension.

Content validity "refers to the extent to which the items on a test are fairly representative of the entire domain the test seeks to measure" (Markus & Smith, 2010, p. 239). Content validity looks at relevance to the curriculum, focus on what was taught, comprehensive content coverage, the proportion of the scope of learning and sampled potential content (Babbie, 2007) and is based on expert judgement. In other words, content validity is the extent to which the instrument deals with the content that it has set out to use. For instance, whether the PIRLS test covers the content that would specifically test the Grade 4 learners' reading literacy abilities. As discussed in Chapter Two, the PIRLS assessment instrument comprised two kinds of texts, namely literary and informational texts. These texts are accompanied by a range of items to test learners' reading literacy skills by assessing different levels of comprehension. As such, it could be argued that the PIRLS assessment instruments, in terms of the content, cover what they set out to cover, namely various aspects of text content that test learners' reading literacy levels.

Construct validity refers to the notion that "the appropriateness, meaningfulness, and usefulness of score-based inferences are inseparable and that the unifying force behind this integration is the trustworthiness of empirically grounded score interpretation" (Messick, 1989, p. 8). There is more than one way to establish construct validity:

- the measure is able to reflect clients' developmental changes;
- the measure correlates with previous measures that have proven construct validity;
- the underlying dimensions or traits of the construct have been identified using factor analysis;
- the measure's internal consistency has been demonstrated using statistical techniques such as biserial correlations;
- the measure has convergent and discriminant validity; and
- the construct is tested experimentally to see whether supposed interventions alter the construct in the hypothesized direction (Anastasi, 1988, pp. 147-150).

Construct validity could be seen as the degree to which the construct, that is, the test, measures what it set out to measure (Maruyama & Ryan, 2014; Taylor, 2013). For example, the PIRLS instruments (containing both text and items) set out to test learners' reading literacy

comprehension by covering specific content that examines reading literacy comprehension. Trochim and Donnelly (2006) clarify that face and content validity are actually types of translation validity, where the validator of the text determines whether the operationalisation (translated text) is an accurate reflection of the construct (original text). Under construct validity, face validity is seen as whether the operationalisation appears to be a good translation of the construct.

3.4.3.3 Validity and Translation of PIRLS Assessments

According to Arffman (2013), when ILSAs translate their assessment instruments into multiple languages, these translated versions of the source texts should be "equivalent, or comparable, to each other" (p. 2). ILSAs, in other words, should ensure that the same construct is measured among the different versions and that the test instrument should be equally easy or difficult for the learners to answer. There should not be a higher cognitive load for some learners than for others based on the translation of the texts (Rueda, 2011). If, however, one of the translated texts is cognitively more taxing for readers and results in more working memory needed to decode and comprehend it, this could be seen as a threat to the validity of the assessment instrument, as one group of learners is at a disadvantage. For example, if a ST provides a literal answer to a question, then the TT must be such that it also provides a literal answer to the same question as in the ST, and not an answer that needs to be inferred.

Even though ILSAs take great care to create equivalent translations, Arffman (2013) suggests that there have been problems translating assessment instruments and that equivalence may not always be attained. Hilton (2006) has criticised the manner in which PIRLS is conducted. She has raised questions regarding the validity of not only PIRLS but also ILSAs in general, and bases her criticisms on cultural validity, construct validity, methodological issues and the PIRLS survey in England. Her research indicated that there is cultural strangeness and linguistic bias in the PIRLS test, and she bases her argument on Bonnet's (2002) view of ILSAs:

> This scale is supposed to be the same in every country, which makes it possible to measure it in the same way everywhere. Not everyone believes this to be the case because of the influence of linguistic and cultural factors on the test (Bonnet, 2002, p. 389).

Hilton (2006) further argues that there are underlying issues of testing reading comprehension across different cultures, which include issues of the learners' understanding being reliant on variables such as culture, language, ethnicity, schooling and socio-economic status. To illustrate this aspect, she referred to the released PIRLS text *Puppy Walking* about the training of a guide dog. The following is a selected piece from the text:

> She also goes with Tim to soccer and tennis and with Whitney to ballet and gym.
>
> Like all puppies Goldie can be naughty. She likes to take smelly socks outside and hide them. And if no one's looking, she tries to lick the plates in the dishwasher (Campbell et al., 2001, p. 75).

Hilton argues that the above excerpt of the text contains cultural assumptions and understandings that may not be true for all cultures participating in PIRLS and as a result does not have unidimensionality. However, Whetton et al. (2007) rebuts Hilton's arguments by conducting their own statistical analysis and explained that the issue of dimensionality was addressed by the analysis conducted by the PIRLS team to determine which items functioned differently for individual countries; if some items functioned differently between countries during the field trial, those items were rejected. It should be noted that if Hilton's argument is taken to its logical conclusion, it suggests that we can only really read about things that we know, yet reading can be a means of learning about concepts, ideas and other phenomena with which we are not familiar.

Based on the above arguments, it is important to establish the equivalence of ILSAs across the different languages and this is done by conducting meticulous translation procedures (*cf.* Martin et al., 2017). The TIMSS & PIRLS International Study Center have over the years developed a rigorous translation verification procedure to ensure that the translations of the PIRLS texts are equivalent or comparable. As discussed in Chapter Two, the texts used for the PIRLS 2016 and PIRLS Literacy 2016 assessments are developed in English (the international source version) and then sent to the participating countries to adapt to their cultural context as well as to translate the texts to their languages of teaching and learning (Martin et al., 2017). The TIMSS and PIRLS International Study Center has developed Survey Operations Procedures (SOP) that deal with the translations of the texts, items and other materials to ensure that high quality and comparable translations are accepted for testing. "The ultimate goal of the translation and adaptation process was to create national versions of the PIRLS 2016

instruments that accommodate national languages and context while maintaining international comparability" (Ebbs & Wry, 2017, p. 7.1).

In South Africa, the translation and adaptation processes were substantial, as the testing material (text and items) was adapted to English UK and then translated into 10 of the 11 official languages of the country. The learner and parent questionnaires were also translated into all the official languages whilst the teacher and school questionnaires were only available in Afrikaans and English.

The translation of the PIRLS assessment instruments in South Africa is reported on in *PIRLS 2016 Encyclopedia: Education Policy and Curriculum in Reading: South Africa Encyclopedia*. (Howie et al., 2017). After the South African National Research Coordinator (NRC) received the testing materials from the IEA, the adaptation and translation processes could begin. The US English versions of the international source material were versioned to UK English and adapted to the South African context, with input from the SA Steering Committee. Thereafter the UK English versions of the different texts were translated into the remaining ten official languages. According to Howie et al. (2017) the South African team, along with other southern hemisphere participating countries, had less time to complete the translation processes as the international materials were released late. The translation process which includes translations, back-translation, translation verification, formatting and layout verification, was completed in six weeks. The NRC employed several Grade 4 language teachers to review the translated texts for accurate and appropriate translation.

Howie et al. (2017) explains that, for the South African PIRLS, the aim of the translations was to generate equivalent versions of the source text. However, it proved difficult with particular texts as certain phrases and vocabulary were not available in some of the languages. Nevertheless, the translation team endeavoured to translate all texts and items as accurately as possible. Any changes made to the texts were carefully recorded on the National Adaptation Forms (NAF). Recording all the changes provides some form of quality assurance as the changes had to be accepted by the IEA in order to keep changes minimal, while at the same time acknowledging the various participating countries' national contexts.

3.4.3.4 Validity of Assessment Instruments

As part of the validity of a test instrument, such as PIRLS, it is also important to look at the interpretation of the test scores. To evaluate the validity of test scores, one should gather

validity evidence to show appropriateness of the interpretation and use (Wools, 2015). Taylor (2013, p. 2) describes validation in assessment as "evaluating logical arguments and empirical evidence to determine whether they support proposed inferences from, as well as interpretations and uses of, assessment results". Validity has been operationalised, at least in education research, around the use of standardised tests. According to Wools (2015), there is no current set of criteria available that can be used to evaluate the validity of assessments.

Messick (1995) has looked at validity in terms of interpreting test scores and the generalisability and possible replicability of tests. Based on Messick's conditions, the PIRLS 2016 and PIRLS Literacy 2016 instruments would be labelled as valid as the results of the aforementioned studies were generalizable to the broader learner population, specifically Grade 5 and 4 learners, respectively. Furthermore, Messick (1992) postulates that "what is to be validated is not the test or observation device as such but the inferences derived from test scores or other indicators – inferences about score meaning or interpretation and about the implications for action that the interpretation entails" (p. 1487). Messick (1995) continues by stating that validation should focus on the association between the test instrument (content) and the interpretations of the learners' results.

In terms of PIRLS Literacy 2016 and PIRLS 2016, the embodiment of the test content relates to the satisfactory portrayal of the comprehension processes (*cf*. Chapter Two) used to examine learners' reading literacy abilities. The PIRLS assessment framework stipulates that if, for example, an item is testing learners' inferential abilities such as details about themes or about the author's intent, then the item should ask learners to answer an *interpret and integrate ideas and information* type question (Mullis et al., 2015). Furthermore, test validity also considers relevance and utility of the content specifications of assessments (Cumming & Berwick, 1996). Within this study, English, Afrikaans and isiZulu learners' performance is validated by looking at the translation of the released PIRLS Literacy 2016 and PIRLS 2016 texts, where relevance and utility could be understood as the purpose of the text. It also refers to functionality, which relates to the operationalisation of the target text, that is, the Afrikaans and isiZulu texts, in terms of the text producer, translator and text reader.

With regard to the validity of PIRLS Literacy 2016 and PIRLS 2016, it is prudent to first look at the aim of these studies, which is to conduct international reading literacy assessments and to explore associated factors that may have some bearing on learners' literacy achievement across countries (Mullis & Martin, 2015). According to Venuti (2012), validation of instrument

translation includes the verification as well as the quality assurance thereof. In the case of PIRLS Literacy 2016 and PIRLS 2016, the texts and items selected for testing should draw on the learners' comprehension processes so that the learners are able to construct meaning from the text, and in turn, answer questions based on the different processes of comprehension. The validity of the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments could be determined by the interpretation of test scores, such as Classical Test Theory (Labuschagne, 2014).

Section 3.4 dealt with translations, adaptations and validity of translations. As this study focuses on the equivalence of ST and TT of the PIRLS Literacy 2016 and PIRLS 2016 released texts, special consideration is given to the translation and validity of these texts.

3.5 TEXTUAL COHERENCE

This section considers what constitutes a text (3.5.1) and in sub-section 3.5.2, discusses the coherence of a text, which includes a discussion of translation and coherence.

3.5.1 What is a Text?

Before one can analyse a text based on the principles of textuality, it would be prudent to define what a 'text' is. According to Richards and Schmidt (2010), a text is:

> a segment of spoken or written language that has the following characteristics:
>
> 1. It is normally made up of several sentences that together create a structure or unit, such as a letter, a report, or an essay (however one word texts also occur, such as DANGER on a warning sign).
> 2. It has distinctive structural and discourse characteristics.
> 3. It has a particular communicative function or purpose.
> 4. It can often only be fully understood in relation to the context in which it occurs (p. 594).

However, for the purposes of this study, the definition of text will be further explored within the ambit of coherence.

A piece of language is recognised as a text as soon as the reader recognises that it has a communicative purpose (Widdowson, 2007), in other words to create a connection between

the reader and the text. The aforementioned connection is used to convey the meaning of the text to the reader. If it happens that a text does not effectively convey its meaning, then the reader may not realise the meaning of said text.

3.5.2 Coherence

Coherence refers to the notion of unity or 'togetherness' in a text, that is, that the whole text makes sense. Coherence includes both reader and text features and refers to the logical flow of a text, creating a mental representation of the text being read (Menzel et al., 2017). It involves how the different components of textual matter are accessible and appropriate (De Beaugrande, 1995; De Beaugrande & Dressler, 1981); for example, the organisation of *concepts* and *relations* that underpin surface text. A *concept* is an organisation of cognitive content that can be recalled or activated with some consistency in the mind and *relations* are the connections between concepts, that is, *relations* that join sentences, clauses or paragraphs in particular ways. Sometimes the *relations* are signalled explicitly with linguistic devices as in (a) below, and sometimes they are implicit, and the reader has to infer the connection (Menzel et al., 2017) as in (b). For example:

- (a) '*It was a beautiful day and as a result we decided to have a picnic.*'
- (b) '*It was a beautiful day. We decided to have a picnic.*'

Coherence *relations* are created in the reader's mind with the assistance of linguistic and discourse markers (Louwerse & Graesser, 2007). These markers may include a specific word, phrase, and sentence or feature that helps the reader to interpret events or ideas and their connections to other events or ideas contained in a text. It can be argued that coherence is the outcome of a reader's cognitive processes as text does not make sense on its own; it requires the reader's knowledge and experience to be able to interact with what is captured in the text (*cf.* Petöfi, 1973). In their work, Graesser et al. (2003) state that "a text is perceived to be coherent to the reader when the ideas hang together in a meaningful and organised manner" (p. 4). Table 3.2 depicts the different classes of coherence relations.

Table 3.2: Classes of Coherence Relations

Class of Coherence Relations	Description
Coreference	Coreference takes place when two words refer to the same thing, person, concept or idea. Coreference includes pronouns

Class of Coherence Relations	Description
	for example *he*, *she*, *them*, *their*, *it*, *what*, *who* and so on. *They* is coreference for a group of people.
Deixis	Refers to people, locations and time in a conversation among the reading participants. It may occur between a reader and writer or between two characters in a story. The pronouns refer to people (*you*, *me*), to locations (*here*, *there*) and time (*now*, *later*).
Given-new structuring of information	Content of each sentence in a text is divided into given/old information and new information. Given information has already been introduced (or is assumed to be known by the reader) whereas new information is added to advance the discourse of new content.
Conjunctive relations	Conjunctive relations are text-connecting relations that connect sentences or adjunct clauses. The subcategories include additive (*and*, *so*), temporal (*then*, *when*), causal (*consequently*, *because*), intentional (*in order to*), adversative (*but*, *however*) and logical (*so*, *therefore*).
Verb tense and chronology	Events in a story unfold chronologically, or when using an expository text, the events are described in a sequential order. Tenses (*past*, *present*, *future*) used also help in keeping track of the order of events. Parallel structures also contribute, for example, consistent use of the passive to describe something rather than switching between active and passive (e.g. *passive*, *passive*, *passive*, instead of *passive*, *active*, *passive*).
Scene changes	Scene is a spatial construct and includes the characters in a story who interact with each other. A scene change is signalled by stating for example, *meanwhile*, *back at home*.
Headers and highlighting	Headers, sub-headers and highlighted words operate as discourse functions. These aspects assist in organising text and may also help the reader to determine the text genre.
Topic sentences	In expository texts, the first sentence in each paragraph typically captures the main topic or theme for that particular

Class of Coherence Relations	Description
	paragraph. The subsequent sentences embellish the main topic sentence.
Punctuation	Quotation marks are used to signify speech acts of characters in a story. Other punctuation also has conventional meanings such as periods (.), commas (,), colons (:) and dashes (-).
Signals of rhetorical structure	Coherence relations have an important role in the recognition and organisation of the rhetorical structure. For example, lists and orderings (*first*, *second*, *last*). These signals are distinctive.

Note. From *What do readers need to learn in order to process coherence relations in narrative and expository text* (p. 90-96), by A. C. Graesser et al., 2003, Guilford Publications. Copyright 2003 by Guilford Publications.

Text linguists (*cf*. De Beaugrande & Dressler 1981; Halliday & Hasan, 1976; 1995) argue that when a text makes use of coherence markers, as outlined in Table 3.2, the content of the text is organised in such a manner that it is easy to read (Najafi et al., 2020). When phrases, sentences and paragraphs contribute to the meaning of a text, then that text could be considered coherent. Mikhchi (2011) explains that "a coherent text has an underlying logical structure that acts to guide the reader through the text" (p. 53). Therefore, a coherent text is where all the sentences, paragraphs and other parts of text form a whole.

Van Dijk and Kintsch (1983) developed a model of reading comprehension. Reading comprehension occurs when the reader builds a mental representation of the text (*cf*. Kintsch & Rawson, 2005). Van Dijk and Kintsch's (1983) model distinguishes between two levels of representation of a text, namely *text-based* and a *situation-based* representation of what readers build up while they read. During both levels, meaning construction is thought to be an incremental process – in other words, where the text is processed through several cycles with only a portion of the text processed during each cycle (Albrecht & O'Brien, 1993).

The distinction between the two levels is not new – a singular aspect motivates the notion behind *text-based* and *situation-based* representation of text, and that is to understand what the text is about. *Text-based* representation includes both text and knowledge elements (van Dijk & Kintsch, 1983). *Text-based* questions often focus on fact-finding or literal type items (Sabatini et al., 2012).

Situation-based representation includes creating a mental representation of the information contained in a text (Bos et al., 2015). The *situation-based* representation gradually builds up on text features such as main characters of the story, location, and time, amongst others. *Situation-based* question items usually refer to inference type items, interpretations, predictions and speculations (Sabatini et al., 2012). Readers can build a good situation model if they read attentively and work through the text – in so doing, it enables the reader to answer higher cognitive type questions and it helps the reader gain a deeper understanding of the text (Bos et al., 2015).

According to van Dijk and Kintsch (1983), if a reader is not able to imagine the situation in which an individual has certain characteristics (as shown in the text) and if the reader does not understand the relations between local and global facts within the text, it might provide evidence that the reader does not understand the text. When the reader reads a piece of text, she/he activates her/his comprehension strategy to facilitate the mapping of new information onto existing information. The type of mapping process is reliant on the reader's selection of a strategy that is best for keeping information in focus (Albrecht & O'Brien, 1993). The mapping process can be divided into two processes, namely *local* and *global* coherence. The former refers to how clear a sentence or paragraph in a text links within and across adjacent sentences. The latter refers to links across larger units of text and how united the entire text is (van Dijk & Kintsch, 1983), in other words, establishing connections between the already processed information and information which has already occurred, that is, across paragraphs or sections. Readers establish both local and global coherence during reading comprehension. Establishing local and global coherence is developmental. It is easier for younger and weaker readers to make local connections. As readers mature and become more skilled, they increasingly make global connections (Pretorius, 2000).

In terms of translation, van Dijk and Kintsch (1983) explain that translation is not simply one surface form taking another form. When the ST is translated into the TT, the situation model should be applied; however, if the SL and TL differ, a more explicit situation model is required during translation. It is argued that when translating a coherent text, the translated version should also be coherent (*cf.* Menzel et al., 2017).

3.6 TRANSLATION EQUIVALENCE

The importance of translation equivalence is still best described by Jakobson (1959)

"equivalence in difference is the cardinal problem of language and the pivotal concern of linguistics" (p. 233). Ideally, the piece of text or instrument being translated from the SL into the TT should be textually equivalent. As such, this sub-section focuses on the four main threats to validity concerning the instruments' translational equivalence, *viz*. linguistic, functional, cultural and metric (Peña, 2007).

3.6.1 Linguistic Equivalence: Translation of Instruments

Linguistic equivalence concerns the translation of text and ensuring the translated version of the text is similar across the SL and TL (*cf*. Chesterman, 2016). In practice, researchers or translators will make use of two main translating techniques. The first being translation and back-translation (Beck et al., 2003). During this process, the translator translates the text from the ST into the TT. Thereafter a different translator independently translates the TT back to the language used in the ST (Behr, 2017). After the translation and back-translation processes are conducted, the researcher can compare the original text with the translated text in order to find errors and correct them. It is important to make use of a translator who speaks the language at home language level to ensure accuracy of the translation.

According to Peña (2007), the objective of linguistic equivalence "is to make certain that the words and linguistic meaning used in the instruments and instructions are the same for both versions" (pp.1256-1257). However, it may not be easy to meet this objective. When a text is translated into a different language, the words may be similar but differences may arise in responses. In other words, if a language test or questionnaire is translated into another language, the two groups of respondents completing these tests, may have different responses due to cultural differences, familiarity and interpretation of the text, amongst other reasons (*cf*. Behr, 2017). For example, in a questionnaire that is developed in English and is translated into isiZulu, the English respondents may respond differently to the isiZulu respondents. Moreover, scholars have debated the pros and cons of back translations, some arguing that issues and discrepancies of the ST and TT could be due to forward translation errors as well as during the back translation (*cf*. Colina et al., 2017; Harkness, 2003).

If, however, the aim of the translation is only to look at the linguistic equivalence, then conducting translation and back-translation should be sufficient. Although, if the researcher or translator wants to use a translated text that is culturally sensitive, but which also functions the same across the different language groups and has the same item difficulty for the different

groups, then additional consideration should be paid to cultural, functional and metric equivalences to reduce bias and guard against validity threats. In the case of the PIRLS assessments, they are developed in English, and are then translated into a target language (TL). The source language (SL), namely English, might have different typological and orthographic[11] features. English has an opaque orthography whereas African languages such as isiZulu have a transparent orthography (Spaull et al., 2020). For example, in English each letter could represent a different sound (Miller, 2018); for example, the grapheme '*c'* in *car* makes a '*k*' sound whereas in *celery* it is an '*s*' sound. IsiZulu is also an agglutinative[12] language where each letter represents a specific sound, meaning it has a transparent orthography, for example '*umama*' (mother) is spoken as u + ma:ma.

The above example shows that translating a text from a source language (SL) to target language (TL) is complex and that linguistic equivalence cannot be established simply by using back-translation for languages other than English. It reveals one of the threats when translating assessment instruments for cross-cultural studies when different languages across the world are used. When translation is required in different languages with different orthographies, it is vital to ensure that the assessment instrument is equal for all learners participating in the assessment to demonstrate their abilities. In the case of PIRLS, learners from different countries, cultures and languages should then be afforded equal and fair chance of completing the reading assessment. Peña (2007) states that in order to guard against these threats of validity against the assessment instrument, different kinds of equivalences should be considered such as linguistic equivalence, discussed above, as well as functional, cultural and metric equivalence, discussed in the subsequent sections.

3.6.2 Functional Equivalence: Instrument Behaviour

In the 1960s Nida (1964) described translating texts as reproducing a target text in the target language in the closest possible equivalence to the source text and language, thus ensuring equivalence in terms of meaning and also in terms of style. Functional equivalence aims for a "naturalness" (Bermann & Porter, 2014, p. 72) of the translation, meaning that it aims to elicit the same behaviour across the SL and TL (Greenfield et al., 2006). Nida made an immense contribution to the field of translation and although his work was mostly based on the

[11] Orthography describes the conventional spelling of a language.
[12] Agglutinative languages contain root words that may include several short morphemes (Land, 2015).

translation of the Bible, his term of functional equivalence is upheld today (*cf*. Bermann & Porter, 2014; Chunhua, 2014; Fengling, 2017). The successful translation of text is determined by the receptors of the translated text, this means that the translated text should be comprehended by the persons reading it as if they were the original receptors reading the original text (Nida & Taber, 1969).

Functional equivalence is able to deal with some of the threats to validity "by ensuring that the instrument and elicitation method allow examination of the same construct" (Peña, 2007, p. 1257). When a piece of text is translated into the TL, it is entirely plausible that the linguistic equivalence of that text may contain biases or insensitivity towards the TL if there are functional differences between the TL and SL (Rogler, 1999). Functional differences indicate that translation from the TL to the SL may result in incongruity in the meaning of the text. This mismatch in meaning threatens the content validity of that text. It is therefore necessary to employ functional equivalence when translating a text to ensure that the same construct is assessed during an assessment. There are two main ways to obtain functional equivalence, the first being *decentering* and the second a *dual-focus approach*.

Translators make use of the *decentering* approach in order to obtain equivalence in meaning between different texts (Sechrest et al., 1972). Decentering also uses the translation back-translation method. The decentering approach is used to avoid construct bias in ILSAs or cross-cultural studies (*cf*. Boer & Fischer, 2010; van de Vijver & Leung, 1997) as it takes into consideration inputs from different cultures. When a text is translated using the decentering approach, the translator may produce a text that has changed from the original words in the ST to represent the words and concepts of that text in a linguistically familiar way in the TL (Peña, 2007), meaning that the wording may have changed to enhance the translation (Weiner, 2003).

The *dual-focus approach*, as explained by Erkut et al. (1999), employs a bilingual and bicultural team consisting of two groups of professionals, which include cultural and linguistic translators, to deal with the translation process. The dual-focus approach consists of two major features, namely horizontal collaboration and concept-driven translation. Horizontal collaboration includes the researchers from indigenous cultures as part of the team and concept-driven translation, rather than translation-driven, enables the translators to produce a translated text that is conceptually equivalent across the two languages (Erkut, et al., 1999; Gathercole, 2013). In essence, this approach uses the knowledge and experience from both the cultural and linguistic groups to perform a translation.

A key aspect of any translation is coherence between the translated text and the source text as well as the style of said text. Translators should strive to produce a translation that is equivalent. According to Chunhua (2014, p. 153), "a good translation is not only a version that is grammatically and lexically coherent with the original text, but also one that is correspondent to what the author intends to express". In terms of functional equivalence, the translated text should be stylistically the same as the original as any changes, whether it be improvement or decline, would violate key principles of functional equivalence.

Furthermore, there are three key aspects to ensuring that a translated text is the best possible translation from a source text in order to obtain functional equivalence. The three aspects include target language and culture, original language and culture, as well as the target audience (Chunhua, 2014). When functional equivalence is adhered to, it means that the learners taking an assessment that has been translated into different languages, are able to demonstrate their abilities in a context that is familiar to them (Peña, 2001). In essence, functional equivalence[13] calls attention to the information contained in the text rather than the direct formal equivalence in translation (Fengling, 2017). In Rasch theory, functional equivalence could be confirmed by using Differential Item Functioning (DIF).

3.6.3 Cultural Equivalence: Same Meaning for Each Group

Cultural equivalence is defined as having the same meaning and relevance of the construct in more than one culture (Chan & So, 2017). A study conducted by van der Veer et al. (2003) examined the meaning shift of items across different languages. They found that items may have different salience for the different groups of participants based on language and culture. More recently, du Plessis et al. (2015) found that cross-cultural differences have, to some extent, some bearing on test items. Their study indicated that there are distinctions that occur solely within a cultural grouping which create difficulties when studying a phenomenon across cultures. These two studies have shown that the same test items could be differently understood

[13] There are four key aspects of functional equivalence: lexical, sentence, passage and stylistic (Fengling, 2017).

or interpreted based on culture[14] and language. These differences in salience, may partially be due to how cultural groups interpret concepts (Usunier, 2011).

Cultural equivalence is important when text or an instrument is used for different cultural and linguistic groups. The focal point of cultural equivalence is the manner in which members of different cultural and linguistic groups understand the underlying meaning of an item or text (Peña, 2007). It is possible that the understanding of the underlying meaning or construct may be influenced by how the persons from these diverse backgrounds respond to the items. Moreover, there are other culturally determined definitions and understandings of aspects such as knowledge, creativity and language. These aspects also have some bearing on individuals' responses on items from different cultural and linguistic backgrounds. In other words, different parts of the world and indeed cultures have different theories of characteristics, like intelligence. For example, Western theories of self are based on specific cultural assumptions and may be different to those from the East and Africa. African cultures focus on community and *Ubuntu* while Western cultures, according to Boucher (2010), have an intense desire for self-enhancement. Eastern cultures also focuses more on communal traits. Nida considered cultural equivalence to be important for all translations, as "for truly successful translating, biculturalism is even more important than bilingualism, since words only have meaning in terms of the cultures in which they function" (Nida, 1993, p. 110.)

3.6.4 Metric Equivalence: Item Difficulty across Different Languages

Metric equivalence concerns equivalence in item difficulty and is crucial in assessment instruments when assessment is undertaken in more than one language (Peña, 2007). Often with ILSAs, such as the Progress in International Reading Literacy Study (PIRLS), an assessment instrument is developed in English and then sent to participating countries in order to make necessary adaptations and translations (Mullis & Prendergast, 2017) so that those countries can participate in their national languages. It is crucial that the test items be at the same difficulty level across the different test languages and it often gives rise to methodological complexities. Fischer et al. (2018) argue that even though one of the major goals of ILSAs is to explain why learners are performing differently across countries, most research focuses on

[14] Culture is a complex concept as it has multiple meanings, is widely used across different disciplines (Rashed, 2013) and is determined by different cultural groups based on knowledge, beliefs, arts, traditions, customs, values and language.

pedagogic or systemic factors and not necessarily translation factors. They contend that testing for measurement invariance is overlooked but that a possible reason why these differences occur between countries may be reliant on the learners' cultural similarity or dissimilarity (Fischer et al., 2018).

For most ILSAs, the assessment instrument relies heavily on an English corpus as the source test (and language) and selections are made for test item difficulty based on English. This means test items are drafted in English and, after approval by the stakeholders, are then sent for adaptations and translations (*cf.* Mullis & Prendergast, 2017). Thus, the original item difficulty is gauged in one language. Peña (2007) cautions that with any adaptation or translation, the issue of dialects may arise. She further cautions that to start with only an English corpus would not be appropriate, as words have different frequency levels in other languages than in English.

A study by Stubbe (2011) examined how different versions of a test instrument function when translated into a single language and found that even though only one language, German, was used, there was measurement invariance, especially with those items with differing translations. Based on these differing translations, it was found that one version of the same test instrument was easier than the other two. This example of assessment of instrument translation highlights the importance of having metric equivalence across all assessment instruments.

One way of ensuring metric equivalence could be to develop a source language and target language word list, such as the one Tamayo created in 1987 (Tamayo, 1987). The list had English and Spanish words that matched on either translation of the word or on the basis of the frequency of the words. However, this exercise may not always be achievable as not all languages have this kind of corpus linguistic data available. Peña (2007) explains that researchers and item developers should consider lexical frequency in all the target languages if no psychometric data is available.

When any form of assessment is created for multiple languages, it is worthwhile to consider that the typology of the languages, word frequencies and word classes could be different. In her research, Ntshangase-Mtolo (2009) explained that learners who speak isiZulu, might have difficulty with certain English concepts, such as '*a bearing* (direction)' or '*hypothesis*', as isiZulu lacks equivalent terms which speaks to the translatability of English into isiZulu. Other

researchers have also looked at this issue as there is an underdevelopment of the isiZulu lexicon and standardisation of the language (*cf*. Foley 2002; Murray 2002). However, Gauton and Schryver (2004) found a way in which English to isiZulu translators could solve the non-equivalence issue, by using computerised corpora methodology.

As this study aims to determine whether the PIRLS Literacy 2016 and PIRLS 2016 assessments are equivalent across English, Afrikaans and isiZulu, it was necessary to discuss the four types of equivalence considered for this study. Descriptive and Rasch analysis are used to determine the functional, cultural and metric equivalence. Furthermore, this study makes use of qualitative data collection techniques such as workshop notes, open-ended questionnaires and discussions with language experts to determine cultural and linguistic equivalence.

3.7 CONCEPTUAL FRAMEWORK FOR THE STUDY

The conceptual framework for this study is informed by the main research question:

To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?

The study's main research question addresses the equivalence of the PIRLS Literacy 2016 and PIRLS 2016 released texts across three languages. The four types of equivalences investigated in this study include linguistic, functional, cultural and metric equivalence. By examining the released texts across English, Afrikaans and isiZulu, the study may determine whether the translated versions of the ST, within the South African context, are indeed equivalent. To put it another way, if any of the four equivalences were found to be low across the English, Afrikaans and isiZulu PIRLS texts, such differences may have affected the learners' performance negatively. For that reason, it is essential to use a conceptual framework that is robust enough to focus specifically on linguistic, functional, cultural and metric equivalences of assessment instruments in relation to the Grades 4 and 5 learners' performance. The conceptual framework (*cf*. Figure 3.1) underpinning this study emphasises equivalence across-the-board for cross-cultural research and learner achievement. This study's conceptual framework draws mainly on the work by Peña (2007) where she identifies four considerations for cross-cultural research, particularly in translations of text. The conceptual framework is also linked to the design of this study (*cf*. Chapter Four). Figure 3.1 shows the schema of the conceptual framework:

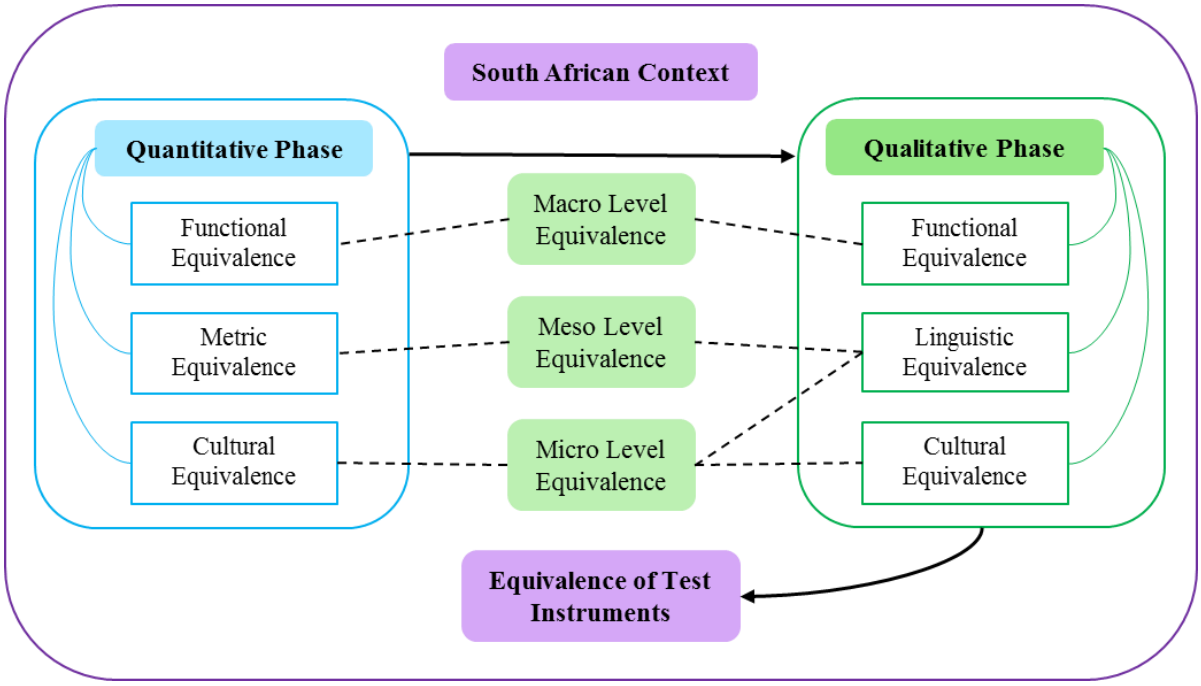

Figure 3.1: Translation Equivalence for Cross-Cultural Research (informed by Peña, 2007)

This study extends on Peña's (2007) four considerations for cross-cultural research to judge the equivalence of PIRLS assessments in South Africa. In Section 3.6, each of the four considerations were discussed. In summary, *functional equivalence* indicates that the assessment instrument will produce the same behaviour across different groups (*cf.* Greenfield et al., 2006); *metric equivalence* refers to the difficulty level of the specific item expressed in two distinct languages (*cf.* Kim et al., 2003); *cultural equivalence* considers how readers will interpret a given direction or test item and develops items that infer the same cultural meaning for each cultural linguistic group (*cf.* Chan & So, 2017); *linguistic equivalence* concerns translating instructions and instruments and checking the translation using the back-translation method (*cf.* Chesterman, 2016). Peña's (2007) research and theories regarding cross-cultural translation are based on equivalence and relate to bias and validity of instruments in different cultural settings. Section 3.7.1 deals with the quantitative phase (Phase One) of the conceptual framework and discusses the importance of functional and metric equivalence while Section 3.7.2 discusses the qualitative phase (Phase Two), specifically cultural and linguistic equivalences in terms of the South African learners' results on the PIRLS Literacy 2016 and PIRLS 2016 assessments.

3.7.1 Phase One of the Conceptual Framework

A goal of this study is to determine to what extent the PIRLS Literacy 2016 and PIRLS 2016 test scores have the same meaning across English, Afrikaans and isiZulu. Van de Vijver and Leung (1997) regard this 'same meaning' as an important aspect in cross-cultural testing. One way of determining whether the test scores across the aforementioned languages are equivalent, is to investigate the scores at a technical level. According to Koch (2009), if non-equivalence or bias exists in the assessment instrument, then it cannot be used for comparison between different groups or for the language versions, as the translated texts then do not have the same meaning. When bias occurs, the equivalence or invariance of the assessment is contested.

The term 'bias' in testing refers to when the test scores differ between groups even though there are no other differences in the underlying trait (van de Vijver & Leung, 1997). It is considered as a threat to the validity of cross-cultural assessments and may include aspects such as poor translation, inappropriate content, lack of standard operation procedures during translation and the lack of standardisation of administration procedures. Van de Vijver and Leung (1997) state that there are three types of bias, namely construct, method and item bias. *Construct bias* happens when the construct that is being assessed is not the same across the different groups and that the measure (the test) favours one group only. An example of construct bias is when a test is developed to test a specific set of skills and knowledge but it includes construct irrelevant features for one group but not for the other; for instance, if a learner with poor reading literacy skills completes a mathematics test, the learner may obtain a lower mathematics score because of his/her literacy skills. *Method bias* refers to sample, instrument and administration bias and deals with methodological aspects perplexing the comparison, and determining the reasons why there are differences between the groups; for example, one group is mistakenly given more time to complete the same test as another group. It typically has an effect on the entire assessment instrument. *Item bias* concerns item level nuisances which are a measurement problem that poses a threat to the validity of cross-cultural assessments. Item bias, for instance, could occur when there are differences in the suitability of the item content, such as certain topics covered in the curriculum for one group but not the other. It could also occur when the item formulation was inadequately done or when the translation was not equivalent across the different languages.

For the purposes of this study, emphasis is placed on construct and item bias as the study aims to determine whether the selected PIRLS Literacy 2016 and PIRLS 2016 texts are equivalent

across different languages. Construct and item bias focus specifically on the text and items. Table 3.3 shows the link between this study's Phase One sub-questions and the conceptual framework.

Table 3.3: Phase One Sub-Questions and Conceptual Framework Alignment

Main Research Question	To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?	
Research Question	**Sub-question 1:** How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?	**Sub-question 2:** How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016?
Data Analysis Method	Utilising IDB-analyser to obtain percentages and means of learner achievement and conducting DIF analysis to determine whether there are differences between language groups. Focuses on item bias or non-equivalence.	Utilising IDB-analyser to obtain percentages and means of learner achievement and conducting DIF analysis to determine whether there are differences between language groups. Focuses on item bias or non-equivalence.
Link to Conceptual Framework	Functional, cultural and metric equivalence	Functional, cultural and metric equivalence

Sub-question 1 and 2 examine learner performance on the PIRLS assessments and how it differs across the three languages. Both questions focus on item non-equivalence or bias which is examined by using Rasch theory, specifically differential item functioning (DIF), to determine whether the bias is the same (uniform DIF) or not the same (non-uniform DIF) across the different ability groups. When DIF is present in an assessment instrument, it means that the items may favour either of the three languages. This research question relates to functional, cultural and metric equivalence. The purpose of Phase One of the conceptual framework is to analyse the equivalence of the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments (texts and items) across different languages in order to address the broad question of whether

one can interpret and use the test scores obtained from two or more different versions of a test in the same way.

Functional equivalence looks at the assessment instrument and determines whether it measures the same construct across different versions. As this study incorporates three of the 11 official languages of South Africa, it is important to take into consideration the cultural equivalence of the PIRLS assessments. *Cultural equivalence* is whether members of different cultural or linguistic groups understand the meaning of the text or item. In terms of *metric equivalence*, this study aims to determine whether the item difficulty across different languages is equivalent (*cf*. Peña, 2007).

As indicated above, Phase One is the quantitative part of the study and focuses on sub-questions 1 and 2 and their alignment to the study's conceptual framework. Phase Two of the study's conceptual framework is discussed in Section 3.7.2 and involves the qualitative nature of this study.

3.7.2 Phase Two of the Conceptual Framework

As stated in the previous section, this study seeks to determine whether the versions of the PIRLS Literacy 2016 and PIRLS 2016 assessments are in fact equivalent. Any ILSA that is conducted in more than one language has specific translation purposes, especially in a reading comprehension test such as PIRLS, where each text has its own translation purpose. For example, a fable will not be translated in the same manner as an informational text as the purpose, structure and language features are different.

In terms of ILSAs, the translation company should be well informed about the type of translation that is required (Arffman, 2013). ILSAs should create task specifications which contain information about the assessment, purpose of the translation, as well as other considerations such as the target language, context age group and/or reading level. The IEA provides such task specifications in the Survey Operations Procedures (SOP) for PIRLS Literacy 2016 and PIRLS 2016. The aforementioned SOP explicitly details how the translator should handle the texts and items, and that all adaptations must be recorded in the National Adaptation Forms (NAFs). The direct translations are not required to be recorded in the NAFs. The SOP further gives clarity in terms of preparation for translation, actual translation and that the translations, with the NAFs, should be sent to the reviewer.

This phase of the conceptual framework focuses on linguistic and cultural equivalence. To recap, *Linguistic equivalence* refers to the linguistic meaning of the words used in the assessment instruments and whether they are the same for the different versions (Peña, 2007). *Cultural equivalence* looks at how different cultural groups understand the underlying construct of the text (Peña, 2007). For example, if two or more cultural groups write an assessment, would they understand the construct in the same manner. Table 3.4 (following page) depicts the Phase Two sub-questions and how they link with the proposed conceptual framework.

Sub-question 3 deals with the construct non-equivalence of the PIRLS assessments and relates to the face validity of the assessments. Face validity concerns subjective judgement of an assessment instrument to determine whether it appears valid (Martinez, 2017; Mostert, 2007). As discussed in Section 3.4.3.2, the PIRLS assessments are examined for face value, whether they look similar to other reading assessments for Grade 4 and 5 learners as well as to determine whether the texts used in the PIRLS assessments are comparable to those with which learners are familiar. Moreover, this sub-question links to the functional equivalence of this study's conceptual framework as it will compare the PIRLS assessments to what is required in the national curriculum.

Table 3.4: Phase Two Sub-Question and Conceptual Framework Alignment

Main Research Question	To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?		
Research Question	**Sub-question 3:** To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?	**Sub-question 4:** How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?	**Sub-question 5:** How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?
Data Analysis Method	Makes use of comparison between the PIRLS assessments and what is required in CAPS. Using questionnaires completed by experts. Focuses on construct bias or non-equivalence.	Utilising workshop notes, discussions, questionnaires and text notes to determine whether the texts are equivalent. Focuses on content bias or non-equivalence.	Utilising workshop notes, discussions, questionnaires and text notes to determine whether the texts are equivalent. Focuses on content bias or non-equivalence.
Link to Conceptual Framework	Functional equivalence	Linguistic and cultural equivalence	Linguistic and cultural equivalence

Research sub-questions 4 and 5 concern the textual coherence of the PIRLS texts by looking at macro, meso and micro levels of equivalence. In terms of *macro* level of equivalence, this study provided an overview of the PIRLS texts which describes text characteristics such as genre and themes. At this level, the layout of each of the texts as well as the length of the text is considered. The *meso* level includes passage mapping for each text as well as the level of comprehension assessed for each item. The *micro* level focuses on aspects such as the vocabulary used in the text, rhetorical devices used, as well as readability of the texts, all of which are further unpacked in Chapter Seven and Eight. The levels of equivalence link with the linguistic and cultural equivalence of this study's conceptual framework as these types of equivalences focus on whether the text is the same across the SL and TL. This study looks at the textual coherence of the PIRLS texts in conjunction to the selected languages (English, Afrikaans and isiZulu) to determine whether each of these versions are equivalent to one another.

3.7.3 Equivalence of Test Instruments

The evaluation of equivalence of the PIRLS assessments instruments is the outcome for this study's conceptual framework. As mentioned earlier, this study's conceptual framework extends the cross-cultural equivalence theory put forward by Peña (2007), to look at the equivalence and validity of ILSAs, specifically PIRLS 2016, in the South African context. By looking at metric, functional, linguistic and cultural equivalences in the translation of the PIRLS Literacy 2016 and PIRLS 2016 texts, it may be possible to determine whether the translated versions, Afrikaans and isiZulu, are equivalent to the source version, English.

The examination of the PIRLS texts in terms of functional and metric equivalences determine whether these texts behaved similarly across the language groups. This is done by quantitative measures such as percentages, means and Rasch measurement theory. Any anomalies found during Phase One of the study are then to be addressed by exploring the linguistic and cultural equivalences of those texts in Phase Two. By applying the tenets of translation equivalence in this context, it is possible to determine whether the translations of the PIRLS Literacy 2016 and PIRLS 2016 texts were equivalent in English, Afrikaans and isiZulu.

To recap: the main research question of this study asks: *To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?*

In terms of the conceptual framework, five subsidiary questions are posed which assist in addressing the main research question in two phases. These questions are as follows:

1. *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?* This sub-question investigates whether there are any differences in the English, Afrikaans and isiZulu learners' achievement on the PIRLS Literacy 2016 assessments by conducting quantitative data analysis using techniques such as percentages, means and differential item functioning. This question links with functional, cultural and metric equivalence and focuses on the possible item bias across the different versions.
2. *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016?* See above explanation.
3. *To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?* The third sub-question looks at the PIRLS assessment instruments in relation to what is required by the national curriculum in terms of texts and tests. This question relates to functional equivalence and looks at construct bias and non-equivalence.
4. *How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?* The fourth research sub-question entails exploring the three levels of equivalence in order to determine whether the selected texts have textual coherence. This sub-question links with both linguistic and cultural equivalence and focuses on content bias.
5. *How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?* See above explanation.

3.8 CHAPTER SUMMARY

This chapter discussed various aspects pertaining to the equivalence of assessment instruments of international large-scale studies such as PIRLS. Various researchers have, over time, examined the translations of tests, but more recently the importance of cross-cultural research has taken centre stage (*cf.* du Plessis et al., 2015; Erkut et al., 1999; Fischer et al., 2018; Geisinger, 1994; Peña, 2007). As more countries, especially developing countries like South Africa, participate in ILSAs, it is crucial that equivalent assessment instruments are developed in order to ensure that the original or source text are comparable to the versions across different

participating countries. South Africa has participated in three rounds of PIRLS and is currently in the development phase for the fourth round of PIRLS. Given the emphasis placed by the President of South Africa in his state of the nation address on the importance of reading literacy (South African Government, 2019), with the goal that every 10-year old should be able to read for meaning, this study aims to determine whether the PIRLS Literacy 2016 and PIRLS 2016 assessments are equivalent across different languages and to discuss the possible impact it has on the South African learners' achievement.

-⌘-

CHAPTER 4

Research Design and Methodology

4.1 ORIENTATION

This chapter presents the research design and methodological endeavours that address the research questions posed by this study. Section 4.2 focuses on the research paradigm of the study. Thereafter, the research design for the study is discussed (4.3), followed by the methodological norms that underpin the study (4.4). The ethical considerations for the study are then described (4.5) and the last section (4.6) deals with concluding comments.

4.2 RESEARCH PARADIGM

This study utilised the sequential explanatory mixed methods approach and will first address the ontological and epistemological views of this study (4.2.1). The usefulness of quantitative and qualitative research is argued (4.2.2) which is followed by a discussion about pragmatism as the study's research paradigm in Section 4.2.3.

4.2.1 Ontological and Epistemological Views

Mono-method researchers that use either quantitative or qualitative methods differ not only systematically but also philosophically in their approaches to measurement. Quantitative researchers gather and analyse numerical data in order to perform intricate statistical models to address their research questions whereas qualitative researchers conduct in-depth, 'wordy' research about a specific phenomenon, discussing the meaning of concepts and patterns of interaction and behaviours (Goertz & Mahoney, 2012). Maykut and Morehouse (1994) explain that

> these two paradigms are based on two different and competing ways of understanding the world. As we observe and understand, these competing ways of comprehending the world are reflected in the way research data is collected (words versus numbers) and the perspective of the researcher (perspectival versus objective) (p. 16).

For this research, an alternative to mono-methods was sought. Mixed methods research was selected for the current study as it makes use of both quantitative and qualitative methods to address this study's complex research question regarding the equivalence of the PIRLS

Literacy 2016 and PIRLS 2016 release texts[15] and items across English, Afrikaans and isiZulu languages. The nature of this study requires both quantitative and qualitative methods as it first analysed learners' performance on the PIRLS 2016 cycle to determine possible non-equivalence which was then further explored in the qualitative phase. This section discusses ontology and epistemology and how it links with mixed methods research.

Ontology is a branch of metaphysics that focuses on the study of the nature of reality thus identifying things that exist. In other words, ontology "is concerned with the nature of what exists" (Blaikie, 2004, p. 767). Ontology is the study of different theories about what makes up reality. In terms of social science, theories and methodological views create and shape assumptions about reality, how things exist as well as how things are related (Blaikie, 2011). As such, when a researcher examines a concept, specifically the meaning thereof, the researcher is actually asking about the nature of reality (Goertz & Mahoney, 2012).

Quantitative researchers discuss and debate concepts in terms of data and measurement. Although some discussion of the concept is necessary, less focus is placed on the meaning of the concept. These researchers typically look at the 'operationalisation' and 'measurement' of any given concept (Goertz & Mahoney, 2012). Operationalisation comprises indicators derived from numerical data that indicate some relationship between one another and with the latent variable. These indicators can be used in statistical analysis to generate possible models that generate scores in conjunction with the concept of interest. In contrast, qualitative researchers prefer to discuss and debate the concepts, finding the meaning of those concepts rather than using numerical data or representative samples. Qualitative researchers ask questions about and define concepts; moreover, they specify and list the specific characteristics that make up the concept (Goertz & Mahoney, 2012).

Epistemology is narrowly defined as the theory of knowledge. In other words, "how human beings come to have knowledge of the world around them – of how we know what we know" (Blaikie, 2011, p. 310). Specifically, it studies the nature of knowledge as well as the validity thereof and the difference between opinion and justification (Hathcoat & Nicholas, 2014).

[15] The PIRLS Literacy 2016 and PIRLS 2016 released texts are texts selected from the assessments to be part of the Reader booklet that learners may keep as a classroom resource. The released texts are also available to use by scholars and may be downloaded from the PIRLS website. However, only the original English version is available. To make use of the translated texts, the scholar should apply to the IEA and national NRC for permission to use any of the translated texts.

Epistemology gives a foundation where human beings can decide what kinds of knowledge is possible in order to make informed decisions about how knowledge can be judged, verified and validated (Blaikie, 2011). Furthermore, epistemological considerations inform and generate assumptions regarding the manner in which the research should be conducted. It also assists in selecting appropriate methodological approaches and guides an understanding of the knowledge obtained by the investigation (Hathcoat & Nicholas, 2014).

Crotty (1998) points out that it is not easy to discuss these philosophical positions separately as ontological and epistemological claims are linked. Ontology addresses philosophical questions about the nature of reality and truth whereas epistemology addresses questions about how human beings understand knowledge and reality. Furthermore, epistemology, focusing on knowledge and defining phenomena, has different claims to ontology but according to Blaikie (2011), the differences in the claims cannot be resolved by empirical enquiry, but they can be philosophically debated. He further explains that the researchers who employ ontology and epistemology, "make their claims as an act of faith" (Blaikie, 2004, p. 767).

In mono-methods research, detailed explanations are provided about how these mono-methods should be used and how they differ from each other (Bryman, 1988; 2001; Creswell, 2003; Tashakkori & Teddlie, 2003). Furthermore, mono-methods are characterised based on their ontological and epistemological concerns (*cf.* Crotty, 1998; Lincoln & Guba, 1985). Based on the nature of both ontology and epistemology, different paradigms exist in order to assist in finding answers to questions. Paradigm 'wars' ensued because of the differences outlined above between ontology and epistemology (Bryman 2008; Creamer, 2018; Pinto, 2010). To make mixed methods research viable, a way needed to be found to bridge these differences (Lincoln & Guba, 1985; Creswell & Plano Clark, 2008). Four rationales for conducting mixed methods research were proposed by Onwuegbuzie and Collins (2007): participant amelioration (sample optimization), instrument fidelity (appropriateness of instruments), treatment integrity (trustworthiness of intervention), and significance enhancements (supplementing interpretation of findings). During the paradigm 'wars', Schwandt (2000) took a hard stance on the importance of questioning mono-methods research:

> All research is interpretive, and we face a multiplicity of methods that are suitable for different kinds of understandings. So the traditional means of coming to grips with one's identity as a researcher by aligning oneself with a particular set of methods (or being defined in one's department as a student of "qualitative"

> or "quantitative" methods) is no longer very useful. If we are to go forward, we need to get rid of that distinction (p. 210).

Quantitative and qualitative researchers, based on their worldviews, have different assumptions about the nature of knowledge, which in turn makes it difficult to cross between them. As a result, some quantitative and qualitative researchers reject the principles of each other's paradigms (Morgan, 2007). However, for the purposes of this study, a combination of both quantitative and qualitative approaches was required to address the study's complex research question. I agree with Schwandt that there should be an alternative to mono-method research; this is where mixed methods could provide a suitable alternative. Even though the main data source of this study, namely PIRLS Literacy 2016 and PIRLS 2016, comes from a measurement view where the reality is measurable, the issues surrounding the results of the South African learners need to be investigated beyond league tables in order to determine whether the texts were equivalent across the three selected languages. By using mixed methods research, I was able to utilise qualitative methods to help explain the findings from the quantitative phase and ultimately interpret what was found. Morgan (2007) explained that a new alternative paradigm was needed that could link the quantitative and qualitative fold, namely pragmatism. Even though pragmatism is not completely new, Onwuegbuzie and Johnson (2006) state that there is justification for combining quantitative and qualitative approaches to address research questions as it allows new methods in dealing with methodological issues. By working through the pragmatism worldview, I was able to utilise both quantitative and qualitative methods as I believe that no one single approach is sufficient in telling a story or finding the truth. Sections 4.2.2 and 4.2.3, which follow, discuss the complementarity of quantitative and qualitative research as well as epistemological position for this study.

4.2.2 The Complementarity of Quantitative and Qualitative Research

In Western philosophy, debates about universal truths to understanding the world (Socrates, Plato), as opposed to relative truths (Protagoras) and against a balance of the extreme views (Aristotle), is still observable in current research approaches (Bergman, 2008). When mixed methods research is used, it means that the researcher is situated between the quantitative (Plato) and qualitative (Protagoras) extremes by attempting to use both approaches to address the research problem. Quantitative and qualitative researchers, as mentioned in the previous section, have different worldviews and approaches to finding solutions to a research problem.

The former relies upon the mathematical importance of data and using sophisticated statistical tools to analyse it. The latter relies on understanding the experiences of persons in situations and to richly describe these.

Over the years, an interest in mixed methods research has grown (Archibald, Radil, Zhang & Hanson, 2015) given the limitations of mono-methods. Using both quantitative and qualitative methods provides the opportunity to overcome these limitations in order to address complex research questions (Plano Clark & Ivankova, 2016). By using both quantitative and qualitative methods, it is possible to address a research problem from different perspectives where the researcher can conduct hypothesis testing and inferential analysis but also explore the research problem further by conducting interviews and investigating the themes which emerged during data analysis. All researchers who use either mono-methods or mixed methods need justifiable reasons why they select a particular approach for their research. Methodologists have written extensively on the reasons why mixed methods would be an appropriate approach to conduct research (Bergman, 2008; Johnson, Onwuegbuzie & Turner, 2007).

Bergman (2008) examines the different qualities attributed to quantitative and qualitative methods and explains why these mono-methods were described as being incompatible based on the arguments from the originators that lead to the 'paradigm wars'. He then comments that "… researchers engaging in mixed methods design have to maintain a strangely schizophrenic position toward the division of labor [sic] between QL and QN methods" (p. 13).

Mixed methods advocates have to accept the opposing qualities ascribed to the quantitative and qualitative paradigms[16], yet propose that the strong point(s) of both can be merged into a single research design. It then allows the mixed methods researcher to make use of a repertoire of data collection and analysis techniques to solve the research problem. Bergman (2008) continues to ask scholars and researchers to 'rethink' the separation of quantitative and qualitative methods as these can viably be used in mixed methods research. He outlined four principles that could aid in using both quantitative and qualitative paradigms: (a) differentiate between quantitative and qualitative data collection and data analysis methods; (b) being aware of the inductive and deductive features of the different analytic phases of the study; (c) bearing in mind the researcher and her habits alongside the different data collection and analysis

[16] Based on ontological, epistemological and axiological differences.

methods; and (d) if possible, to refrain from expounding on the limits of either methods as it brings about the binary classification of these methods. As such, mixed methods advocates should rather focus on justifying using the different methods in reference to their research problem, theoretical framework and data needs. The justification for using mixed methods in this study, is based on the complex research problem, that is, determining the equivalence of the PIRLS Literacy 2016[17] and PIRLS 2016[18] assessments within the South African context by specifically looking at the English, Afrikaans and isiZulu versions of the assessments. The PIRLS assessments align with a measurement view where the reality is measurable; however, by employing mixed methods research and using the pragmatist lens, I introduced a qualitative view which implies that reality can be interpreted too.

Utilising both quantitative and qualitative techniques may provide the researcher with a better understanding of the problem than either of the techniques could give individually (Creswell & Plano Clark, 2011). When quantitative and qualitative methods are used together, it can provide validity through 'complementarity'. Complementarity could be understood as the "elaboration, illustration, enhancement, and clarification of the findings from one method with the results from the other method" (Onwuegbuzie & Leech, 2004, p. 770). Hence, the aim for this study was to select a research method that best allowed the researcher to address the research question that is not based solely on the presumed ontological and epistemological link to the research question. According to Bergman (2008), mixed methods research does not allow one to bridge the perceived gap between the two mono-methods, but rather offers an alternative method to conduct research for a specific problem under specific circumstances.

When researchers employ mixed methods research, the main philosophy is pragmatism as it focuses on gaining knowledge from multiple views, positions and standpoints (Bergman, 2008). Taking the above discussion into account, the next section provides the rationale for selecting pragmatism as the study's research paradigm.

[17] PIRLS Literacy 2016 texts analysed in this study: The Pearl, African Rhinos and the Oxpecker Birds and Flowers on the Roof.
[18] PIRLS 2016 texts analysed in this study: Macy and the Red Hen, The Green Sea Turtle's Journey of a Lifetime and Flowers on the Roof.

4.2.3 Pragmatic Research Paradigm

In light of the argument for the complementarity of quantitative and qualitative research methods as outlined above, and given the extensiveness of the research questions, mixed methods design was selected as the design for this study. As the popularity of mixed methods research grows (Tashakkori & Teddlie, 2010), it has led to discussions about the 'pluralism' of how researchers conduct mixed methods research. As a result, an assortment of mixed methods typologies was suggested (*cf.* Creswell & Plano Clark, 2008; Johnson et al., 2007; Teddlie & Tashakkori, 2009). Several typologies focus on methodological issues[19], which is discussed later.

After the paradigm 'wars' of the 1970s and 1980s, mixed methods was accepted as a "third methodological movement" (Hall, 2013, p. 72), which has led to searching for a paradigm that justifies and legitimises the use of both quantitative and qualitative methods as one research design. Some researchers have suggested that pragmatism is a suitable research paradigm for mixed methods research (Feilzer, 2010; Johnson et al., 2007; Morgan; 2007).

Pragmatism, according to Onwuegbuzie et al. (2009), is only one of the various philosophical stances that underpin mixed methods research. Yet pragmatism searches for some common ground between philosophical dogmatisms (Johnson & Onwuegbuzie, 2004) that enables researchers, on a philosophical level, to use both quantitative and qualitative methods to build on current knowledge (Doyle et al., 2016). Pragmatism was first documented by William James in his book, *Pragmatism,* in 1907 (Morgan, 2014) with other main pragmatists including Peirce, James, Mead and Dewey (Scheffler, 2011). This study establishes itself with these pragmatists. Pragmatism is a philosophical system that focuses on what works in a particular situation. A key aspect of pragmatism is "that the meaning of any concept is determined by its practical implications" (Hammersley, 2011, p. 848).

In recent years, pragmatism has been seen as a primary philosophy when conducting mixed methods research as it is takes into consideration multiple viewpoints and positions (Johnson et al., 2007; Morgan, 2007). Furthermore, pragmatism emphasises multiple realities where meaning and knowledge are tentative and ever changing. Knowledge is constructed and based upon the experience and reality of the world. Even though the approach to pragmatism can be

[19] See Section 4.3.1 for a discussion on mixed methods research that includes the main topologies.

dialectical and pluralistic, it focuses on determining what could be used to best solve a research problem. In terms of pragmatism, knowledge is amassed by inquiry, belief as well as the improvement on past understandings (Onwuegbuzie et al., 2009). Perhaps more importantly, pragmatism "takes an explicitly value-oriented approach to research that is derived from cultural values; specifically endorses shared values such as democracy, freedom, equality and progress" (Onwuegbuzie et al., 2009, p. 123).

In addition to pragmatism, the dialectic stance within a pragmatism epistemological paradigm is adopted for this study. It is generally accepted that a researcher within this philosophical situation would opt to select the most appropriate method to address the research problem (Greene & Caracelli, 1997). As such, the dialectic stance presumes that the use of multiple paradigms during the research process provides a better understanding of the problem being researched (Green & Hall, 2010). Greene and Caracelli (1997) argue that the purpose of the dialectic stance is to gain insight into the research problem. Quantitative and qualitative methods were thus incorporated into this study, taking the form of a sequential explanatory mixed methods research design.

4.3 RESEARCH DESIGN

This section describes the research design for this study. The selected mixed method research design as well as the reasons for selecting it are explained in Section 4.3.1. Thereafter, the study's research questions are posed (4.3.2). The research methods, sampling, data collection and analysis procedures for Phase One (quantitative) of the study are discussed (4.3.3) followed by a discussion of the same methodological aspects for Phase Two (qualitative) of the study (4.3.4).

4.3.1 Mixed Methods as Research Design

This section is divided into two sub-sections. The first sub-section (4.3.1.1) provides the rationale for selecting mixed methods as the research design for this study, followed by elucidation of the specific mixed methods design selected (4.3.1.2).

4.3.1.1 Rationale for Selecting Mixed Methods Research Design

An advantage of using mixed methods is that it enables the researcher to make use of multiple viewpoints, perspectives and positions (Johnson et al., 2007). Based on the different mixed

method designs available, there are also different typologies or rationales for using mixed methods research, such as those discussed extensively by Plano Clark and Ivankova (2016). This study focuses on Greene, Caracelli and Graham's (1989) mixed method typology that has five purposes for conducting mixed methods research.

The first purpose, namely *triangulation* within mixed methods research, seeks "convergence and corroboration of results from different methods studying the same phenomenon" (Johnson et al., 2007, p. 115). Webb, Campbell, Schwartz and Secherest (1966) coined the term 'triangulation' in the 1960s. The second purpose, namely *complementarity*, which was discussed earlier[20], seeks amongst others, the elaboration of both mono-method results. Making use of the information contained in one method to inform the other method is the third purpose, namely *development*, where the researcher is able to draw conclusions that are more meaningful. The fourth purpose is *initiation*, or uncovering paradoxes and disagreements, which may result in amending the research question. The last purpose is that of *expansion* and entails the increase of the range of inquiry by means of using diverse methods for the various inquiry components (Johnson et al., 2007; Plano Clark & Ivankova, 2016).

Other than initiation, the remaining four purposes played either a major or minor role in this study. There are, according to Denzin (1978), four types of triangulation, namely data triangulation (utilising a range of sources in a research study), investigator triangulation (utilising several researchers), theory triangulation (utilising multiple perspectives and theories to interpret results) and methodological triangulation (utilising multiple methods to study a specific research problem). The latter consists of two types, namely simultaneous and sequential triangulation (Morse, 1991). This study made use of sequential (methodological) triangulation as it utilises the results from one approach to inform the other.

In terms of complementarity, this study made use of both quantitative and qualitative research to complement the findings of each (Plano Clark & Ivankova, 2016). The research design also made use of development as it enables the researcher to utilise the benefits of a large-scale study, its sample and results to inform the qualitative method by selecting items that displayed differential functioning across the English, Afrikaans and isiZulu language groups. As development is selected, it links with sequential mixed method studies (Creswell & Plano

[20] See Section 4.2.2.

Clark, 2011). The last purpose used in this mixed method research design is expansion. It adds to the scope and extensiveness of the study and to the body of knowledge as the study examined the PIRLS Literacy 2016 and PIRLS 2016 released texts to understand the complexities of the phenomenon (Onwuegbuzie & Collins, 2007), for example, equivalence of PIRLS texts across different languages. Based on the above arguments, a sequential mixed methods design was seen as appropriate as there are several research questions that require answering to understand the PIRLS texts used in the South African context. In addition, by using the aforementioned design it "will accomplish more" (Morgan, 2014, p. xiii) than either mono-method design.

4.3.1.2 Specific Mixed Methods Design Used

There is a multitude of mixed methods research designs available, depending on the needs of the research problem. As each research problem is unique, different typologies exist (with different complexity levels) which can be utilised to address the specific research problem within mixed methods research. Ivankova, Creswell and Stick's (2006) typology on sequential explanatory mixed method research is represented in three methodological dimensions: *priority*, *implementation* and *integration*. This typology was used to assist in selecting an appropriate mixed methods design for this study.

Priority involves which approach, quantitative or qualitative, receives more weight or whether both approaches should receive equal weight. Within this specific design, the quantitative approach receives more weight as the quantitative data is first collected and analysed.

Implementation refers to whether the quantitative and the qualitative phases occur at approximately the same time, that is, independent of each other (concurrent) or if these two phases occur one after the other indicating that the second phase is dependent, to some extent, on the first phase (sequential). During a sequential explanatory design, the quantitative data are collected and analysed in the first phase of the study and the qualitative data are collected and analysed in the second phase. Additionally, when this design is selected, the research questions usually seek contextual explanation of the statistical results (Ivankova et al., 2006).

Integration involves the mixing of the quantitative and qualitative methods at a certain stage during the research process. There are a range of possibilities when mixing the two methods; the choice of which could, for example, rest on the purpose and the research questions (Teddlie & Tashakkori 2003) or integrating the findings from both methods at the interpretation stage of the study (Onwuegbuzie & Teddlie 2003). However, for the purpose of this study, the

connecting point is the development of qualitative procedures based on the results obtained from the quantitative phase in order to investigate the quantitative results in more depth (Ivankova et al., 2006). This investigation is done by collecting and analysing the qualitative data during the second phase of this study.

For the purpose of this study, and in light of the research question posed, a *sequential explanatory design* (Creswell, 2013; Creswell & Plano Clark, 2011; Ivankova et al., 2006) was adopted as it comprises two phases that occur sequentially and where the quantitative data is first collected and analysed followed by qualitative data that elaborate on the quantitative results. During the first phase, the South African Grades 4 and 5 learner results from PIRLS Literacy 2016 and PIRLS 2016 are used to describe the learners' performance. In addition, a deeper look into the learners' performance by language group is conducted to determine whether items functioned differently, indicating item bias. The second phase of the study entails the qualitative leg of the mixed methods design. During this phase, the researcher examines the PIRLS released texts based on meso, macro and micro levels of equivalence. The results obtained from the quantitative phase inform the qualitative phase of this study. Figure 4.1 presents a general schema of a sequential explanatory design (as informed by Creswell, 2013; Creswell & Plano Clark, 2011; Ivankova et al., 2006).

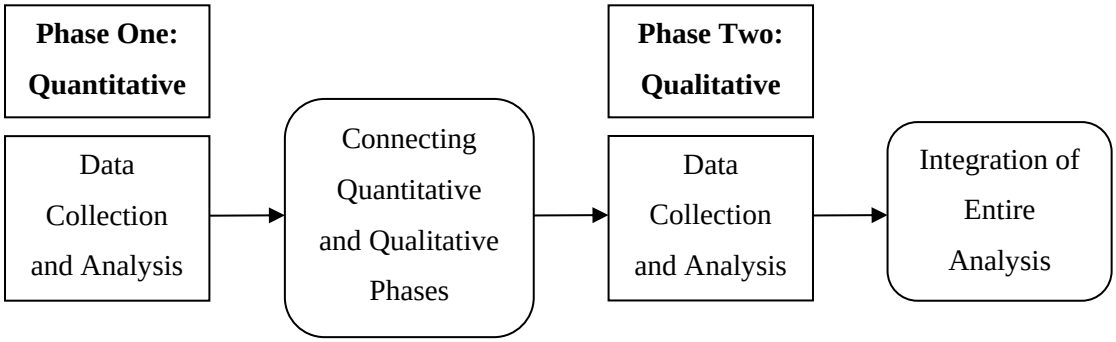

Figure 4.1: Sequential Explanatory Design

Figure 4.1 shows the general visual model of a sequential explanatory design. During this design, the researcher first collects and analyses quantitative data, thereafter the qualitative data is collected and analysed. The interpretations rests upon the qualitative data explaining the quantitative results (Creswell, 2013). A visual model for this study's sequential explanatory design is further elaborated in Figure 4.2.

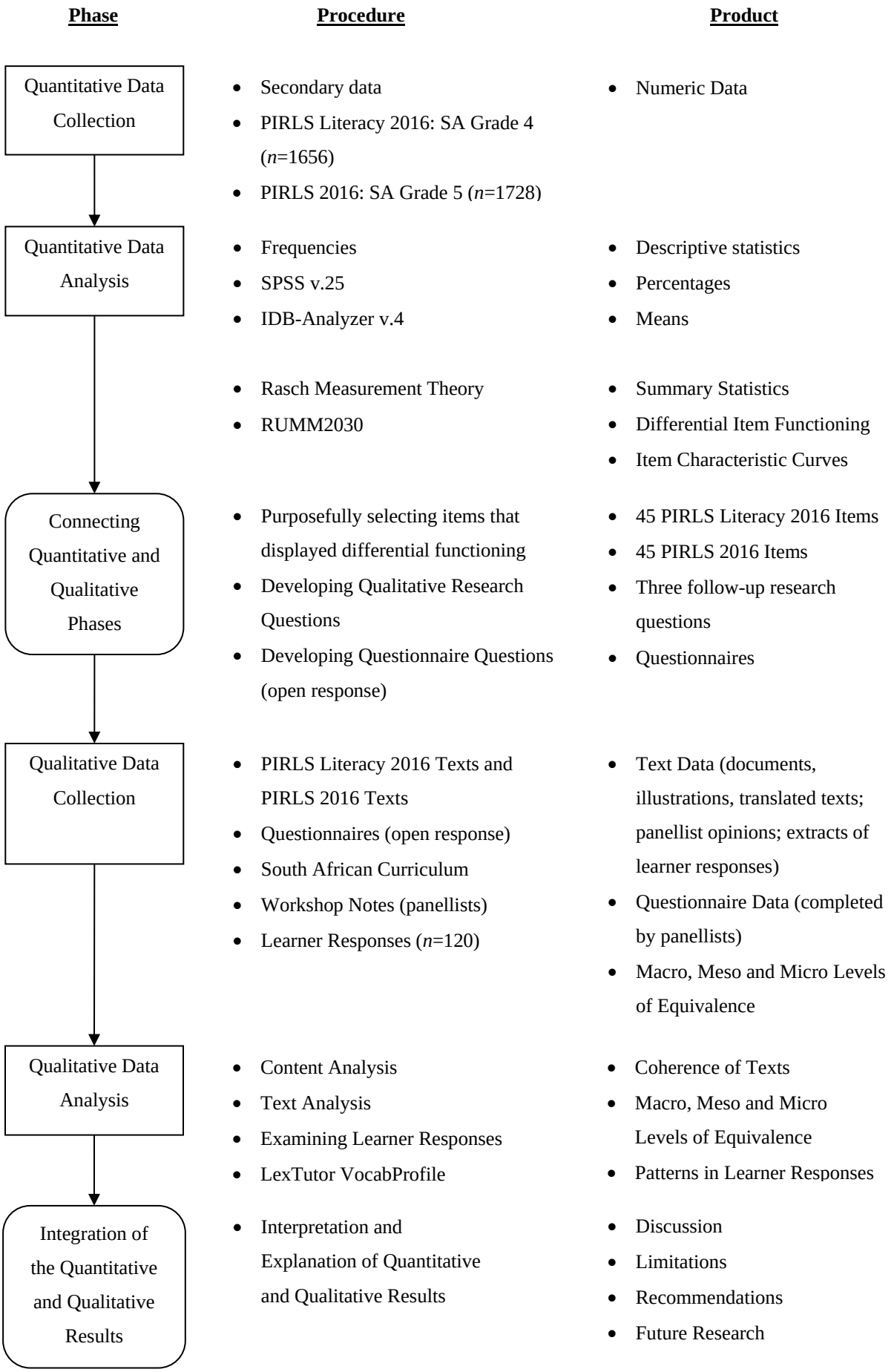

Figure 4.2: Sequential Explanatory Mixed Methods Design for This Study

4.3.2 Overview of the Study's Research Questions

This study's research design is informed by the research questions, with the study's main research question asking the following:

To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?

In order to address the main research question, five sub-research questions were posed.

- Phase One: Quantitative

Two of the five sub-questions manifest at the first quantitative phase of the research process. To reiterate, the first and second sub-questions in the quantitative phase are:

1. *How does learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?*
2. *How does learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016?*

Both these sub-questions investigate learner performance on the PIRLS assessments and how it differs across the three language groups for Grades 4 and 5 learners, respectively. The rationale for selecting the three languages is based on the number of languages tested at the Grade 5 level. Only English, Afrikaans and isiZulu were tested; as such, the same three languages were selected for the Grade 4 group. Selected data from PIRLS Literacy 2016 and PIRLS 2016 are used to provide a descriptive overview of South African Grades 4 and 5 learners' achievement by language. The differences in achievement were determined in two ways: the percent correct per item and differential item functioning (DIF) to establish item non-equivalence.

As these sub-questions focus on item non-equivalence, Rasch measurement theory was applied to determine whether differential functioning occurred on the selected items across the three languages. DIF analysis was used to establish the validity (in terms of both functional and cultural equivalence) of the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments. Language bias could be identified during the Rasch measurement. By utilising Rasch measurement, it is possible to identify test items which may threaten test validity (Wright & Stone, 1999).

- Phase 2: Qualitative

The remaining three sub-questions manifest at the second qualitative phase of this research. To recapitulate, the third sub-question deals with face validity and is as follows:

3. *To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?*

To address this question, selected PIRLS Literacy 2016 and PIRLS 2016 texts were examined and compared against the South African curriculum and assessment policy statement (CAPS) in order to determine whether the PIRLS Literacy 2016 and PIRLS 2016 texts have face validity. Wools (2015) explains that "validity is to be interpreted as a continuous property of test score interpretation" (p. 10) which could be seen as a tenet of test quality. The main goal of PIRLS is to assess learners' reading literacy levels at the Grade 4 level since it is a crucial shift from learning to read to reading to learn (Mullis & Martin, 2015). This study examined the word length requirements for Grades 4 and 5 to ascertain whether these texts are typical texts appropriate for use in the South African context.

The fourth and fifth sub-questions of this research asked the following:

4. *How does textual coherence, as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?*
5. *How does textual coherence, as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?*

The above sub-questions are multifaceted as they aimed to establish whether the textual coherence of the PIRLS released texts differ between the ST (English) and TTs (Afrikaans and isiZulu). As such, the texts were analysed by looking at *macro*, *meso* and *micro* levels of equivalence. The *macro* level provided an overview of each text which included the genre, main characters, themes, purpose and setting. This level also included a comparison of the layout of each text as well as text length. The *meso* level included passage mapping of each text as well as the difficulty of the items in terms of the PIRLS *Processes of Comprehension*. The *micro* level investigated the readability of each text including the vocabulary frequency levels, literary devices used, average word length per language and the average sentence length for each language. Thus, content analysis was conducted as it "provides knowledge and understanding of the phenomenon" (Downe-Wamboldt, 1992, p. 314).

Based on the quantitative results, a deeper look into the PIRLS texts was necessary to determine whether the texts are indeed equivalent across the languages of English, Afrikaans and isiZulu. In order to do this, workshops were held during which the texts were discussed with panellists reviewing the texts and completing open-end questionnaires to provide further comments.

4.3.3 Phase One: Achievement Data from PIRLS Literacy 2016 and PIRLS 2016

A summary of the research process used to address the first and second sub-questions by way of the first phase or quantitative phase of this study is deliberated in Section 4.3.3.1. In addition, the sampling of the study is discussed (4.3.3.2) and the exposition of the analytical strategies used during Phase One of the study can be found in Section 4.3.3.3.

4.3.3.1 Secondary Analysis of PIRLS Literacy 2016 and PIRLS 2016 Data

The goal of the first quantitative phase was to identify items from the PIRLS Literacy 2016 and PIRLS 2016 released texts that display possible non-equivalence (bias) among the English, Afrikaans and isiZulu language groups. As such, this study takes the form of a secondary analysis. Secondary analysis is a re-analysis of existing data, which was collected for a different primary purpose, in order to address a new research question (Cheng & Phillips, 2014; Clark & Maynard, 2012; Johnston, 2014; Payne & Payne, 2004). Secondary analysis is versatile as it may be used in research to understand the past, change over time (i.e. trend analysis) and extend previous studies (Kiecolt & Nathan, 1985). It is also advantageous as it less costly than collecting data. Beyond economic reasons for using secondary analysis, most of the studies do not fully utilise all the data available, and so the secondary researcher delves further into existing data to extend it (Clark & Maynard, 2012).

However, before doing so it was necessary to first determine the South African learners' overall achievement on both PIRLS assessments. The PIRLS Literacy 2016 and PIRLS 2016 data were collected by using test booklets with two texts and accompanying items, as described in Chapter Two. The learners' mean score by language group were calculated by using statistical software, IDB-Analyzer that allows the complex weighting as well as the plausible values for the PIRLS assessments. These results were used to partially assist in addressing the study's first sub-question. Subsequently, an in-depth quantitative analysis, to assist in determining whether potential non-equivalence was present in the PIRLS assessments among the selected three language groups, was conducted. The in-depth analysis consisted of Rasch measurement

theory where item statistics, summary statistics, differential item functioning and distractor analysis were conducted.

4.3.3.2 Phase One Sampling

Both the Grades 4 and 5 PIRLS data were used in the first phase of the study. As such, the sampling approach for the original studies are outlined in this section. South Africa in conjunction with 50 other countries participated in PIRLS 2016 with 11 additional benchmark participants. The Grade 4 learners took part in PIRLS Literacy 2016 whereas the Grade 5 learners participated in PIRLS as a benchmark participant. PIRLS provides valid and reliable measurement of trends of learner achievement across participating countries (LaRoche, Joncas & Foy, 2017). For each participating country a representative random sample of schools and classes were tested. The benchmark participants did not have representative data as it does not include the entire population of the participating country, only specific languages or areas. In terms of South Africa, a nationally representative sample for Grade 4 was selected, meaning that this sample was representative of the 11 official languages in the nine provinces. However, at the Grade 5 level, South Africa only participated as a benchmark participant with representative samples of only three language groups, namely English, Afrikaans and isiZulu; however, these were representative of the three languages in the nine provinces (Combrinck et al., 2017).

Both the PIRLS Literacy 2016 and PIRLS 2016 assessments made use of a two-stage random sample design where the schools were selected as the first stage and then one or more intact classes were selected from these sampled schools as the second stage (LaRoche et al., 2017). Participating countries had to define their national target population, thereafter these countries had to apply the PIRLS sampling methods in order to achieve a nationally representative sample. The target population is defined by PIRLS as the amount of schooling learners have received which is further broken down as four years of formal schooling, by using the International Standard of Classification of Education (ISCED) (UNESCO, 2012). Furthermore, the learners who participate in PIRLS should not fall below the 9.5 years minimum average (LaRoche et al., 2017).

In South Africa, the target population for PIRLS Literacy 2016 was Grade 4 learners and that of PIRLS 2016 was Grade 5 learners. The first stage was based on the Language of Learning and Teaching (LoLT) in the Foundation Phase. Consequently, more learners were selected in

certain areas due to the place where the language is located (Combrinck et al., 2017). As part of the first stage, a random sample of schools was selected, representative of language and province, in proportion to size. As mentioned earlier, the second stage involved randomly sampling of classrooms from the selected schools. Note that the PIRLS Literacy 2016 sample was representative of both language and province but not together (Combrinck et al., 2017).

In terms of the PIRLS 2016 South African sample, the learners were approximately a year older than the above defined target population in order to do benchmarking. The population for PIRLS 2016 in South Africa was Grade 5 learners in schools were the LoLT was either English, Afrikaans or isiZulu in the Foundation Phase (Grade 1 to 3). In South Africa, the LoLT changes for the majority of learners at Grade 4, which means that the schools for PIRLS 2016 were selected, based on the LoLT in the Foundation Phase. As with PIRLS Literacy 2016, the first stage included schools which were randomly selected in proportion to size. The second stage included randomly sampled classrooms within schools. Note that because the Grade 5 sample does not include all 11 official languages, it is not a nationally representative sample, but because both PIRLS Literacy 2016 and PIRLS 2016 had samples for English, Afrikaans and isiZulu subgroups, the same schools were selected where possible (Combrinck et al., 2017).

The South African PIRLS Literacy 2016 sample comprised 293 out of 304 schools, with a total of 324 intact classrooms, resulting in a 96% school participation rate. The nationally representative sample size for PIRLS Literacy 2016 was 12 810 Grade 4 learners. As PIRLS Literacy 2016 was a voluntary study, some schools refused to participate and were replaced; however, 11 schools were not able to participate due to not meeting the target grade, schools that had closed down or in some cases, the schools were unreachable (Combrinck et al., 2017). For the South African PIRLS 2016 sample, 125 schools with 147 intact classes were sampled, resulting in a 94% school participation rate. The sample size was 5 282 Grade 5 learners; however, the sample was representative only of English, Afrikaans and isiZulu.

For the purposes of this study, both the Grades 4 and 5 PIRLS data were selected. However, only three languages were examined namely English, Afrikaans and isiZulu as explained earlier. As this study specifically looks at these three languages, it is pivotal to show the percentage of each language group tested in both PIRLS Literacy and PIRLS. Table 4.1 depicts the percentage of Grade 4 and 5 learners by PIRLS Literacy and PIRLS test language.

Table 4.1: PIRLS Literacy 2016 and PIRLS 2016 Test Languages

Test Language	PIRLS Literacy % of Population	PIRLS % of Population
Afrikaans	9.2	17
English	23.0	46
isiZulu	21.8	37
isiNdebele	0.3	Not applicable*
isiXhosa	15.9	Not applicable*
Sepedi	9.3	Not applicable*
Sesotho	5.2	Not applicable*
Setswana	7.1	Not applicable*
siSwati	2.3	Not applicable*
Tshivenda	2.2	Not applicable*
Xitsonga	3.8%	Not applicable*

*These languages were not tested as part of PIRLS 2016.

Only released texts from both PIRLS Literacy 2016 and PIRLS 2016 were used, as the other texts will be used in future cycles of PIRLS. The PIRLS assessment instruments were discussed in Chapter Two of this study, including the texts used.

4.3.3.3 Phase One Secondary Data Analysis

The first phase of the research made use of secondary data analysis (*cf.* sub-section 4.3.3.1). For this study, the existing PIRLS Literacy 2016 and PIRLS 2016 data were used as the data sources for the secondary analysis. By using ILSAs such as PIRLS, the data that is provided are already categorised and tabulated, which enables the secondary researcher to explore the data via descriptive and inferential statistics. It is a useful technique when elaboration is required on a specific research topic (Kiecolt & Nathan, 1985). Secondary analysis has been gaining traction in social sciences (*cf.* Payne & Williams, 2011) with some limitations. For example, the data source availability and the quality of the data could be considered as limitations as the researcher has no control over these aspects (Roux, 2014). Secondary analysis

of the PIRLS Literacy 2016 and PIRLS 2016 datasets was considered meaningful for this study as I had access to these databases, having worked on the PIRLS South African studies.

- Descriptive Statistics

The original PIRLS assessments recorded the learners' mean reading achievement data by text title, item number, processes of comprehension as well as purpose for reading. Each item then became a unique variable within the dataset. The international (Mullis et al., 2017) and national (Howie et al., 2017) reports presented learner achievement using descriptive summaries (Pietersen & Maree, 2016) between the overall mean score and other contextual variables derived from the four questionnaires, namely the school, teacher, parent or home and learner questionnaires. For this study, descriptive reporting also occurred where the mean score was calculated for each of the three selected languages, including the overall mean score for each grade. Percentages were also calculated to give a broader picture of what the data holds. Moreover, descriptive statics were used to identify any differences in reading literacy achievement between English, Afrikaans and isiZulu language sub-group responses. The descriptive statistics was conducted by using the International Database Analyzer (IDB Analyzer) software. IDB Analyzer is a plug-in program for the Statistical Package for the Social Sciences (SPSS), developed by the IEA in order to combine and analyse data from ILSA datasets such as PIRLS (Foy, 2017). Raw mean scores as well as additional percentages were calculated by using SPSS. The learner overall performance by language was determined and is presented in Chapters 5 and 6.

- Rasch Measurement

After the initial descriptive quantitative analyses were completed, more sophisticated analysis techniques were used in order to assist in addressing the first two sub-research questions. Rasch theory is measurement driven and the data must fit the model, while Item Response Theory (IRT) uses statistical models, rather than measurement models, where the model must fit the data (Engelhard, 2013). As such, Rasch theory varies from IRT as each perceives measurement differently (Engelhard, 2013). Rasch measurement theory was utilised in this study to analyse the PIRLS Literacy 2016 and PIRLS 2016 secondary data. The Rasch IRT model is a one-parameter model that measures learners' probability or chance to answer an item correctly (Smith & Smith, 2004). Here *probability* refers to the chance that a learner has to correctly respond to a test item taking into consideration that the answer depends on item bias. *Measures*,

according to Combrinck (2018), have to be "equally ordered, additive and meaningful approximations of the construct under consideration" (p. 25). In terms of this study, the construct is the learners' reading literacy achievement on PIRLS. When learners write a test, it is generally understood that learners who are able to correctly answer the more difficult items, should also answer the easier items correctly. In the same way, learners who struggled with the easier items are unlikely to correctly answer the more difficult items. This concept is recognised as the Guttman pattern (Abdi, 2010; Linacre, 2013) where items and persons are ideally ordered. The Guttman pattern, a foundation for Rasch models as it is based on the probability of success on a test item, is seen as a warning that there is extreme overfit (Combrinck, 2018; Wright, 1977).

As such, the Rasch Model uses a scale where the learners' ability and the item difficulty scores are compared. Because testing is an important aspect in many different fields, Rasch measurement theory could be used to examine the invariance of a test by making it more attainable as it takes into consideration the ordering of the items and persons on a single scale. By using the Rasch Model, it is possible to determine whether the test instrument is valid and reliable by providing evidence that if a different group answered the same test instrument, it would produce the same ordering of items (Andrich 2011; Linacre 2016). Furthermore, a crucial advantage of using Rasch IRT is when heterogeneous groups, such as linguistic profile, race, socio-economic status amongst others are tested and in turn, invariance could be prevalent which may render the test to become unfair (ETS, 2014). This means that the construct (for example, learner achievement) that is being measured by a test instrument is invariant for one group in comparison to the other and requires additional investigation to ensure that the construct measures the same way across the different groups (Combrinck, 2018). Ideally, the item order should remain the same for all groups but the test itself may be more or less difficult for each group (Randall & Engelhard, 2010).

In order to determine whether items functioned differently across different groups, Differential Item Functioning (DIF) may be utilised. DIF is a technique used to analyse survey and more importantly, test data, which is conducted via Rasch measurement (Boone, Staver & Yale, 2014). The following is an example question for conducting DIF: Are there items on the test that are biased in some manner towards gender? DIF may be used when a researcher is interested in "evaluating if the manner in which items define a measurement scale does so in the same way for different groups of respondents" (Boone et al., 2014, p. 274). In other words,

determining whether the 'yardstick' used when measuring different groups operates in the same way. DIF is often used when a test instrument is translated into more than one language (Gierl & Khaliq, 2001).

DIF is founded on the Rasch between-fit-statistic, using a characteristic of the persons, who are in the analysis, in order to separate them into groups (van Staden, 2018). It is a useful technique as it is able to detect differences in the item-level performance for different sub-groups as person ability and item difficulty is included on the same scale in Rasch measurement. As such, the aim of employing Rasch analysis in this study was to establish whether the items selected from PIRLS Literacy 2016 and PIRLS 2016 functioned differently for learners from different language backgrounds, in other words whether non-equivalence, also known as item bias was present. Put differently, the South African Grade 4 and 5 learners' probability of answering the item correctly was dependent on the English, Afrikaans and isiZulu group differences. The differences between the languages may be associated with different probabilities of a person answering an item correctly (Gierl & Khaliq, 2001).

For the purposes of this study, RUMM2030 software (Andrich, Sheridan & Luo, 2012) was used to analyse the PIRLS Literacy 2016 and PIRLS 2016 data. During the first phase of this study, after the initial descriptive analyses, I conducted summary statistics to determine whether the data fitted the Rasch model. Thereafter, I carried out reliability analysis for each language group and text. In terms of the reliability analysis, Cronbach Alpha (α) was used where the .7 was regarded as sufficient (Field, 2009). Cronbach Alpha is used to measure internal consistency; differently put, it is how closely the items are grouped together. It is considered as a measure of test reliability on the raw score of a sample and can only be reported on complete datasets, missing data is not allowed (Linacre, n.d; Linacre, 2012). Most of the texts (by language) had acceptable α coefficients. However, for the isiZulu *MacyTurtle* analysis, two items[21] were removed as they did not fit the Rasch model. After the removal of Turtle Item 7 and Turtle Item 11, the isiZulu group had an α coefficient of .6.

Next, individual item-fit statistics were conducted for the *MacyTurtle*, *PearlRhino* and *Flowers* analyses for the combined languages and then separately for each language. This was done in

[21] These items had zero persons for some categories and would appear to have been very difficult for the isiZulu learners.

order to partially assist in accepting or rejecting the null hypothesis. This study had the following hypotheses[22]:

- H_0: μEnglish = μAfrikaans = μisiZulu
- H_a: μEnglish ≠ μAfrikaans ≠ μisiZulu

Thereafter DIF was conducted to determine which items displayed differential functioning across the English, Afrikaans and isiZulu language groups. The output given by RUMM2030 provided an analysis summary of the data including an Analysis of Variance (ANOVA) test. It displayed a person factor ANOVA of each set of items. Table 4.2 shows the number of items per RUMM2030 analysis file.

Table 4.2: Number of Item per RUMM2030 Analysis File

Text Title	RUMM2030 Analysis Name	PIRLS or PIRLS Literacy	Brief Summary [a]	Total Number of Items
The Pearl	PearlRhino	PIRLS Literacy	This story, about a young pearl merchant, illustrates the power of home, friendship and generosity above greed.	15
African Rhinos and the Oxpecker Birds	PearlRhino	PIRLS Literacy	This passage presents information about African rhinos and oxpecker birds and describes how the two animals depend on one another for food and survival.	17
Flowers on the Roof	Flowers	PIRLS Literacy & PIRLS	This contemporary story portrays friendship between the generations.	13
Macy and the Red Hen	MacyTurtle	PIRLS	This contemporary story portrays a complex character who meets a challenge when caring for a red hen.	16

[22] The hypotheses are the same for the PIRLS Literacy 2016 and PIRLS 2016 analyses.

Text Title	RUMM2030 Analysis Name	PIRLS or PIRLS Literacy	Brief Summary [a]	Total Number of Items
The Green Sea Turtle's Journey of a Lifetime	MacyTurtle	PIRLS	This passage describes the life cycle of a female green sea turtle from the time she hatches from an egg to the time she lays her own eggs.	16

[a] From *PIRLS 2016 International Results in Reading* (p. 1.21), by I. V. S. Mullis et al. (Eds.), 2017, TIMSS & PIRLS International Study Center. Copyright 2017 by TIMSS & PIRLS International Study Center.

The *PearlRhino* included a total of 32 items. For the *MacyTurtle* analysis, a total of 32 items was used. For both *Flowers* analyses, 13 items were selected. *Flowers on the Roof* was a linking text between PIRLS Literacy 2016 and PIRLS 2016. The ANOVA results, presented in Chapters Five and Six of this study, compare the mean scores of independent groups (English, Afrikaans and isiZulu language sub-groups) and also assist in testing the null hypothesis where the group means are all equal (Field, 2009; Tabachnick & Fidell, 2007). The ANOVA output provides the *F*-ratio and *p*-value, where the *F*-ratio is a measure where two or more quantities are expected to be equal under the null hypothesis and the *p*-value is the statistical model used to give evidence whether or not the null hypothesis should be accepted. As such, the ANOVA gives evidence whether the mean score differences (between groups) are statistically significant. The Bonferroni correction was selected for the ANOVAs conducted with RUMM2030 (Andrich & Marais, 2019). This type of correction is used as some scholars have concerns regarding test of fit, and that many could be significant due to chance. The Bonferroni correction is commonly used and is straightforward: "the chosen probability value of significance is simply divided by the number of tests of fit" (Andrich & Marais, 2019, p. 196) as it reduces the risk of type I errors. In addition, person-item maps were also created to assist in gaining a holistic view of the items in conjunction with the learners. The person-item map provides visual information on which items were too easy or too difficult as it plots the learners and items on the same scale.

Lastly, DIF analyses were conducted (*cf.* Chapters Five and Six) and were graphically presented by item characteristic curve (ICC) graphs. ICCs were generated to graphically show the relationship between person ability and probability of answering the item correctly. Based

on the evidence gathered from the Rasch analyses, the hypothesis was that there are differences in the English, Afrikaans and isiZulu learner item responses. The different probabilities of the three language groups may have influenced the manner in which the learners correctly answered the item. In this study, the learners from the three different language groups had varied performances across the selected items, indicating non-equivalence.

4.3.4 Phase Two: Exploration of PIRLS Literacy 2016 and PIRLS 2016 Texts

The second phase of the research addresses the third, fourth and fifth sub-questions. This qualitative phase of the study makes use of case studies of selected PIRLS Literacy 2016 and PIRLS 2016 texts. This sub-section discusses the second phase or qualitative phase of the mixed methods study. The value of using qualitative research methods is discussed (4.3.4.1) followed by a description of the PIRLS texts selected for this study (4.3.4.2). Subsequently, the data analysis (4.3.4.3) strategies for the second phase of this study is deliberated.

4.3.4.1 Value of Qualitative Research

Because quantitative data entails numerical data and testing hypotheses, it may not be sufficient to address complex research problems. As such, qualitative research can be used to contextualise and interpret the quantitative data (Östlund et al., 2011). Qualitative research, according to Creswell et al. (2006), is invaluable to mixed methods research and can add to the mixed methods explanations (that is, help explain results). Within this study, qualitative methods were employed to expand and elaborate on the quantitative results obtained in the first phase.

The second, qualitative phase of this research addresses the PIRLS Literacy 2016 and PIRLS 2016 limited released texts and includes a text analysis of these texts. In doing so, the goal was to use the results from the second phase to complement and extend the findings from the first phase of this research. The ability to delve deeper is what makes qualitative methods so appealing for this study. Mason (2006) stated that “a distinctive strength of qualitative research lies in its intimate and habitual concern with context, with the ‘particular’ and with understanding the situatedness of social experience, processes and change” (p. 17). For this research, the methodological vehicle involves content analysis and questionnaires.

4.3.4.2 Nature of the Data

As already stated, for this research, five PIRLS released texts were analysed. The English versions of these texts were released by the IEA and the translated versions were released after permission was sought from the South African PIRLS National Research Coordinator (NRC). The texts are either narrative (literary) or informational in nature. *The Pearl*, *Flowers on the Roof* and *Macy and the Red Hen* were literary type texts while *African Rhinos and the Oxpecker Birds* and *The Green Sea Turtle's Journey of a Lifetime* were informational texts. Table 4.3 provides a breakdown of word count as well as the Flesch-Kincaid grade level (see description below) for each of the released texts in English.

Table 4.3: Word Count and Flesch-Kincaid Grade Level of PIRLS Released Texts

Study	Text Title	Word Count	Flesch-Kincaid Grade Level [a]
PIRLS Literacy 2016	The Pearl	536	2.9
PIRLS Literacy 2016	African Rhinos and the Oxpecker Birds	449	4.7
PIRLS & PIRLS Literacy	Flowers on the Roof	811	2.8
PIRLS 2016	Macy and the Red Hen	913	4.4
PIRLS 2016	The Green Sea Turtle's Journey of a Lifetime	943	4.0

[a] The Flesch-Kincaid grade level readability formula is used by the IEA as it uses average syllables for each word as well as the average sentence length to produces a number output representing the United States grades in which learners can read the specific text.

4.3.4.3 Phase Two Data Analysis

This study made use of summative content analysis to determine whether the content of the PIRLS Literacy 2016 and PIRLS 2016 released texts were appropriate for South African Grade 4 and 5 learners, taking into consideration three of the four components of the study's conceptual framework, namely *linguistic*, *cultural* and *functional* equivalence. To recap, *linguistic equivalence* refers to the linguistic meaning of the words used in the test and that the words are the same for the different versions whereas *cultural equivalence* involves how different cultural groups understand the underlying construct of the text (Peña, 2007).

Functional equivalence entails whether the test measures the same construct across different versions (Peña, 2007).

Content analysis is a systemic process of making sense of qualitative data by recognising patterns or themes that may offer meaningful answers to a research question (Vaismoradi, Jones, Turunen & Snelgrove, 2016). This type of analysis "starts with identifying and quantifying certain words or content in text with the purpose of understanding the contextual use of the words or content" (Hsieh & Shannon, 2005, p. 1283). In order to answer the sub-questions for the qualitative phase or Phase Two, of this study, content analysis was used to determine whether the released PIRLS texts were in fact linguistically, culturally and functionally equivalent across the different languages. Content analysis requires reading and judgement and it involves "summarising and reporting on written data" (Cohen et al., 2007, p. 475).

For this research, the second qualitative phase involves an in-depth look at the translation of the given texts from English into Afrikaans and isiZulu. The translation process of the original study was conducted under the supervision of the PIRLS 2016 NRC and any changes made were recorded on the National Adaptation Forms (NAF). The translators, selected for the mammoth task in South Africa, were vetted by the NRC and had to meet all the requirements set out by the IEA (*cf*. Chapter Two Section 2.6.2.2). For the purposes of this study, a panel of experts was asked to assist in scrutinising the Afrikaans and isiZulu versions of the released texts in order to determine whether these texts and their accompanying items tested what they purportedly set out to test and whether these were equivalent to the English version (source language). The panellists were selected based on whether they were able to speak and write in the home language (Afrikaans/isiZulu); in addition, the panellists were also selected based on their experience with assessments. The panellists consisted of two Afrikaans and three isiZulu persons. The researcher was part of the Afrikaans panellists while the second panellist is a lecturer and researcher at the University of Pretoria, specialising in psychometrics. The first isiZulu panellist is a school teacher who teachers isiZulu to primary school learners and who also assists with an honours course at the University of Pretoria. The second expert is an isiZulu lecturer at the University of Pretoria who has taught isiZulu for many years. The final isiZulu panellist is a professor of isiZulu linguistics at the University of Johannesburg who has delivered many post-graduate students researching aspects of isiZulu language. A set of criteria

was developed for the scrutinising of the texts and included word count, low and high frequency words, translation errors and equivalence to the source text.

Content analysis enables scrutiny of the PIRLS released texts in order to determine whether there are discrepancies in the translation as well as the extent to which these discrepancies hinder or enhance learner performance. Moreover, content analysis is used to validate (Krippendorff, 2004) the fairness of the selected PIRLS texts used in the South African context. The texts were nationally and internationally verified, signifying that they were translated from the source language, English, to the other language and then back-translated into English again. In this way it is possible to identify the discrepancies in the translations and make corrections. By using descriptive statistics, Rasch measurement and content analysis, it is possible to explore problem areas in the texts and items to determine whether the translations were compromised.

4.4 METHODOLOGICAL CONSIDERATIONS FOR THE STUDY

This section provides the methodological considerations for the research. My approach to the methodological norms is discussed in Section 4.4.1. The methodological norms for the first, quantitative phase are discussed in Section 4.4.2 followed by a similar discussion for the second, qualitative phase (4.4.3).

4.4.1 Approach to Methodological Norms for this Study

Over the years, mixed methods research has become a popular method for conducting research as it could enable the researcher to make use of both quantitative and qualitative methods in order to solve complex research problems in many different fields (*cf.* Archibald, et al., 2015; Bergman, 2008; Ivankova et al., 2006; Onwuegbuzie et al., 2009). In other words, mixed methods design is a more robust method of conducting research. According to Zimmerman (2010), there are many terms and definitions relating to methodological norms, such as validity and reliability, amongst others.

For this research, I have discussed the methodological norms for the first, quantitative phase separately from the second, qualitative phase. In order to ensure quality research, the methodological norms for both phases were carefully considered. Key criteria for ensuring quality research include inference quality, which is a combination of design quality and interpretation rigour (O'Cathain, 2010). The former involves design rigour (methods that are

implemented are true to the original design) and design transparency (a description of the design type from a known typology) (Creswell & Plano Clark, 2011) while the latter includes interpretive transparency (it is certain which findings arose from which method) and interpretive consistency (the inferences match the findings) (Teddlie & Tashakkori, 2009) with the conclusions being based on the findings of the research. In order to meet the criteria, my goal was to clearly describe the research design, the data sources used, how the quantitative and qualitative methods can complement each other and how the data were analysed.

4.4.2 Methodological Norms for the First Quantitative Phase

Since this study made use of the PIRLS Literacy 2016 and PIRLS 2016 learner achievement data and the five released texts, it is important to discuss the validity and reliability of these ILSAs. The PIRLS studies were conducted under the auspices of the IEA and as such conform to the stringent technical standards, guidelines and monitoring set out by the IEA. The validity and reliability of PIRLS Literacy 2016 and PIRLS 2016 is based upon the measurement of learners' reading literacy skills and competencies, obtaining contextual data (questionnaires) to assist in understanding the reading literacy achievement scores, data collection strategies for simultaneous collection in each of the hemispheres, using internationally accepted testing instruments (test booklets) and the management of the entire research process. The standards are divided into instrument development, sampling, data collection procedures and reporting (Martin et al., 2017).

In terms of the data collection, Survey Operations Procedures (SOP) were developed to enhance the operations of PIRLS Literacy 2016 and PIRLS 2016 and then ensure that only quality data is collected from each of the participating countries (Martin et al., 2017). In order to do so, quality control measures were put into place guided by the SOP. This means that there was quality control both internationally and nationally. Nationally, the South African NRC ensured that fieldwork monitors were sent to schools to check that the protocols were being adhered to. Moreover, an International Quality Control Monitor (IQCM) was appointed who had randomly selected which schools to visit unbeknown to the NRC (Martin et al., 2017; Howie et al., 2017). The IQCM served as an external quality control measure that recorded the data collection process, made notes where any rules were not adhered to and reported directly to the IEA. The Centre for Evaluation and Assessment (CEA), functioning as the national research centre for PIRLS, ensured high quality control, thus providing the validity and

reliability for this study (see Chapter Two of this study for a discussion of the development of the instruments, translations, data collection, monitoring and scoring).

4.4.3 Methodological Norms for the Second Qualitative Phase

Lincoln and Guba (1985, 2000) explain that in order to gain trustworthiness of qualitative research four criteria should be met. The criteria include credibility, dependability, transferability and confirmability (Tolley, 2016). In the following sub-sections, I discuss each of the four criteria in terms of this study.

4.4.3.1 Credibility

Credibility looks at the confidence in the truth of the qualitative findings (Tolley, 2016) as it establishes whether the research findings are plausible (Korstjens & Moser, 2018). Morrow (2005) explains that credibility may be "enhanced by a thorough description of source data and a fit between the data and the emerging analysis" (p. 252). It involves thick and rich descriptions of the participants and contexts. However, this aspect of the study did not focus on participants, but was based rather on text data sources (PIRLS Literacy 2016 and PIRLS 2016) used in this study. These data sources were described in detail (*cf*. Chapter Two) as well as through the statistical analysis (*cf*. Chapters Five and Six). The statistical analysis was linked to the qualitative phase of this research, specifically to the content analysis. Accordingly, the findings and interpretations from this study are linked and confirmable. In her work, Morrow (2005) explained that confirmability is a crucial part of credibility as "it is based on the acknowledgment that research is never objective" (p. 252), meaning that the integrity of the findings is placed on the data and on the researcher's ability to bring the data together.

4.4.3.2 Dependability and Confirmability

Dependability, as with reliability in quantitative research, has to do with the manner in which the study is conducted – it should be consistent in terms of researchers, analysis techniques used and the time the study takes (Morrow, 2005). It is the extent to which the findings of the research can be replicated (Tolley, 2016) and depends on the evaluation of the findings and interpretations to ensure that it is supported by the data. Dependability can be ensured by checking that the research is well documented and has a logical flow, in other words, that there is an 'audit trail' (Korstjens & Moser, 2018). The audit trail means that the confirmability of the study's interpretations can be attained. As such, the data sources, RUMM2030 databases,

research outputs, expert reviews and texts[23] are presentable so that scholars, students and other interested parties are able to access it (Korstjens & Moser, 2018). The original data sources are currently available on the PIRLS website.

4.4.3.3 Transferability

Transferability, also known as external validity, concerns the generalisability of the research (Morrow, 2005). Korstjens and Moser (2018) explain it as "the degree to which the results of qualitative research can be transferred to other contexts or settings with other respondents" (p. 121). Transferability is done by thick description of the research. Even though qualitative research usually has small samples (as they include human participants), some have argued that even though qualitative research does not work with large samples, the findings should be generalised as they contain invaluable information (Mason, 2006). As this research made use of secondary data, no human participants were part of this study. Therefore, transferability refers to the thick description of the qualitative results, the South African educational landscape in which the original study's participants are situated.

4.5 ETHICAL CONSIDERATIONS FOR THE STUDY

The Minister of Education granted the Centre for Evaluation and Assessment (CEA) permission to conduct the PIRLS 2016 study. Thereafter, ethical clearance was received by the University of Pretoria and testing began in 2015. Ethical clearance to undertake this study was obtained from the Ethics Committee of the Faculty of Education at the University of Pretoria (SM 17/07/04). The IEA provided permission to use the PIRLS Literacy 2016 and PIRLS 2016 data. Furthermore, the PIRLS national research centre, the Centre for Evaluation and Assessment (CEA), also gave the researcher permission to use the PIRLS Literacy 2016 and PIRLS 2016 data and released texts for secondary analysis (see Appendix I).

In the original study, anonymity and confidentiality of the participants of the primary study was ensured by using unique identification numbers. This study relied on existing test data as well as expert opinion on the five PIRLS texts. No human participants were assessed in this study as the data had already been collected and was available in the public domain upon request. For this study, the following steps were taken:

[23] To obtain the translated texts, an additional request must be sent to the South African PIRLS NRC.

- Permission from the IEA to conduct secondary research on the PIRLS Literacy 2016 and PIRLS 2016 data (including the released texts);
- Permission from the National Research Co-ordinator (NRC) to conduct secondary research on the PIRLS Literacy 2016 and PIRLS 2016 data (including the released texts in Afrikaans and isiZulu); and
- Ethical clearance from the University of Pretoria's Ethics Committee of the Faculty of Education.

It should be noted that all assessment instruments, questionnaires, manuals and guidelines as well as the PIRLS Literacy and PIRLS data are securely kept in storage at the CEA. The researcher obtained permission from the IEA and the NRC to use and analyse the PIRLS Literacy and PIRLS data and also obtained the released passages which are used during the course of the study.

4.6 CHAPTER SUMMARY

Since the design considerations and methodological endeavours for this research were discussed in this chapter, it is useful to present a summary that depicts how the different data sources, research questions and chapters link with each other. Table 4.4 presents the data sources, main research question, sub-questions and in which chapter the findings of each of the questions are presented.

Table 4.4: Data Sources aligned to Research Questions

Main Research Question	Sub-research Questions	Data Source	Chapter in Thesis
To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?	How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?	PIRLS Literacy 2016 SA Grade 4 learner achievement data	Quantitative Chapter 5
	How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016?	PIRLS 2016 SA Grade 5 learner results	Quantitative Chapter 6
	To what extent do the PIRLS Literacy 2016 and PIRLS 2016	PIRLS Literacy 2016 and PIRLS 2016 texts;	Qualitative Chapter 7, 8

Main Research Question	Sub-research Questions	Data Source	Chapter in Thesis
	assessment instruments have face validity?	South African curriculum	
	How does the textual coherence, as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?	PIRLS literary (narrative) texts; panel of expert opinions; questionnaires and workshop notes	Qualitative Chapter 7
	How does the textual coherence, as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?	PIRLS informational (expository) texts; panel of expert opinions; questionnaires and workshop notes	Qualitative Chapter 8

The quantitative findings for sub-question 1 is presented in the next chapter, Chapter Five.

-⌘-

CHAPTER 5

Phase One: PIRLS Literacy 2016 Grade 4 Quantitative Data Presentation and Results

5.1 ORIENTATION

This chapter presents the findings for the first phase of this study, the quantitative phase, a secondary analysis, using descriptive statistics as well as Rasch measurement theory. For the purposes of this chapter, specific attention is paid to the South African PIRLS Literacy 2016 results in Section 5.2, followed by the Rasch findings and discussion thereof (5.3). Section 5.4 gives a summary of those items which displayed differential item functioning (DIF) across the three languages. These items originated from the three released texts, namely *The Pearl*, *African Rhino and the Oxpecker Bird* and *Flowers on the Roof*. This chapter aims to assist in addressing the first sub-question of the study: *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?*

5.2 DESCRIPTIVE RESULTS OF SUB-QUESTION 1

This section is divided into two sub-sections, each providing evidence for addressing sub-question 1. The first sub-section (5.2.1) provides the South African Grade 4 learners' PIRLS Literacy 2016 results. Thereafter, descriptive statistics is provided in Section 5.2.2 for addressing the sub-research question 1.

5.2.1 PIRLS Literacy 2016 Results

The South African Grade 4 learners participated in the PIRLS Literacy 2016 based on prior evidence[24] and decisions made by the-then National Research Coordinator (NRC) (Howie et al., 2017a). For the South African PIRLS Literacy 2016 assessment, a nationally representative sample was drawn across languages and provinces which meant that the findings of this study are generalizable to the Grade 4 learner population. However, for the purposes of this study, a decision was made to focus on only three of the 11 official languages, namely English, Afrikaans and isiZulu. The reason for this decision is based on the languages used at Grade 5 level (also English, Afrikaans and isiZulu); furthermore, looking at more languages would have required extended time and resources. The total sample for the South African PIRLS Literacy

[24] *cf*. PIRLS 2006 results (Howie et al., 2008; Howie et al., 2017a).

2016 assessment was 12 810 Grade 4 learners in 293 schools within the nine provinces (Howie et al., 2017a). As this study only focuses on the three-abovementioned languages, the sub-sample is reduced to 5 049 learners. This sub-sample was also further reduced as this study only focused on the released texts. As the same learners completed *The Pearl* and *African Rhinos and the Oxpecker Birds*, the sample for this study was 836; while *Flowers on the Roof* included a sample of 820.

The overall South African Grade 4 learner results of PIRLS Literacy 2016 are depicted in Figure 5.1. Note that the scale ranges from zero to 1000 with an average of 500 and a standard deviation of 100. As mentioned in Chapter Two, the scale was calculated by using Item Response Theory (IRT) scaling. This figure includes the South African Grade 4 learner results per language as well as the overall average. These results are shown relative to that of the international mean. The Grade 4 learners achieved a mean score of 320 (SE=4.4) which was the lowest among the participating countries (*cf.* Appendix A).

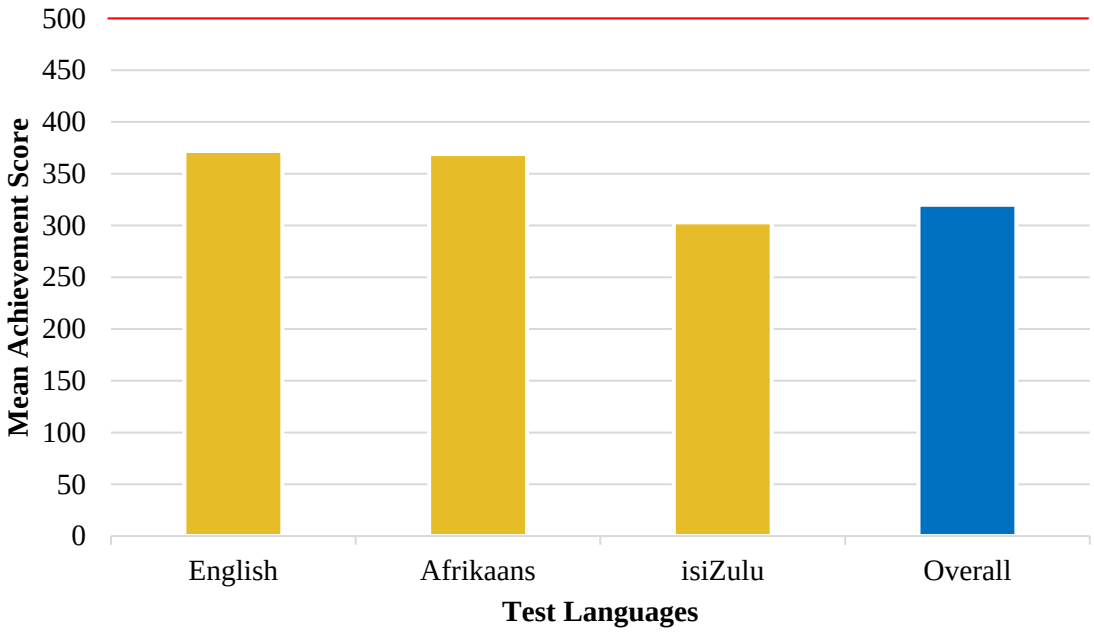

Figure 5.1: South African Grade 4 Learner Achievement across Selected Languages

From the above figure, none of the selected languages nor the overall mean score of the South African Grade 4 learners reached the international mean of 500. Learners who completed the test in English achieved the highest score (372 points, SE=14.4) followed by Afrikaans (369 points, SE=13.4) and isiZulu (303 points, SE=4.3). A 69-point difference exists between the learners who wrote the test in English and those who wrote in isiZulu. Similarly, a 66-point

difference exists between the Afrikaans and isiZulu language groups. This finding provides evidence that there is substantial variation between the languages and both these large differences equate to about two years of education (*cf*. Howie et al., 2017b). These differences were tested to determine whether they were statistically different. Table 5.1 presents the significance in differences among the three language groups.

Table 5.1: Language Comparisons for South African Grade 4 Learner Achievement

Languages	Mean	SE	English	Afrikaans	isiZulu
English	372	14.4		●	▲
Afrikaans	369	13.4	●		▲
isiZulu	303	4.4	▼	▼	

▲ Significantly higher than ▼ Significantly lower than ● Not significantly different
Significance level < 0.05

Note. From *PIRLS Literacy 2016 Progress in International Reading Literacy Study 2016: South African Children's Reading Literacy Achievement* (p. 55), by S. Howie et al., 2017a, Centre for Evaluation and Assessment. Copyright 2017 by Centre for Evaluation and Assessment.

No statistical difference was found between those learners who wrote the PIRLS Literacy 2016 assessment in English and Afrikaans, while both these languages achieved mean scores significantly higher than the isiZulu language sub-group. Various factors could be at play in the achievement of learners who completed PIRLS Literacy 2016, such as parental involvement, teacher qualifications, learning and teaching support materials, reading books and learner attitudes towards reading, to name but a few. A possible reason for the stark difference between English, Afrikaans and isiZulu achievement might be the issues relating to quality of translation.

Did the language of the test align with the learners' home language? In general, the learners who completed PIRLS Literacy 2016 indicated that most (66%) of them always spoke the language of the test at home, while 19% of the learners stated that they sometimes spoke the language of the test at home and 6% said they never spoke the language of the test at home (Howie et al., 2017a). Moreover, the South African PIRLS Literacy 2016 data has a variable that combines the data from the learner and parent questionnaires regarding the language learners speak at home and comparing it to the language of the test. The result indicated that 71% of the learners spoke the same language as the test language and obtained a score of 315

points (SE=4.1). Twenty-nine per cent of the learners stated that they spoke a different language than the test language. Surprisingly, these learners obtained a mean score of 333 (SE=7.5 (Howie et al., 2017a) which was significantly different (t=-2.74) from those where the language at home and the test language were the same.

5.2.2 Descriptive Statistics to Address Sub-Question 1

Sub-question 1, as part of the quantitative phase of this study, asked the following: *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?* In order to address this question, items from the three released texts from the PIRLS Literacy 2016 assessment, namely *The Pearl, African Rhino and the Oxpecker Bird* and *Flowers on the Roof* were analysed. Table 5.2 depicts the total number of items per text as well as the maximum points for each of the texts.

Table 5.2: Total Items and Maximum Points per PIRLS Literacy 2016 Text

Text	Total Number of Items	Maximum Points
The Pearl	15	18
African Rhino and the Oxpecker Bird	17	19
Flowers on the Roof	13	16
Total	45	53

The Pearl and *African Rhino and the Oxpecker Bird* were part of the Reader booklet and *Flowers on the Roof* were part of booklets that were not released in the public domain. The *Flowers on the Roof* was developed as a trend passage in 2001 which was released during the last cycle of PIRLS (2016). A total of 45 items were released with a maximum of 53 points. Table 5.3 shows the overall raw mean score for each of the texts across the three languages.

Table 5.3: PIRLS Literacy 2016 Raw Mean Score per Text by Language

Text	Overall Mean Score	English Mean Score	Afrikaans Mean Score	isiZulu Mean Score
The Pearl (18)	8.19	9.39	8.69	6.41
African Rhino and the Oxpecker Bird (19)	9.00	10.10	8.82	7.84
Flowers on the Roof (16)	4.02	4.78	4.40	2.79

It would appear that South African Grade 4 learners struggled the most with the *Flowers on the Roof text* obtaining an overall mean score of 4.02, with the isiZulu (2.79) learners achieving the lowest mean score in comparison to the other two languages for this text. Learners who wrote the test in isiZulu (6.41) scored the lowest mark for *The Pearl* text. Based on the information contained in Table 5.3, the English sub-group performed the best on the *African Rhino and Oxpecker Bird* text.

Table 5.4[25] presents the percentage of learners who correctly answered the 15 individual items for *The Pearl*. This text is a literary type text initially set on a beach where a group of young children played around in the water. One of the children found a pearl and gave it to a boy who really liked it. The boy then became a wealthy pearl merchant but realised that having all the money did not make him happy, so he returned to his childhood home and found other things that were more important.

Table 5.4: The Pearl: Number and Percentage of Grade 4 Learners Who Correctly Answered Items in English, Afrikaans and IsiZulu

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
1	344	233	68	188	129	69	287	136	47*
2	343	205	60	187	109	58	281	79	28*
3	346	197	57	194	110	57	286	126	44*
4	342	189	55	196	106	54	285	131	46*
5	340	124	36*	193	74	38*	282	54	19*
6	339	203	60	193	99	51	286	140	49*
7	343	218	64	188	104	55	278	123	44*
8	343	176	51	189	99	52	276	126	46*
9	333	99	30*	191	67	35*	276	50	18*
10	338	196	58	181	102	56	270	142	53
11	325	64	20*	177	40	23*	270	37	14*

[25] Does not include missing data.

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
12	336	212	63	180	91	51	267	128	48*
13	324	196	60	181	91	50	265	102	38*
14	339	251	74	181	118	65	258	114	44*
15	336	65	19*	184	28	15*	265	12	5*

**Less than 50% of the responses were correct.*

From the table above, it is clear that the learners who completed the test in isiZulu found this text considerably more difficult as only one item, item 10, was answered correctly by more than 50% of learners. In contrast, the English and Afrikaans sub-groups struggled with only four of the 15 items, and with the same items, namely items 5, 9, 11 and 15, where less than 50% of the learners were able to answer the items correctly. This particular text may have demanded a higher cognitive level of reading for isiZulu learners.

Table 5.5 shows the percentage of learners who correctly answered items from an information text, *African Rhinos and the Oxpecker Bird*. This text is an informational text describing the symbiotic relationship between the African rhinoceros and the oxpecker birds relating to ticks. The text is divided into four sub-headings, each focusing either on the relationship between the two animals or the animals themselves. A total of 17 items were given to the learners to complete.

Table 5.5: African Rhinos and the Oxpecker Bird: Number and Percentage of Grade 4 Learners Who Correctly Answered Items in English, Afrikaans and IsiZulu

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
1	324	187	58	192	93	48*	280	97	35*
2	329	209	64	187	107	57	281	104	37*
3	324	116	36*	187	74	40*	278	101	36*
4	325	233	72	194	87	45*	281	83	30*
5	325	205	63	194	94	48*	283	72	25*

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
6	325	199	61	187	121	65	283	189	67
7	330	271	82	196	155	79	284	200	70
8	330	181	55	187	95	51	283	132	47*
9	325	106	33*	192	45	23*	281	59	21*
10	328	204	62	184	137	74	279	184	66
11	317	160	50	186	89	48*	269	120	45*
12	320	164	51	191	66	35*	279	126	45*
13	313	174	56	183	79	43*	264	163	62
14	310	218	70	184	112	61	262	137	52
15	317	152	48*	184	77	42*	269	104	39*
16	310	151	49*	189	56	30*	265	56	21*
17	307	49	16*	178	31	17*	261	13	5*

* Less than 50% of the responses were correct.

It would appear that the learners who completed the test in either Afrikaans or isiZulu found the test considerably more difficult than those who completed the test in English. For learners who completed the test in Afrikaans, only 6 of the 17 items (items 2, 6, 7, 8, 10 and 14) were correctly answered by more than 50% of the learners. For those who completed the test in isiZulu, only 5 of the 17 items were correctly answered by more than 50% of the learners (items 6, 7, 10, 13 and 14). In contrast, for learners who completed the test in English, 12 of the 17 items were answered correctly by 50% and more learners, while items 3, 9, 15, 16 and 17 were found to be more difficult. It seems that the English learners found this text and items to be less difficult.

Table 5.6 shows the percentage of learners who correctly answered the items for the *Flowers on the Roof* text. This text is a literary text about a young boy who had hoped to make a new friend, close to his own age, but became friends with a granny instead. When the granny moved to her new flat, she was unhappy at first but the boy helped her to make her flat into a home.

The boy was surprised about the grass she planted on the roof and the farm animals she had in her home.

Table 5.6: Flowers on the Roof: Number and Percentage of Grade 4 Learners Who Correctly Answered Items in English, Afrikaans and IsiZulu

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
1	334	151	45*	196	59	30*	270	61	23*
2	323	199	62	191	109	57	251	105	42*
3	323	161	50	197	86	44*	260	92	35*
4	333	176	53	198	82	41*	272	107	39*
5	332	76	23*	198	85	43*	272	103	38*
6	326	65	20*	196	41	21*	252	11	4*
7	327	22	7*	197	19	10*	266	4	2*
8	320	49	15*	196	34	17*	265	10	4*
9	318	72	23*	195	20	10*	255	7	3*
10	313	99	32*	194	81	42*	243	71	29*
11	321	108	34*	191	75	39*	248	52	21*
12	300	0	0*	181	0	0*	237	0	0*
13	304	138	45*	182	64	35*	235	82	35*

*Less than 50% of the responses were correct.

The items based on the *Flowers on the Roof* text appear to be extremely difficult for the isiZulu sub-group as none of the items were correctly answered by at least 50% of the learners, closely followed by the Afrikaans sub-group who responded correctly to only one of the 13 items - item 2 with 57% of learners answering correctly (*cf.* Table 5.6). The English sub-group answered 3 of the 13 items correctly (items 2, 3 and 4). Item 12 appeared to be the most difficult item as none of the learners across the three languages correctly answered it. Based on the information contained in the above table, the learners who responded to the text in Afrikaans and isiZulu learners had found the text to be difficult. The English sub-group also struggled with the text but had at least 50% of learners which correctly responded to three items above 50%.

5.3 RASCH EVIDENCE FOR ADDRESSING SUB-QUESTION 1

This section provides the overall Rasch results of *The Pearl*, *African Rhinos and Oxpecker Birds* and *Flowers on the Roof* for South African Grade 4 English, Afrikaans and isiZulu learners. In Section 5.3.1, the reliability of *The Pearl* and *African Rhinos and the Oxpecker Birds* texts is given. Thereafter, the individual item-fit statistics (5.3.2) of the texts are discussed. The next section (5.3.3) provides the differential item functioning (DIF) results per item as a possible source of misfit between the three language groups. Section 5.3.4 provides the reliability of the *Flowers on the Roof* text followed by the individual item-fit statistics of the text (5.3.5.). The differential functioning of the *Flowers on the Roof* items is presented in Section 5.3.6. A summary of the problematic items across PIRLS Literacy 2016 released texts is provided in Section 5.4.

5.3.1 Reliability of PIRLS Literacy 2016 Pearl and Rhino Texts

Table 5.7 presents the summary statistics for the passages *The Pearl* and the *African Rhinos and Oxpecker Birds*. For the PIRLS Literacy 2016 analyses, the file *PearlRhino* was created which included the items of *The Pearl* and the *African Rhinos and Oxpecker Birds*. The file also contains the learners who responded to these texts in English, Afrikaans and isiZulu. The summary statistics, in particular the power of analysis and person reliability indices, were calculated for the English, Afrikaans and isiZulu languages combined. This section also provides these statistics for each language.

Table 5.7: Reliability of PIRLS Literacy 2016 Pearl and Rhino Texts

Analysis Name	Power of Analysis of Fit	PerSepIdx with extremes	PerSepIdx no extremes	Coefficient Alpha with extremes	Coefficient Alpha no extremes
PearlRhino [a]	Reasonable	0.53986	0.53596	N/A	N/A
English Only	Good	0.68384	0.67271	0.76884	0.75525
Afrikaans Only	Good	0.66821	0.65360	0.77425	0.75263
isiZulu Only	Too Low	0.30891	0.24412	0.53673	0.48474

[a] The analysis file *PearlRhino* included all languages.

In order to obtain the summary statistics for the PIRLS Literacy 2016 texts, these texts had to be calculated individually. However, this section also provides the summary statistics for the

PearlRhino (combined languages) but it does not include the Cronbach Alpha due to the missing data in the data set (as show in Table 5.7). The summary statistics of the English and Afrikaans groups showed that for these two language groups, the tests were able to differentiate between the different groups of readers (Combrinck, 2019). In other words, the person separation index (*PerSepInx*) shows "how efficiently a set of items is able to separate those persons measured" (Wright & Stone, 1999, p. 151). If the separation is too wide, it signifies that there is a gap between item difficulty and person ability and indicates that the measurement is inaccurate. On the other hand, if the separation is too narrow, the items are seen as redundant as they do not differentiate enough between person abilities (Wright & Stone, 1999; 2004). The α coefficient with no extremes for both English and Afrikaans sub-groups were above .70 whilst the isiZulu sub-group had an α coefficient below the acceptable range (.50) but the power of analysis is too low, meaning that the test cannot differentiate between different groups of people. Based on Fisher's (1992) table for interpreting the person separation index (*PerSepInx*), the isiZulu sub-group can only differentiate one distinct group. To develop greater understanding of the possible misfit of the items, the following section deals with the individual item-fit statistics for the two PIRLS Literacy 2016 released texts.

5.3.2 Individual Item Statistics for Pearl and Rhino Texts

This section provides the individual item-fit information for the overall sample by item and it is ordered by difficulty.[26] Table 5.8 provides evidence of the items and persons and how they link to the fit of the model. The assumption of the model states that as the ability of the persons increases, the probability of responding correctly to more difficult items should also increase (Combrinck, 2019). If there is a lack of fit, it then violates this assumption. Furthermore, the item-fit statistics also presents the chi-square, which shows the invariance across the trait (Pallant & Tennant, 2007). It determines whether the hierarchical ordering of items is consistent across the increasing levels of the trait (Combrinck, 2019). The chi-square is used to ascertain the goodness-of-fit of the item-trait interaction. By using the Rasch measurement theory, it is possible to determine if there is a misfit between the items and the persons (Linacre, 2012) within this study, the items and readers. The mean-square fit statistics include either *overfit* or *underfit*, taking into consideration that the statistics have a range of -5 to 5 with a mean of 0.0 and standard deviation of 1.0 (Bond & Fox, 2015). When the item indicates *overfit*

[26] Also known as item location.

it means that the responses are too predictable while *underfit* indicates too much unpredictability and noise (Linacre, 2002). In other words, when items display *overfit*, it indicates that there is too much discrimination amongst the items while *underfit*, means that the items are discriminating too little between more able and less able learners (van Staden, 2018).

Table 5.8 presents the items for *The Pearl* and the *African Rhinos and the Oxpecker Birds* across languages. In order to test the null hypothesis for this study, the individual item-fit statistics are used. The null hypothesis states that the mean score of the English, Afrikaans and isiZulu learners who completed the PIRLS Literacy 2016 assessment are equal (μEnglish = μAfrikaans = μisiZulu). If, however, the null hypothesis is rejected, then the alternative hypothesis may be accepted. In this case, the alternative hypothesis is that the English, Afrikaans and isiZulu mean scores are not equal to each other (H_a = μEnglish ≠ μAfrikaans ≠ μisiZulu).

Table 5.8: Individual Item-Fit Statistics for Pearl and Rhino Texts

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Rhino Item 7	-1,885	0,098	-1,170	33,23	0,000**
Rhino Item 10	-0,999	0,087	-1,580	13,11	0,158
Rhino Item 6	-0,869	0,085	-3,313*	21,23	0,012
Rhino Item 14	-0,788	0,087	-4,910*	53,68	0,000**
Pearl Item 14	-0,769	0,086	-3,900*	38,64	0,000**
Pearl Item 1	-0,695	0,083	-2,060	18,00	0,035
Pearl Item 13	-0,463	0,055	-1,947	10,91	0,282
Pearl Item 10	-0,373	0,084	1,260	21,36	0,011
Pearl Item 7	-0,346	0,082	-0,426	15,26	0,084
Pearl Item 6	-0,329	0,082	-3,625*	17,89	0,036
Pearl Item 12	-0,291	0,084	-2,180	23,59	0,005
Rhino Item 13	-0,243	0,085	2,117	11,09	0,269
Rhino Item 2	-0,225	0,083	-2,920*	17,29	0,044
Pearl Item 3	-0,219	0,081	0,334	7,98	0,536
Pearl Item 4	-0,185	0,082	-0,709	26,12	0,002
Rhino Item 8	-0,124	0,083	4,578*	55,34	0,000**

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Rhino Item 4	-0,088	0,083	-4,518*	38,13	0,000**
Rhino Item 12	-0,080	0,053	-2,289	13,74	0,132
Pearl Item 8	-0,057	0,082	5,316*	73,33	0,000**
Pearl Item 2	-0,001	0,082	2,743*	22,29	0,008
Rhino Item 1	0,070	0,083	0,670	5,22	0,814
Rhino Item 11	0,110	0,085	-5,274*	42,38	0,000**
Rhino Item 5	0,175	0,083	-5,714*	47,36	0,000**
Rhino Item 9	0,179	0,058	6,773*	85,66	0,000**
Pearl Item 5	0,245	0,055	0,754	16,33	0,060
Rhino Item 15	0,303	0,085	2,708*	39,62	0,000**
Rhino Item 3	0,535	0,085	6,432*	118,51	0,000**
Rhino Item 16	0,917	0,090	-3,526*	51,41	0,000**
Pearl Item 15	0,964	0,065	-0,236	9,73	0,373
Pearl Item 9	1,250	0,092	-0,826	10,28	0,329
Pearl Item 11	1,710	0,101	2,006	46,11	0,000**
Rhino Item 17	2,573	0,124	-1,205	6,53	0,686

* Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000313 for 32 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Eleven of the 32 items displayed significant misfit where the fit residuals were either above 2.5 or below -2.5:

- *Rhino Item 14*, *Pearl Item 14*, *Rhino Item 4*, *Rhino Item 11*, *Rhino Item 5* and *Rhino Item 16* displayed overfit, indicating that these items are discriminating too much among the language sub-groups.
- The five items that displayed underfit (i.e. items with too little discrimination) include *Rhino Item 8*, *Pearl Item 8*, *Pearl Item 2*, *Rhino Item 9*, *Rhino Item 15* and *Rhino Item 3*.

The hypothesis, as mentioned earlier, states that the learners who completed the PIRLS Literacy 2016 test in English, Afrikaans and isiZulu should have equal mean scores, therefore,

it is important to consider the item-fit statistics per language. Tables 5.9, 5.10 and 5.11 present the item-fit statistics for the items per language.

Table 5.9: Individual Item-Fit Statistics for Pearl and Rhino Texts English Only

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Rhino Item 7	-2,013	0,170	-0,652	15,57	0,076
Pearl Item 14	-1,340	0,148	-2,446	24,25	0,004
Rhino Item 4	-1,137	0,148	-3,508*	35,57	0,000**
Rhino Item 14	-0,946	0,148	-2,002	15,07	0,089
Pearl Item 1	-0,838	0,137	-1,062	17,35	0,043
Pearl Item 7	-0,535	0,134	-0,331	5,55	0,784
Rhino Item 2	-0,521	0,137	-1,849	15,04	0,090
Pearl Item 13	-0,494	0,092	-1,064	9,99	0,351
Pearl Item 12	-0,448	0,135	-0,768	9,18	0,421
Rhino Item 5	-0,402	0,137	-2,272	23,23	0,006
Rhino Item 10	-0,392	0,136	-1,071	6,96	0,642
Rhino Item 6	-0,342	0,137	-4,252*	24,39	0,004
Pearl Item 6	-0,331	0,134	-3,913*	22,99	0,006
Pearl Item 2	-0,295	0,132	0,959	13,28	0,150
Pearl Item 10	-0,193	0,133	0,293	10,22	0,333
Pearl Item 3	-0,139	0,131	-0,155	11,12	0,267
Rhino Item 1	-0,104	0,136	-0,769	4,16	0,900
Rhino Item 12	-0,052	0,091	-2,257	8,40	0,494
Pearl Item 4	-0,018	0,131	0,516	14,47	0,107
Rhino Item 8	0,040	0,133	3,789*	48,17	0,000**
Rhino Item 13	0,058	0,136	0,267	10,00	0,351
Pearl Item 8	0,150	0,130	3,295*	47,38	0,000**
Pearl Item 5	0,377	0,088	1,313	18,55	0,029
Rhino Item 11	0,383	0,136	-3,925*	29,59	0,001
Rhino Item 9	0,386	0,091	5,859*	90,55	0,000**
Rhino Item 15	0,440	0,135	1,048	15,92	0,069

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Rhino Item 16	0,479	0,137	-2,310	22,59	0,007
Pearl Item 15	0,927	0,096	0,791	9,67	0,377
Rhino Item 3	0,931	0,136	4,351*	72,67	0,000**
Pearl Item 9	1,571	0,142	-0,654	13,23	0,153
Pearl Item 11	2,117	0,156	0,640	8,50	0,484
Rhino Item 17	2,683	0,179	-0,724	5,15	0,821

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000313 for 32 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Table 5.9 depicts the item-fit statistics of *The Pearl* and *African Rhinos and the Oxpecker Birds* texts for the English language sub-set. Of the 32 items, only five of these items displayed misfit, which is significant. *Rhino Item 4* was the only item that displayed overfit. This item was too easy to discriminate against the English sub-group. The remaining four items, namely *Rhino Item 8*, *Pearl Item 8*, *Rhino Item 9* and *Rhino Item 3*, displayed underfit. These underfitting items do not adequately discriminate between the more and less able English learners.

Table 5.10 shows the item-fit statistics for the Afrikaans language sub-group.

Table 5.10: Individual Item-Fit Statistics for Pearl and Rhino Texts Afrikaans Only

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Rhino Item 7	-2,226	0,219	-1,475	23,91	0,004
Rhino Item 10	-1,501	0,199	-0,474	10,52	0,310
Pearl Item 1	-1,017	0,182	-0,274	8,49	0,486
Pearl Item 14	-0,814	0,183	-1,270	13,99	0,123
Rhino Item 6	-0,774	0,179	-1,210	8,01	0,533
Rhino Item 14	-0,729	0,180	-3,009*	17,60	0,040
Pearl Item 13	-0,482	0,117	-0,569	10,40	0,319
Pearl Item 2	-0,359	0,174	0,869	17,83	0,037
Pearl Item 3	-0,359	0,173	1,339	12,43	0,190
Rhino Item 2	-0,347	0,174	-0,882	8,11	0,523

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Pearl Item 4	-0,257	0,171	-1,082	14,56	0,104
Pearl Item 7	-0,255	0,174	0,317	8,03	0,531
Pearl Item 10	-0,252	0,177	1,041	11,47	0,245
Pearl Item 5	-0,225	0,119	-0,420	5,81	0,759
Pearl Item 8	-0,084	0,172	2,785*	23,84	0,005
Pearl Item 6	-0,082	0,172	-1,576	14,38	0,110
Rhino Item 8	0,022	0,173	1,189	11,12	0,267
Rhino Item 1	0,088	0,171	0,968	7,05	0,632
Pearl Item 12	0,141	0,176	-3,202*	20,69	0,014
Rhino Item 5	0,144	0,171	-3,063*	15,85	0,070
Rhino Item 11	0,203	0,175	-3,195*	23,06	0,006
Rhino Item 9	0,215	0,125	3,418*	41,04	0,000**
Rhino Item 4	0,318	0,172	-0,229	7,75	0,559
Rhino Item 12	0,435	0,114	-1,418	8,22	0,512
Rhino Item 15	0,485	0,176	2,307	19,36	0,022
Rhino Item 13	0,492	0,176	0,643	5,17	0,819
Rhino Item 3	0,541	0,175	2,133	42,71	0,000**
Pearl Item 9	0,822	0,178	-0,279	15,78	0,072
Pearl Item 15	0,963	0,133	0,715	10,52	0,310
Rhino Item 16	1,300	0,189	-0,847	17,49	0,042
Pearl Item 11	1,487	0,200	1,155	21,64	0,010
Rhino Item 17	2,108	0,220	-0,252	7,18	0,619

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000313 for 32 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Only one item displayed misfit and was significant, namely *Rhino Item 9*. This item displayed in particular, underfit as it does not adequately discriminate between the different ability groups of learners. The Afrikaans sub-group had a slightly better fit in comparison to the other two languages.

Table 5.11 shows the item-fit statistics for the isiZulu language sub-group.

Table 5.11: Individual Item-Fit Statistics for Pearl and Rhino Texts isiZulu Only

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Rhino Item 7	-1,646	0,144	0,425	9,33	0,408
Rhino Item 6	-1,417	0,140	-1,557	6,73	0,666
Rhino Item 10	-1,411	0,141	-2,376	18,93	0,026
Rhino Item 13	-1,021	0,140	0,683	11,18	0,264
Rhino Item 14	-0,673	0,138	-4,078*	27,03	0,001
Pearl Item 10	-0,646	0,136	1,262	17,25	0,045
Rhino Item 12	-0,532	0,081	-1,561	9,55	0,388
Pearl Item 6	-0,467	0,132	-0,364	6,17	0,723
Pearl Item 13	-0,443	0,086	-1,350	13,96	0,124
Rhino Item 8	-0,405	0,132	2,840*	19,93	0,018
Pearl Item 1	-0,385	0,132	-1,501	6,55	0,684
Pearl Item 12	-0,375	0,137	-0,375	9,17	0,422
Pearl Item 4	-0,323	0,132	-1,448	11,59	0,237
Pearl Item 8	-0,267	0,135	3,848*	29,15	0,001
Pearl Item 14	-0,249	0,140	-1,149	30,94	0,000**
Pearl Item 7	-0,216	0,134	0,676	14,32	0,111
Pearl Item 3	-0,213	0,133	-0,039	18,11	0,034
Rhino Item 11	-0,208	0,137	-2,410	14,31	0,112
Rhino Item 9	-0,043	0,093	1,249	13,80	0,130
Rhino Item 15	0,032	0,138	1,612	13,50	0,141
Rhino Item 3	0,116	0,137	3,651*	31,04	0,000**
Rhino Item 2	0,126	0,137	-0,665	3,24	0,954
Rhino Item 1	0,239	0,139	2,027	8,36	0,498
Pearl Item 5	0,353	0,092	0,042	4,38	0,885
Pearl Item 2	0,527	0,143	3,802*	44,76	0,000**
Rhino Item 4	0,550	0,144	-2,581*	16,33	0,060
Rhino Item 5	0,854	0,151	-2,560*	19,30	0,023

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Pearl Item 15	1,192	0,121	-1,003	5,97	0,743
Rhino Item 16	1,211	0,169	-2,341	30,28	0,000
Pearl Item 9	1,267	0,167	-0,105	7,73	0,561
Pearl Item 11	1,459	0,178	1,448	38,50	0,000**
Rhino Item 17	3,014	0,303	-0,789	12,69	0,177

*Fit residuals are indicated if above +2.5 or below -2.5

**Bonferroni adjustment is 0.000313 for 32 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Two of the 32 items displayed significant misfit, namely *Rhino Item 3* and *Pearl Item 2*. Both of these items showed underfit as the items do not adequately discriminate between the more and less able learners. Based on the evidence from the three item-fit statistics, by language, it would appear that there is some variation between the languages as the English sub-group had five items that displayed significant misfit whereas the learners who wrote the test in Afrikaans and isiZulu had only one and two items, respectively, that had significant misfit. As such, the mean scores are not equal between the three language sub-groups. As a result, this finding partially provides evidence for rejecting the null hypothesis.

The following section examines the differential item functioning (DIF) of *The Pearl* and *African Rhinos and the Oxpecker Birds* PIRLS Literacy 2016 texts for the learners who wrote the assessment in English, Afrikaans and isiZulu.

5.3.3 Differential Item Functioning for Pearl and Rhino Texts

When items function differently for persons from different groups who have the same ability, it means that the persons do not have the same chance of correctly responding to an item (Sandilands, Oliveri, Zumbo & Ercikan, 2013). An ANOVA test is given along with the differential item functioning (DIF) in Table 5.12. It provides the *F*-ratio as well as the *p*-value. The former is a measure where at least two quantities are expected to be equal under the null hypothesis. For the purposes of this study, the quantities are the three language sub-groups. The latter is the statistical model used to provide evidence whether the null hypothesis should be either accepted or rejected. The ANOVA was performed to compare the means of the three languages, namely English, Afrikaans and isiZulu, in order to test the null hypothesis where

the group means are equal (Field, 2009; Tabachnick & Fidell, 2007). The ANOVA provides the values which are necessary to inform the researcher whether the differences between the groups' mean scores are statistically significant. The next table gives the DIF summary for the PIRLS Literacy 2016 released reading texts.

Table 5.12: DIF Summary for PIRLS Literacy 2016 Pearl and Rhino Texts

Item	F-ratio	Probability
Rhino Item 7	3,103	0,045
Rhino Item 10	13,704	0,000*
Rhino Item 6	13,524	0,000*
Rhino Item 14	2,324	0,099
Pearl Item 14	19,097	0,000*
Pearl Item 1	6,810	0,001
Pearl Item 13	1,555	0,212
Pearl Item 10	2,032	0,132
Pearl Item 7	1,812	0,164
Pearl Item 6	1,517	0,220
Pearl Item 12	3,519	0,030
Rhino Item 13	23,910	0,000*
Rhino Item 2	8,386	0,000*
Pearl Item 3	0,347	0,707
Pearl Item 4	1,019	0,362
Rhino Item 8	1,635	0,196
Rhino Item 4	52,452	0,000*
Rhino Item 12	16,706	0,000*
Pearl Item 8	1,849	0,158
Pearl Item 2	13,616	0,000*
Rhino Item 1	2,502	0,083
Rhino Item 11	3,120	0,045
Rhino Item 5	33,405	0,000*
Rhino Item 9	3,606	0,028

Item	F-ratio	Probability
Pearl Item 5	11,554	0,000*
Rhino Item 15	1,758	0,173
Rhino Item 3	9,208	0,000*
Rhino Item 16	13,884	0,000*
Pearl Item 15	0,484	0,616
Pearl Item 9	3,894	0,021
Pearl Item 11	3,739	0,024
Rhino Item 17	3,949	0,020

*Significant at the 5 percent level (Bonferroni 0.000521)

Table 5.12 provides the DIF summary for the two PIRLS Literacy 2016 texts, namely *The Pearl* and *African Rhinos and the Oxpecker Birds*, as demonstrated by ANOVA statistics. Results with a small *p*-value, less than <0.05, are significant in terms of uniform DIF. Uniform DIF is when persons with the same underlying ability have consistently different probability of responding correctly to an item when compared to a different group (Andrich, Sheridan & Luo, 2012). The above table shows that 12 out of 32 items displayed differential functioning across the English, Afrikaans and isiZulu language sub-groups. The items include *Rhino Item 10*, *Rhino Item 6*, *Pearl Item 14*, *Rhino Item 13*, *Rhino Item 2*, *Rhino Item 4*, *Rhino Item 12*, *Pearl Item 2*, *Rhino Item 5*, *Pearl Item 5*, *Rhino Item 3* and *Rhino Item 16*. For each of the aforementioned items displaying DIF, their item characteristic curves (ICC) are examined.

The ICC figures graphically shows each item that has displayed DIF. In this section, these items are described with reference to the ICC, and where the item is a multiple choice question, a distractor analysis is provided. These ICC graphs shows the differential functioning of the specific item across the lower and/or upper class interval.[27]

Item 10 from the African Rhinos and the Oxpecker Birds is a multiple choice (MC) item and requires learners to *Focus on and Retrieve Explicitly Stated Information* by asking: What do ticks need to live?

[27] Lower and upper class intervals look at the spread of observations between certain points on the x-axis of the graph (Van Staden, 2018).

a. trees
b. rhino
c. bushes
d. blood* (correct answer)

Figure 5.2 shows that the item was relatively more difficult for the learners tested in English (blue). More often than not, the English sub-group fell below the IRT model (grey) except at the 2.3 person location. All three language sub-groups between the -3 and -2 person location had less than 20% chance of correctly responding to the item. Extensive inconsistency occurs between the -1 and 1 person locations. Around the -1 person location, the Afrikaans (approximately 70%) sub-group had the highest probability of correctly responding to the item compared to English (approximately 30%) and isiZulu (approximately 40%) sub-groups. The pattern changes slightly at the 0-person location where the learners, who completed the test in Afrikaans, had a 100% probability of correctly responding to the item. The isiZulu sub-group had about an 80% and the English sub-group had about a 50% chance. The item appears to discriminate against the learners at the lower class interval (between -3 and 0). Overall, the item also appears to discriminate against the English sub-group as it seems a higher cognitive level was required by these learners to answer this item.

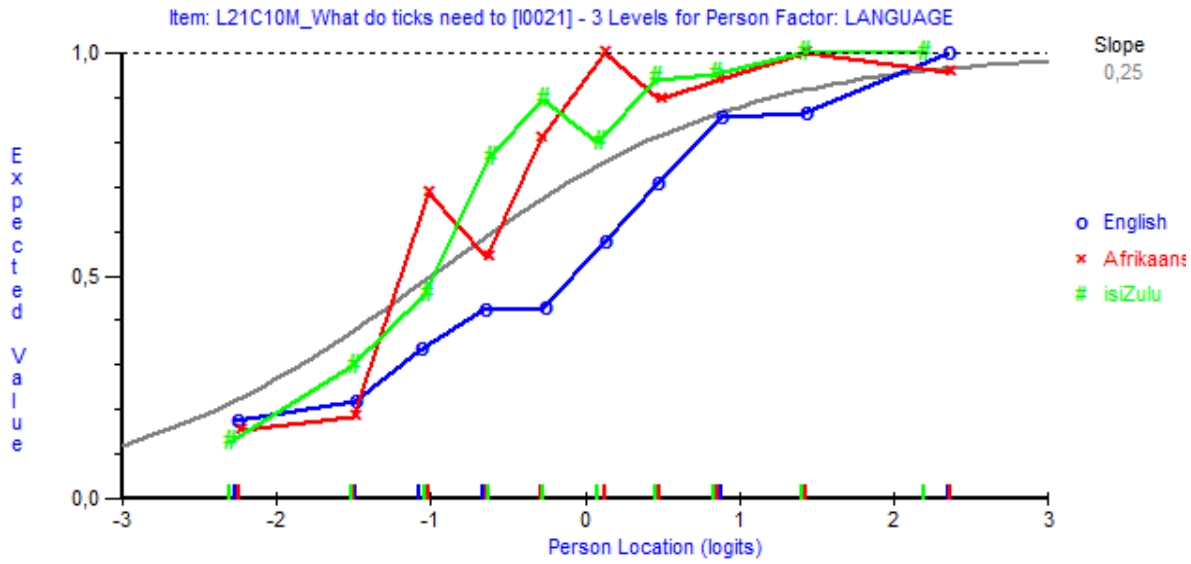

Figure 5.2: Rhino Item 10 Characteristic Curve

The next figure shows the distractor analysis for *Rhino Item 10*. Distractor D(4) was the correct answer although between the -3 and -2 person location, learners were tempted to select distractor B(2). Learners had about a 40% chance of selecting distractor B(2).

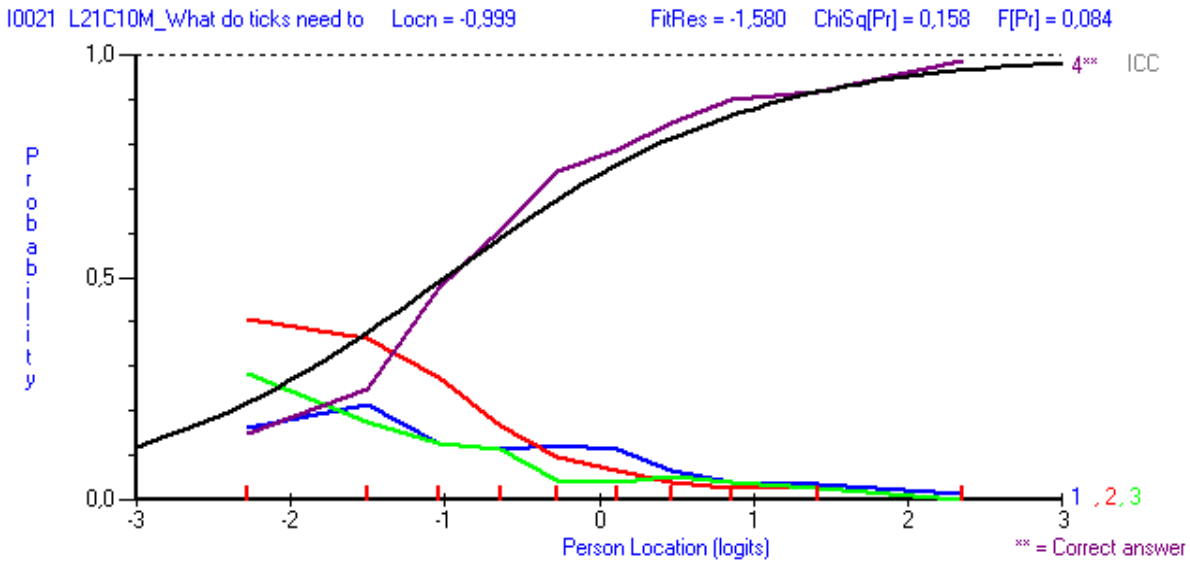

Figure 5.3: Rhino Item 10 Distractor Analysis

Distractor B(2) pulled some attention from the learners between the -3 and -2 person location and had approximately a 40% probability of selecting the distractor. It could be that the misunderstanding stems from the specific versus the general. In other words, it being the rhino specifically that the ticks feed from, yet the text states that the ticks live off the blood of animals. Distractor C(3) was also tempting for the same group of learners as they had about a 30% chance of selecting it. At the -1 person location, the learners saw distractor D(4) as a more plausible answer to the question. Learners across the higher class interval dismissed distractor A(1), B(2) and C(3) as possible answers.

The next item that indicated differential functioning across the three languages was *Rhino Item 6* (as illustrated in Figure 5.4). This item required learners to *Make Straightforward Inferences*. The item asked the following: Why do hunters want to kill rhinos?

a. Rhinos are too dangerous
b. Hunters want rhino meat
c. Hunters want rhino horns* (correct answer)
d. There are too many rhinos

Figure 5.4 shows that *Rhino Item 6* was considerably less difficult for the learners tested in isiZulu (green). The isiZulu language sub-group were above the expected model curve at all points. However, the learners who completed the test in English found the item somewhat more

difficult than the other two languages. There is some inconsistency at the lower class interval (between -3 and 0). The English sub-group at -2.2 had less than 10% probability of correctly responding to the item, whereas both the Afrikaans and isiZulu sub-groups had approximately 20%. At the -1 person location, the English sub-group had about a 10% chance of correctly responding to the item. Again, both the Afrikaans and isiZulu sub-groups had an approximately 50% probability of correctly responding to the item. At the 0.8 person location, the learners who completed the test in English, had an 85% probability of correctly responding to the item. The Afrikaans sub-group, however, had a slightly lower chance of getting the item right. The isiZulu sub-group from person location 1 onwards had a 100% chance of correctly responding to the item. This item discriminated more against the English and Afrikaans sub-group at the lower class interval.

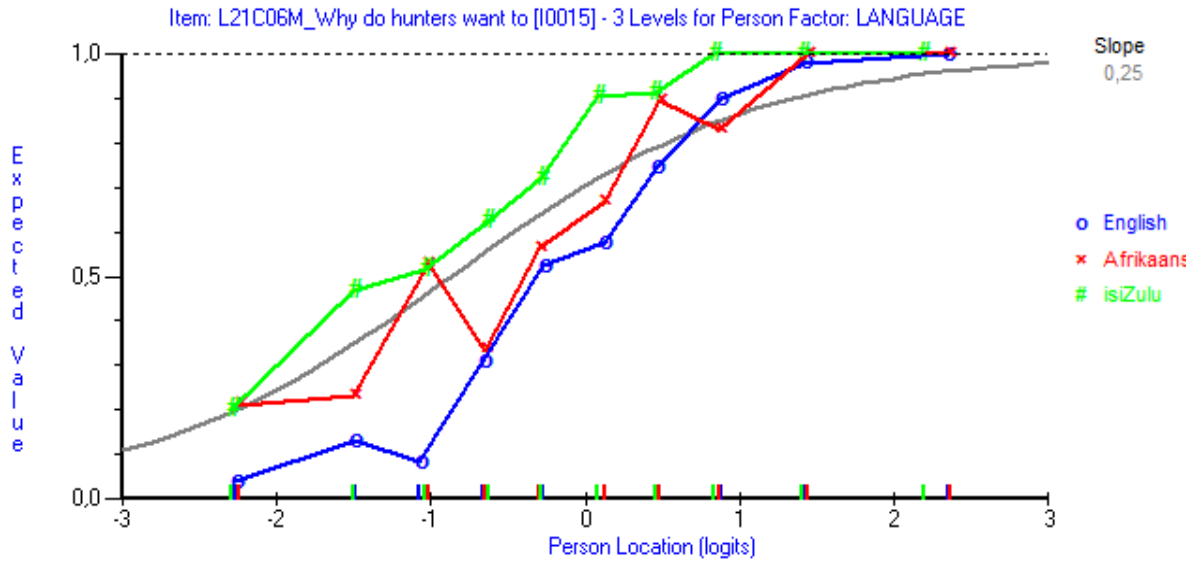

Figure 5.4: Rhino Item 6 Characteristic Curve

Figure 5.5 presents the distractor analysis for *Rhino Item 6*. The correct answer is Distractor C(3), although learners between -3 and -2 person location were tempted by distractor A(1) and had about a 40% chance of selecting it.

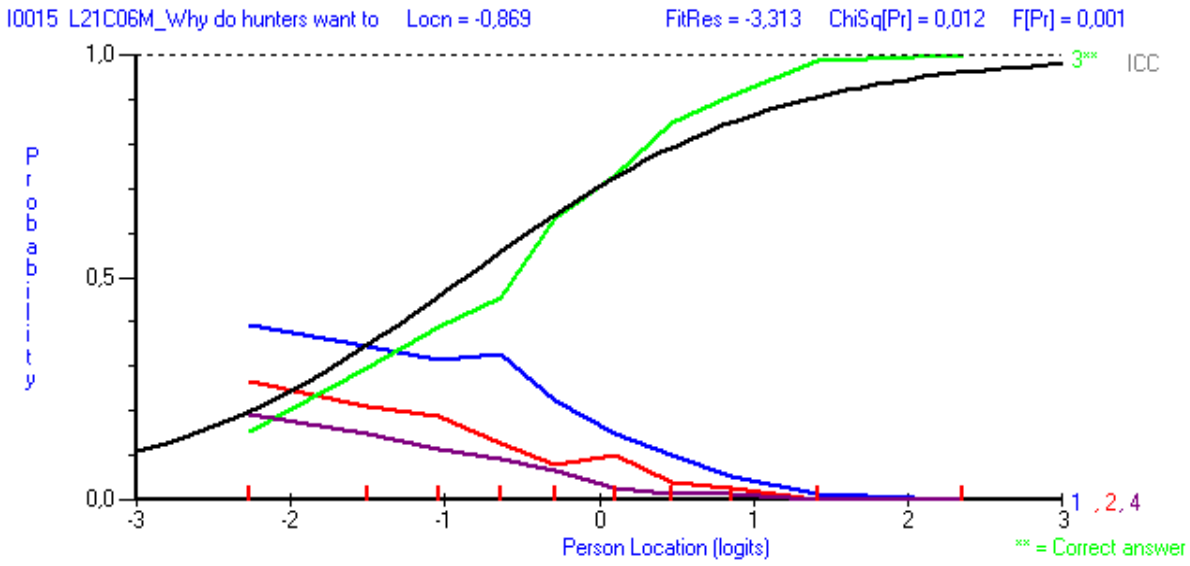

Figure 5.5: Rhino Item 6 Distractor Analysis

The learners between the -3 and -2 person location had slightly less than 30% to select distractor B(2). Distractor A(1) stated that rhinos are dangerous and distractor B(2) stated hunters want the rhino meat. These two options seemed plausible to the learners and could be due to the fact that learners used local knowledge that rhinos are dangerous animals and that hunting the rhino would mean that the hunters may have wanted the meat, as that was the traditional role of a hunter. Between person locations -1 and 0, the learners were able to distinguish between the less plausible answers and the more likely answer. The learners at the upper class interval found the item relatively more easy and at the 1 person location, they had approximately a 90% chance of selecting the correct answer. In addition, the chi-square of the item distractors is significant. This means that the distribution of distractors is significantly different from the expected model.

Pearl Item 14 displayed DIF among the English, Afrikaans and isiZulu sub-groups (as shown in Figure 5.6). To be able to answer this item correctly, learners had to *Focus on and Retrieve Explicitly Stated Information*. The question asked: What does Josh say they should do with Reuben's money?

a. Get a new house
b. Buy lots of pearls
c. Share it with everyone* (correct answer)
d. Take it back to the city

It would seem that there is considerable inconsistency across the three languages, especially at the lower class interval (between -3 and 0) (*cf.* Figure 5.6). The learners who completed the test in English had a steady increase in probability to correctly answer the item between the -3 and 0 person locations. Between the -3 and -2 person location, the Afrikaans sub-group had approximately a 30% probability of correctly responding to the item, whereas the English and isiZulu sub-groups had less than 20%. Between the -2 and -1 person locations, extreme inconsistency occurred as the Afrikaans sub-group at -1.5 had less than 10% correctly responding to the item and at -1, the same learners had approximately 40%. In addition, at the -1 person location, the isiZulu sub-group had about 10% chance of answering the item correctly. At the higher class interval, specifically between 0 and 1, the English and Afrikaans sub-groups were consistently above the model curve, while the isiZulu learners at 0.8 had a 100% chance of correctly responding to the item. Overall, the item discriminated in favour of the Afrikaans and isiZulu learners at the lower class interval.

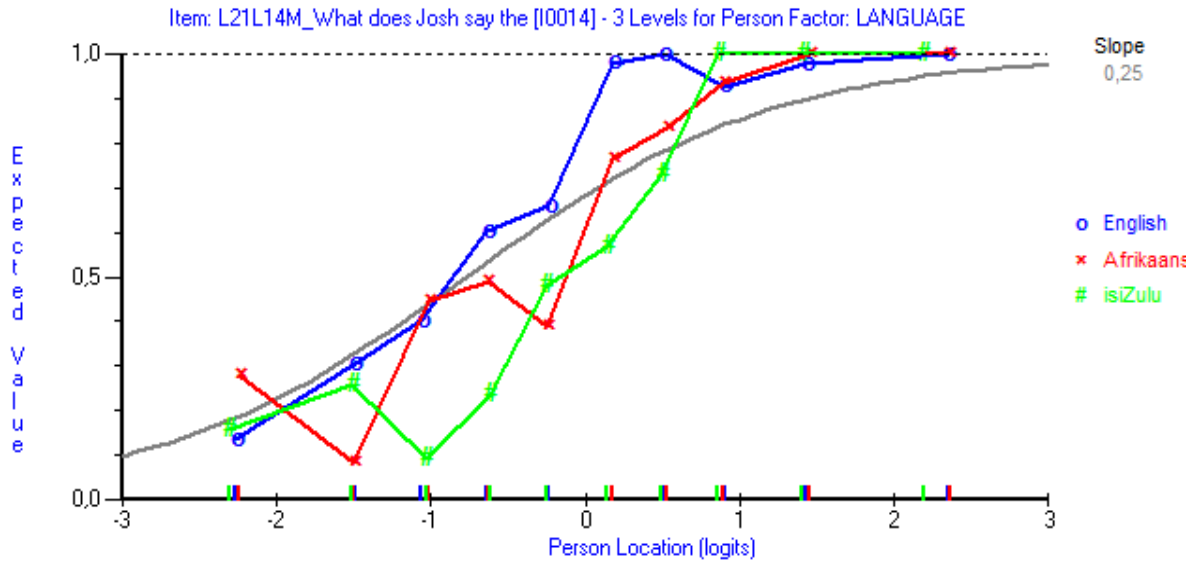

Figure 5.6: Pearl Item 14 Characteristic Curve

The next figure shows the distractor analysis for *Pearl Item 14*. Distractor C(3) was the correct answer.

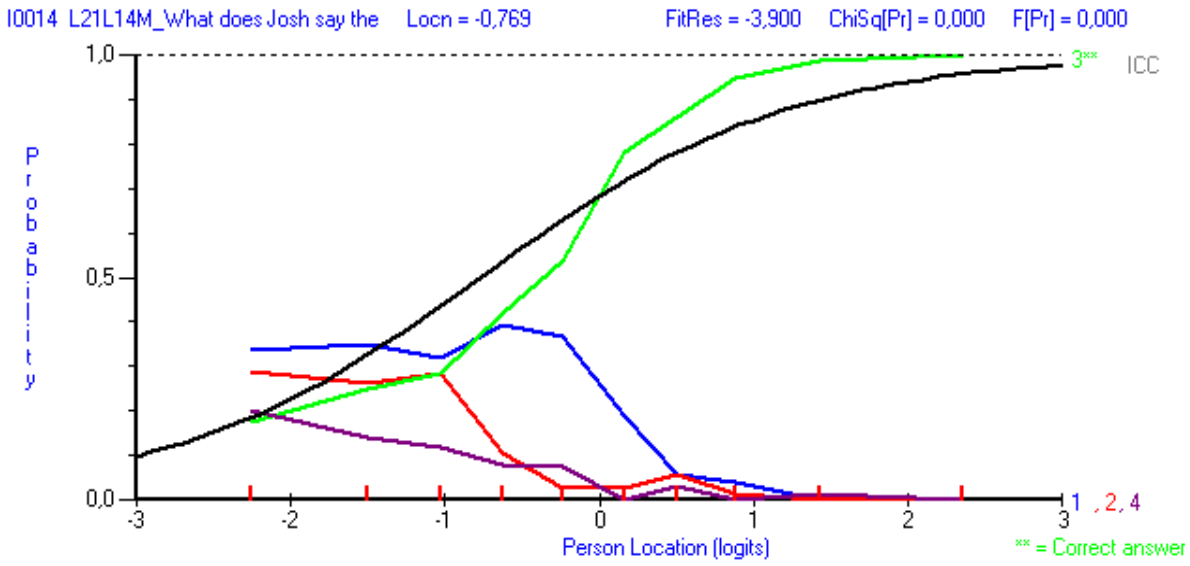

Figure 5.7: Pearl Item 14 Distractor Analysis

The above figure (5.7) shows that learners across the lower class interval struggled to ascertain the correct answer from the four options. Distractors A(1) and B(2) were too tempting for learners between -3 and -2 person location as learners may have relied on local knowledge. These learners had approximately a 30% chance of selecting either distractor. Distractor D(4) across the class intervals became increasingly implausible, meaning that very few learners selected it. This distractor stated that Reuben would take the money back to the city. Learners had to find the answer from the text and distractor D(4) was not part of the text. Learners across the higher class interval had noticed that the only plausible answer was distractor C(3) as it was taken from the text. Moreover, the ICC graph shows that the chi-square is statistically significant which shows that the distribution of distractors is significantly different from the expected model.

The next item displaying differential functioning is *Rhino Item 13* (as illustrated in Figure 5.8). Learners were required to *Make Straightforward Inferences*. The question asked the following: Why do oxpeckers especially like to eat ticks?

a. Because the ticks have been sucking blood* (correct answer)
b. Because there are many ticks on each rhino
c. Because the ticks have been eating leaves
d. Because the ticks are tiny and easy to eat

Figure 5.8 shows that the item was considerably easier for the learners who completed the test in isiZulu. The isiZulu characteristic curve was above that of the model curve across both the lower and upper class intervals, with learners from person location 1 onwards having a 100% probability of correctly responding to the item. The Afrikaans sub-group found the item much more difficult than the isiZulu sub-group. Across the lower class interval, the Afrikaans sub-group had less than a 25% chance of correctly responding to the item. The learners who completed the test in English followed a similar pattern at the lower class interval, with learners having at most a 30% probability of answering the item correctly. At the 0.8 person location, the Afrikaans sub-group had an approximate 40% chance of correctly responding to the item, the English sub-group had about a 60% chance and the isiZulu sub-group had approximately an 80% chance. At the 1.4 person location, the Afrikaans sub-group had a sharp increase in the chance of correctly answering the item with a 70% chance, English roughly the same percentage and the isiZulu had a 100% chance. The item showed discrimination towards the English and Afrikaans sub-group, but particularly in respect of the Afrikaans sub-group.

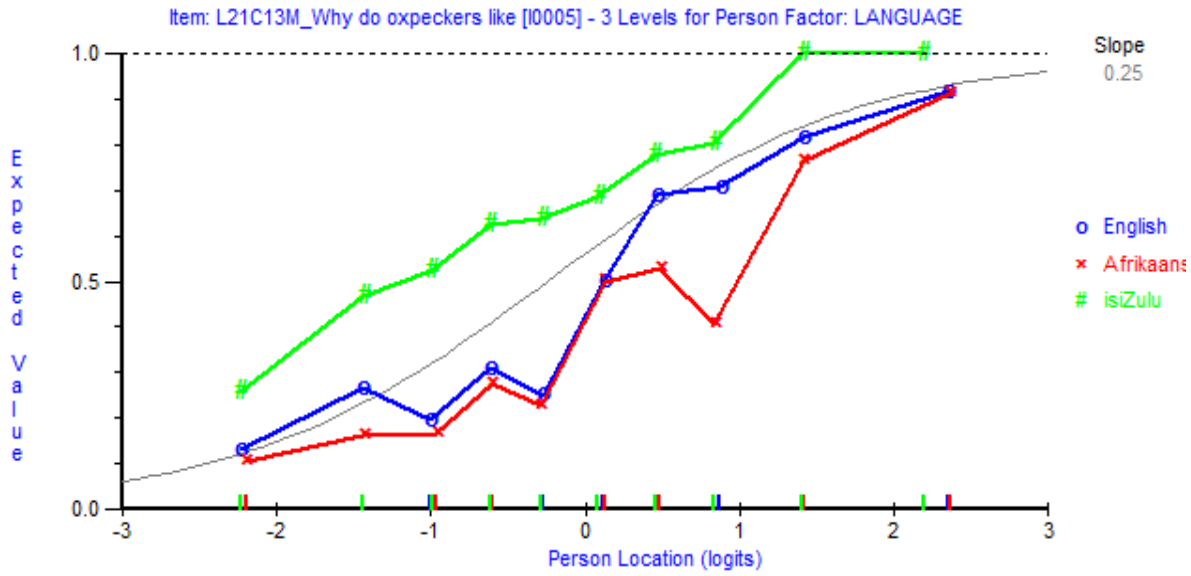

Figure 5.8: Rhino Item 13 Characteristic Curve

The following figure (5.9) shows the distractor analysis for *Rhino Item 13*. Distractor A(1) was the correct answer to the item.

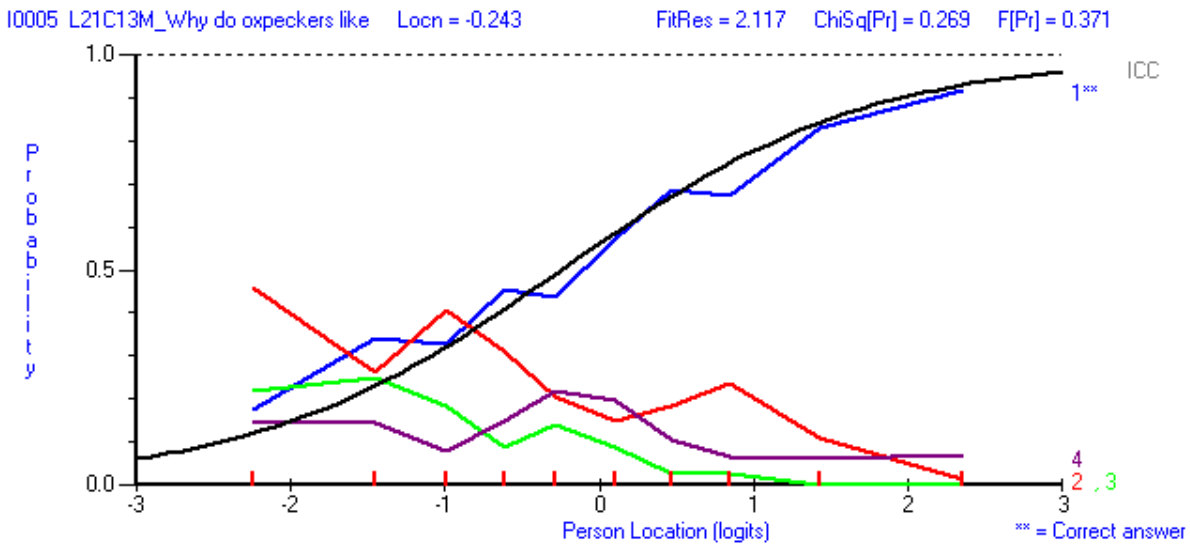

Figure 5.9: Rhino Item 13 Distractor Analysis

Based on the above figure, learners between -3 and -2 person location had almost 50% probability of selecting distractor B(2) and only about a 20% chance of selecting the correct answer. Learners may have accessed information from the text that talked about ticks living off the rhino. However, learners should have inferred from the text that the reason why the oxpecker birds especially eat ticks, is that they are full of blood sucked from the rhino. Neither distractor C(3) or D(3) had drawn much attention as learners were able to deduce that these two options were not possible answers to the item. Across the upper class interval, learners were able to identify incorrect answers and select the correct distractor. Although at the 1 person location, learners were tempted again by distractor B as the correct answer.

Item 2 from the *African Rhinos and the Oxpecker Birds* text showed differential functioning among the English, Afrikaans and isiZulu sub-groups (as illustrated in Figure 5.10). The item requires learners to *Evaluate and Critique Content and Textual Elements* by asking the following: What does the picture in the big red circle help you understand?

a. how oxpeckers fly
b. what oxpeckers sound like
c. an oxpecker's nest
d. what oxpeckers look like* (correct answer)

Figure 5.10 presents the item characteristic curve for *Rhino Item 2*. There is inconsistency between the three language groups across the lower class interval (between -3 and 0) until 0.1

at the upper class interval. At the -2.2 person location, the learners who completed the test in Afrikaans and isiZulu, had less than 10% probability of correctly responding to the item. The English sub-group at the same person location, had a slightly higher percentage of correctly answering the item. At around the -1.5 person location, the English and Afrikaans sub-groups had less than 10% of correctly responding to the item whereas the isiZulu sub-group had approximately 15%. Between the -1 and 0 person location, greater inconsistency occurs between the three language sub-groups. At this interval, the isiZulu sub-group found the item considerably more difficult that the English and Afrikaans sub-groups. At the -1 person l location, all three language sub-groups were below the IRT model curve and at -0.6 person location, the isiZulu sub-group had approximately a 30% chance of correctly responding to this item. The English sub-group had an approximate 40% chance whereas the learners who completed this test in Afrikaans, had the highest chance (60%) of correctly responding to the item. At around the 0.1 person location, extreme inconsistency occurred where the isiZulu sub-group remained below the model curve and had less than 40% probability of correctly answering the item. The Afrikaans sub-group had an approximate 60% chance and the English sub-group had a sharp increase (80%) in their chance of correctly responding to the item. From 0.8 person location onwards, all three language sub-groups were above the expected model curve with the isiZulu learners at the 2.2 person location, having 100% probability of correctly responding to the item.

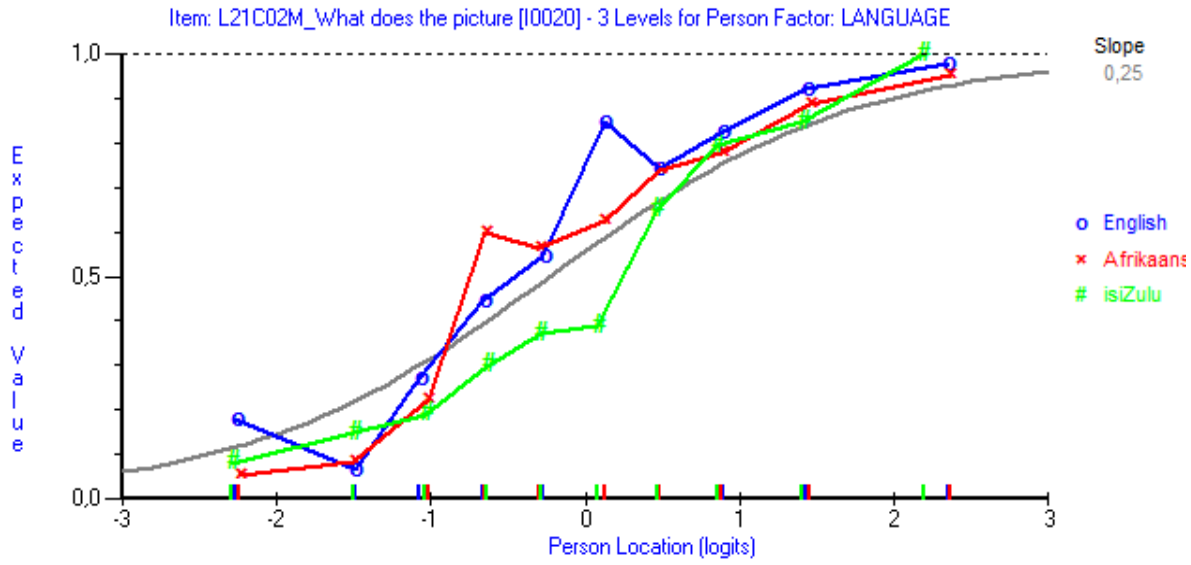

Figure 5.10: Rhino Item 2 Characteristic Curve

The correct answer for *Rhino Item 2* is distractor D(4). Figure 5.11 shows the distractor analysis for the item.

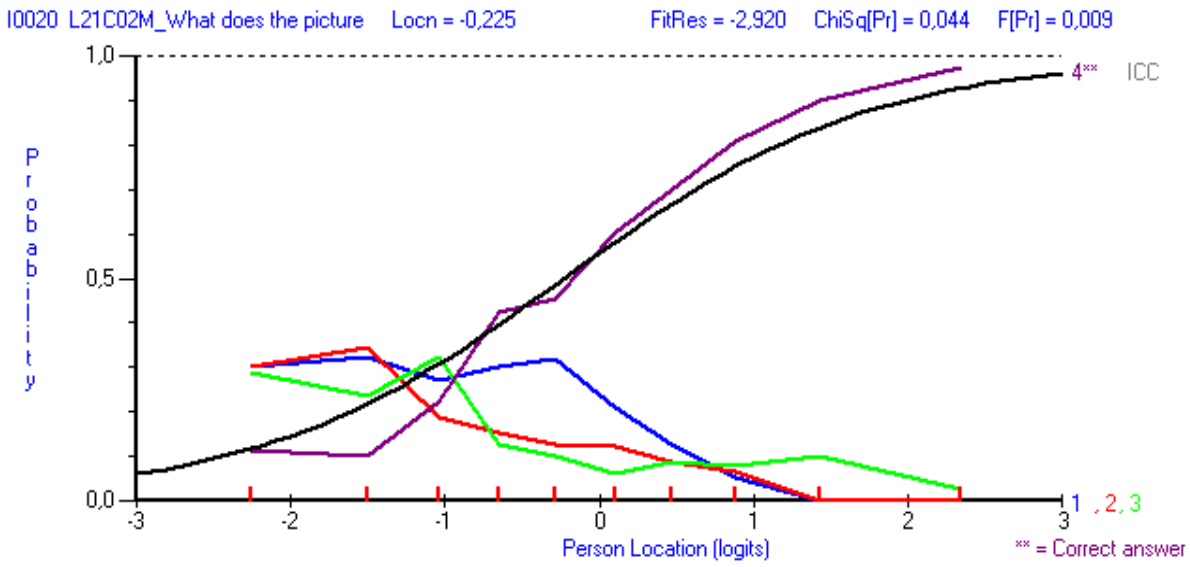

Figure 5.11: Rhino Item 2 Distractor Analysis

The above figure shows the inconsistency across the lower class interval (-3 to 0). Between the -3 and -2 person location, the learners had approximately 10% probability of selecting the correct answer (distractor D). These learners were tempted by distractors A(1), B(2) and C(3) and they had an approximate 30% chance of selecting either of the aforementioned distractors. The item referred specifically to the first picture of the text that shows oxpecker birds sitting on a rhino. The big red circle was an enlargement of the oxpecker birds. Learners were asked to identify what the picture represented. However, it should be noted that for the South African study, the test was not available in colour and as a result, where the item referred to a 'big red circle' the learners where not able to see the 'big red circle' as it was in fact grey. The national version of the question omitted the word 'red' and as such only referred to the 'big circle'. Despite this inconsistency from the international version, the learners at the upper class interval were able to select the correct answer even though it was not in colour, by looking at the picture and reading the text below it.

Rhino Item 4 also displayed differential functioning across the three language sub-groups (as shown in Figure 5.12). It required learners to *Focus on and Retrieve Explicitly Stated Information* from the text in order to answer the following question: Look at the chart. How much does an oxpecker weigh? *Rhino Item 4* is a constructed response (CR) type question with a mark allocation of one (see Appendix C for the complete example of this item).

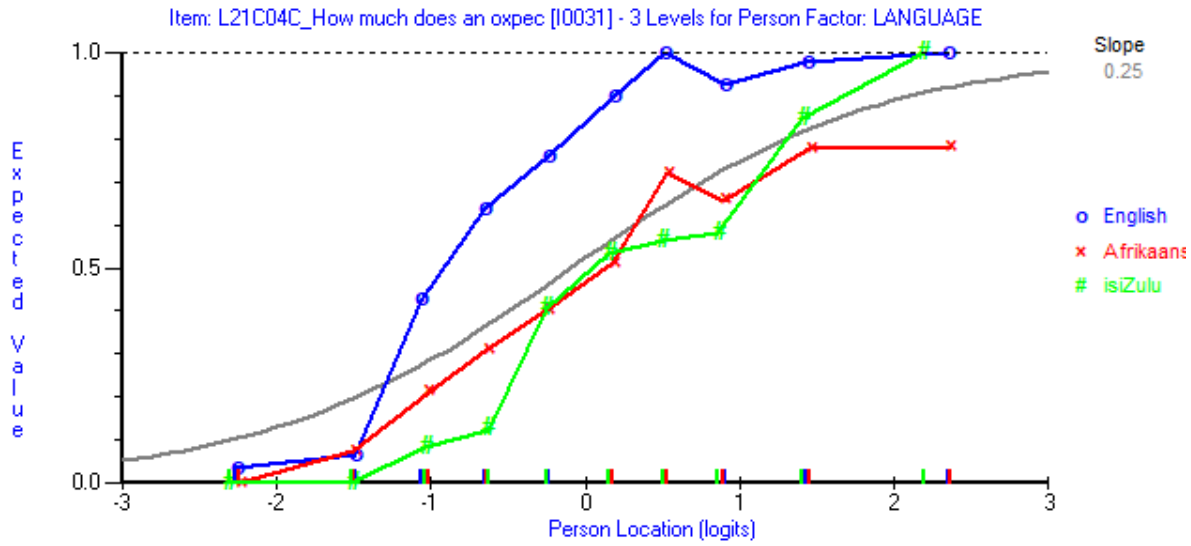

Figure 5.12: Rhino Item 4 Characteristic Curve

There is extensive inconsistency across the lower and upper class intervals among the three language sub-groups (*cf.* Figure 5.12). It would appear that the item was considerably less difficult for the English sub-group compared to the other two language sub-groups. At the -2.2 and -2.3 person locations, the three language groups had almost no chance of correctly responding to the item. At the next interval, the isiZulu learners remain at zero percent of correctly answering the item whereas the English and Afrikaans sub-groups had approximately 5%. At the -1.5 person location, the Afrikaans and isiZulu sub-groups remained below the model curve with approximately 20% and 10% respectively, to correctly responding to the item. Between the -1 and 0 interval, some inconsistency occurred where the isiZulu sub-group at the -0.6 person location had approximately 10% probability of correctly responding to the item and at -0.4, the same sub-group had approximately 40% chance. The English sub-group at these two person locations had an approximate 60% and 70% chance of correctly responding to the item. Between the 0 and 1 interval, the Afrikaans sub-group indicated fluctuation between the three person locations 0.1, 0.5 and 0.9. Between the 2 and 3 intervals, both the English and isiZulu sub-groups had 100% probability of correctly responding to the item, whereas the learners who completed the test in Afrikaans, had an approximate 70% chance. The English sub-group had an approximate 40% chance and was above the model curve. An acceptable response for one mark would be 60 grams. If the learner only provided '60' without

the unit of measurement, the learner would have received an incorrect (zero mark)[28]. Figure 5.12 gives evidence that there is discrimination towards the Afrikaans and isiZulu sub-groups.

The next item that displayed DIF is *Rhino Item 12*, which is a constructed response type question and requires learners to *Focus on and Retrieve Explicitly Stated Information*. The item asks: What are these parts of an oxpecker like? 1. Its tail is…; 2. Its claws are… The item consists of two marks, one mark for describing what the tail is like and the second for what its claws are like. Figure 5.13 displays the item characteristic curve of *Rhino Item 12*.

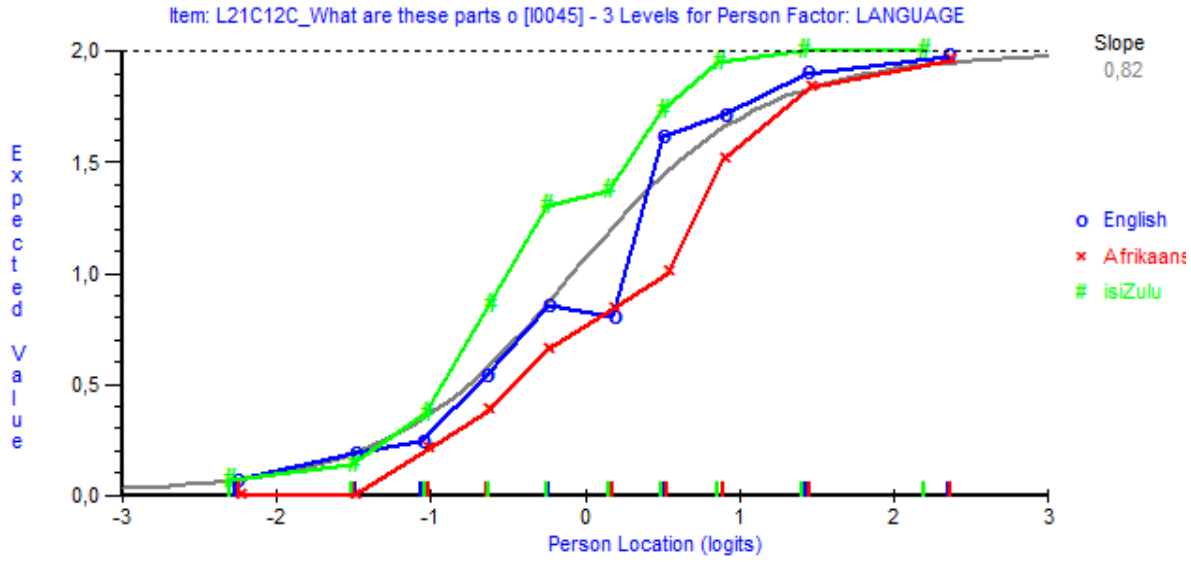

Figure 5.13: Rhino Item 12 Characteristic Curve

The learners who completed the test in English had a characteristic curve close to the model curve. More often than not, the learners who completed the test in Afrikaans found this item more difficult compared to the other two language sub-groups. The Afrikaans sub-group at -2.2 and -1.5 person locations had zero percent chance of obtaining full marks for the item. The English and isiZulu sub-groups at the same person locations had about a 20% chance of responding partially correctly to the item – one out of two marks. The ICC for all three languages across the lower class interval remained relatively close to each other, but at the 0 and 1 interval, greater inconsistency occurs. At the 0.1 person location, the English and Afrikaans sub-groups had an approximate 80% chance of obtaining one out of two marks. At the same location, the isiZulu learners had an approximate 40% chance of correctly responding

[28] See Appendix C for the scoring guide.

to the item – meaning two out of two marks. The isiZulu sub-group continuously outperformed the other two language sub-groups on this particular item. From the 1 person location onwards, the isiZulu sub-group had a 100% probability of correctly responding to the item. In order to obtain full marks, the learners had to find words from the text that describe the oxpecker's tail and claws. For complete comprehension, the response had to include that the tail is stiff and the claws are sharp (*cf.* Appendix C). For partial comprehension, learners only indicated one of the two answers. Based on Figure 5.13, the item discriminated against the learners tested in Afrikaans, as they found the item more difficult than the other two languages and consistently performed below the model curve.

Pearl Item 2 was the next item that showed differential functioning between the three language sub-groups (as illustrated in Figure 5.14). This item took the form of a multiple choice question. It required learners to *Make Straightforward Inferences* from the text. The item asked: Why are the children all eager to touch the pearl?

a. They want to take it away.
b. They think it is special.* (correct answer)
c. They think the boy will drop it.
d. They do not believe it is real.

Figure 5.14 shows the ICC for *Pearl Item 2*. Between the -3 and -2 person locations, the Afrikaans sub-group had less than a 10% chance of correctly answering the item. The English and isiZulu sub-group had an approximate 20% probability of correctly responding to the item. However, at the -1.5 person location, the Afrikaans sub-group had an approximate 50% chance of correctly responding to the item. The English and isiZulu sub-groups did not experience the same sharp increase as the Afrikaans sub-group at the -1.5 person location. Based on the figure below, the learners who completed the test in Afrikaans, had a higher probability across the lower class interval of correctly responding to the item. None of the languages followed a similar curve to the model curve as there is extreme inconsistency between the three languages, especially between the English and Afrikaans sub-groups and the isiZulu sub-group at the 0.4 person location. Here, the isiZulu sub-group had the lowest percentage chance of answering the item correctly across the lower and upper class intervals. This language sub-group had less than 20% chance compared to the English (approx. 85%) and Afrikaans (approx. 80%) sub-groups. What is more, at the 0.8 person location, the learners who completed the test in isiZulu

had an approximate 60% probability of correctly responding to the item and at the 1.4 person location, the same sub-group had an approximate 30% chance. At the same location, the Afrikaans sub-group had an approximate 60% and 70% chance while the English sub-group had an approximate 70% and 80% chance. The figure shows much fluctuation of the isiZulu ICC at the upper class interval. It would appear that the item discriminates against the learners who completed the test in isiZulu, although the isiZulu sub-group was the only group of learners that had a 100% chance of completing the item correctly between the 2 and 3 person locations.

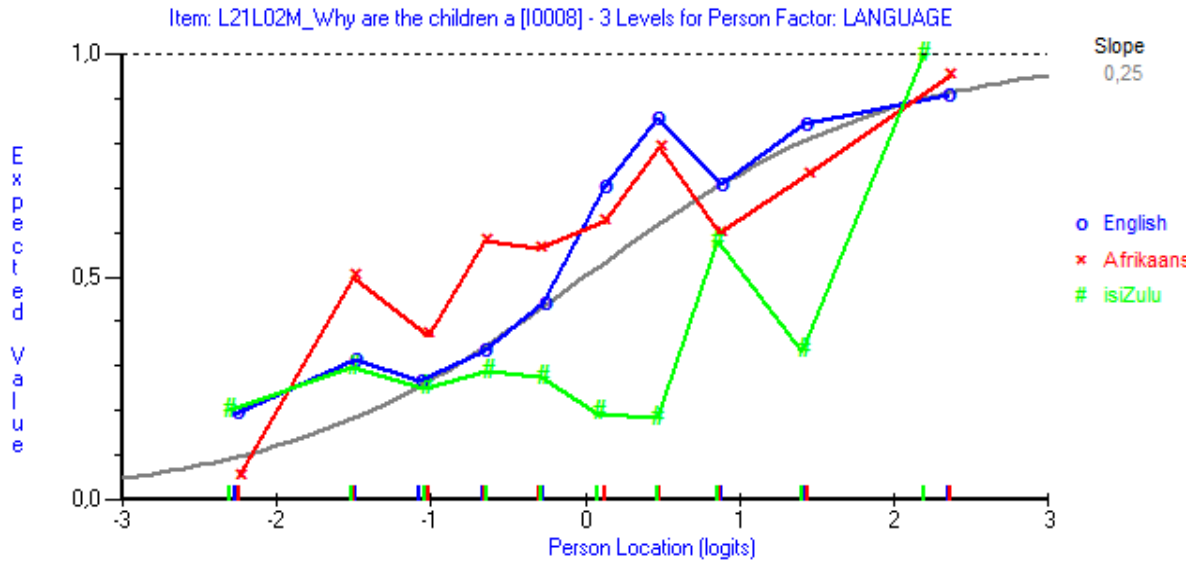

Figure 5.14: Pearl Item 2 Characteristic Curve

The next figure (5.15) depicts the distractor analysis of *Pearl Item 2*. The correct answer is distractor B(2).

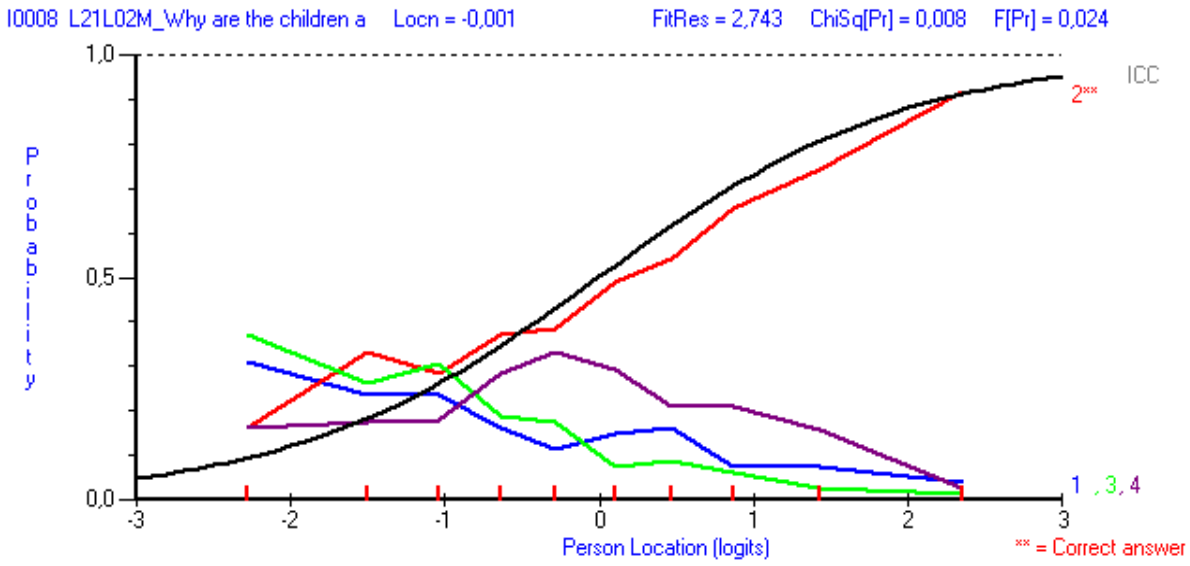

Figure 5.15: Pearl Item 2 Distractor Analysis

Even though the correct answer is distractor B(2), the learners between the -3 and -2 person locations had less than 20% chance of selecting it. It would appear that distractors C(3) and A(1) were too tempting for these learners. The learners were required to make an inference from the text that the pearl was special. In the text, one of the children said that is the pearl was beautiful and in the next paragraph, it stated that all the children were eager to touch it because it was perfect and gleaming. From these descriptions, the learners were required to infer that the pearl was special. The distractors worked well and the learners at the upper class interval were able to distinguish between the distractors for the correct answer.

Rhino Item 5 is the next item that showed DIF between the English, Afrikaans and isiZulu learners (as shown in Figure 5.16). It is a constructed response type item for one mark. It requires learners to *Focus on and Retrieve Explicitly Stated Information* by asking: What is the height of a rhino?

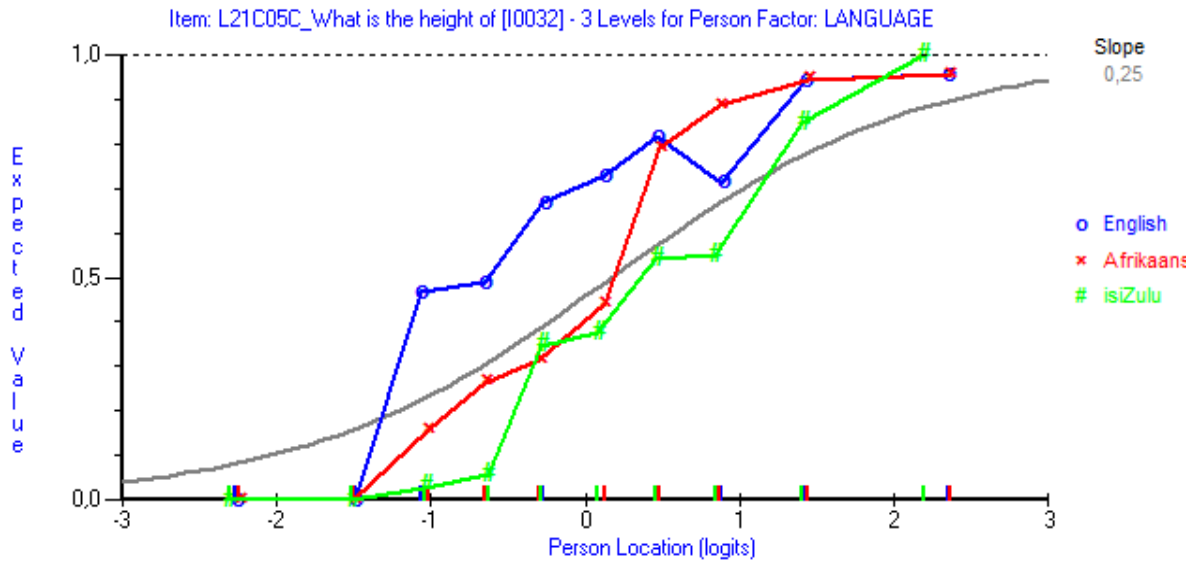

Figure 5.16: Rhino Item 5 Characteristic Curve

Figure 5.16 shows the DIF for *Rhino Item 5*. From the -2.3 to -1.4 person location, the three language sub-groups had zero percent probability of correctly responding to the item. It would appear that the item was too difficult for these learners, as it required learners to retrieve information from a table. At the -1 and -0.6 person locations, there appears to be significant inconsistency between the three languages. The isiZulu learners had the lowest chance of answering the item correctly at the -1 location, with less than 10% whereas the English and Afrikaans sub-groups had approximately a 50% and 30% chance, respectively. The Afrikaans sub-group were above the model curve from the 0.4 person location onwards, the English sub-group were more often than not above the model curve, the exception being at -2.2. Furthermore, it would appear that the English sub-group found the item less difficult than the Afrikaans and isiZulu sub-groups. At the 0.8 person location, the English sub-group had approximately a 70% probability of correctly responding to the item, the Afrikaans and isiZulu sub-groups had an approximate 90% and 50% chance, respectively. At this person location, the Afrikaans sub-group found the item less difficult. In order to obtain the one mark, learners had to write '2 metres'. If, however, the learners only wrote '2', the response would be incorrect as the unit of measurement is absent. Overall, the item displayed discrimination towards the Afrikaans and the isiZulu sub-groups.

The next item that displayed differential functioning between the three language groups is *Pearl Item 5* (as illustrated in Figure 5.17). This item is a constructed response type question

worth two marks. It entailed learners to *Make Straightforward Inferences*. The question is as follows: What does Reuben do differently after he gets the pearl?

Figure 5.17: Pearl Item 5 Characteristic Curve

In order to obtain full marks, learners had to give two responses that list the things Reuben did differently. The item required the learner to make straightforward inferences from the text. For example, Reuben no longer plays with his friends, he reads about pearls, he asks for more pearls or that Reuben wants to become a pearl merchant (*cf.* Appendix C for scoring guide). If the learner only wrote one thing that Reuben did differently, then only one mark would have been awarded. All three language sub-groups between the -3 and -2 person locations found the item very difficult and had less than a 30% chance of obtaining a partial mark – one out of two marks. The grouping of the three languages were close to each other at the -1 person location. At the 0 person location, the Afrikaans learners had approximately a 20% chance of correctly responding to the item and obtaining full marks. At the same person location, the English sub-group had approximately an 80% chance of partially answering the item correctly – one out of two marks. The learners who wrote the test in English and Afrikaans followed a similar pattern between the -1 and 1 person locations. The isiZulu sub-group at 2.2 had a 100% probability of correctly responding to the item and obtaining full marks. Overall, the Afrikaans sub-group were above the model curve, and as such, this group of learners found the item less difficult than the other two language sub-groups. It would appear that the item discriminated against the English and isiZulu sub-groups, specifically between the -1 and 1 person locations.

Rhino Item 3 showed differential functioning between the three language sub-groups (as shown in Figure 5.18). The item tested learners' ability to *Evaluate and Critique Content and Textual Elements*. It asked the following: Why does the writer tell you about the elephant?

a. To show that elephants live near rhinos
b. To show that the rhino is very big* (correct answer)
c. To show that elephants have oxpeckers
d. To show that rhinos and elephants eat the same food

Figure 5.18 graphically presents the item characteristic curve for *Rhino Item 3* as it displays differential functioning across the three language sub-groups. This item has extreme inconsistency across both lower and upper class interval among the three languages. None of the languages follow the model curve. At the -2.2 person location, the Afrikaans sub-group had zero percent probability of correctly responding to the item, whereas the English and isiZulu sub-groups had less than a 20% chance. Yet, at the -1.4 person location, the Afrikaans sub-group had a sharp increase in their chance, as they had an approximate 50% chance of correctly answering the item. The isiZulu sub-group also experienced an increase in their chance to correctly answer the item but the English sub-group experienced a small decrease. At the -1 person location, both the English and isiZulu sub-groups had an approximate 40% probability of correctly responding to the item and the Afrikaans sub-group had a slightly higher chance. Between the -1 and 0 person locations, both the Afrikaans and isiZulu sub-groups experienced a steady decline in their probability to correctly respond to the item, while the English sub-group experienced the opposite and at the -0.2 person location, had a 40% chance of correctly answering the item. Similar inconsistency is present at the upper class interval where at the 0.4 person location, the English sub-group had less than a 10% chance of correctly responding to the item. The learners who completed the test in Afrikaans had approximately a 20% chance with the isiZulu having a much higher chance (approx. 50%) than either the Afrikaans or English sub-groups. The English sub-group across the upper class interval fell below the model curve. It would appear that this sub-group experienced the item as somewhat more difficult than the other two languages. The Afrikaans sub-group at the 2.2 person location, had approximately a 90% probability of correctly responding to the item and were above the model curve. Based on the above figure, it would appear that the learners at the lower class interval had more or less the same probability of getting the item right as the learners at the upper class interval. Learners at both the lower and upper class interval experienced extreme inconsistency

among the English, Afrikaans and isiZulu sub-groups. The figure does not show any discernible pattern; furthermore, it deviates from the model curve.

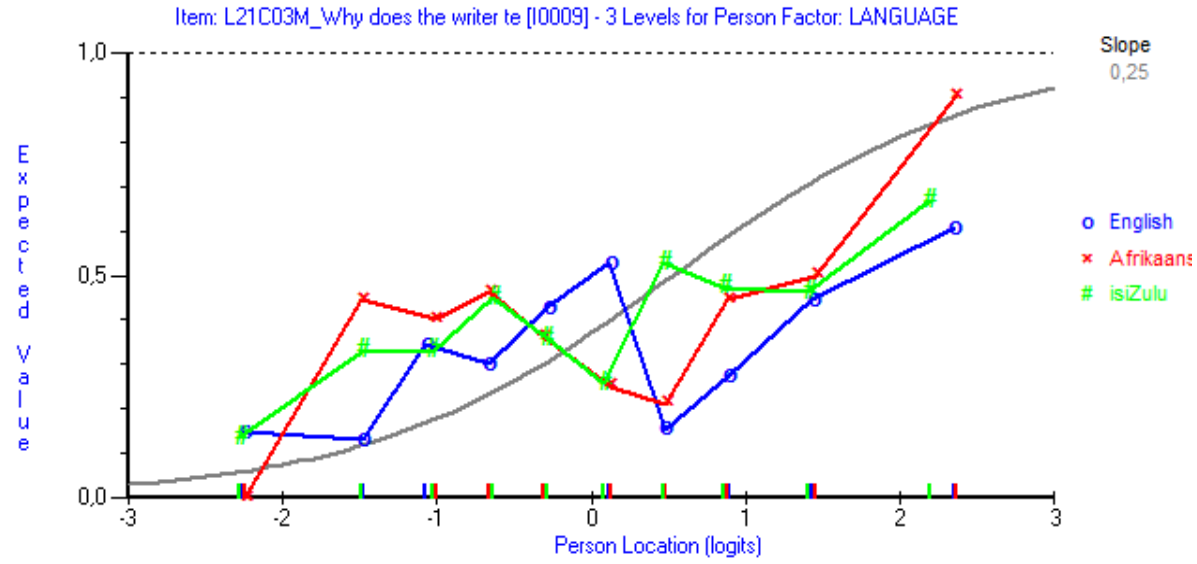

Figure 5.18: Rhino Item 3 Characteristic Curve

The next figure depicts the distractor analysis for *Rhino Item 3*. The correct answer is distractor B(2).

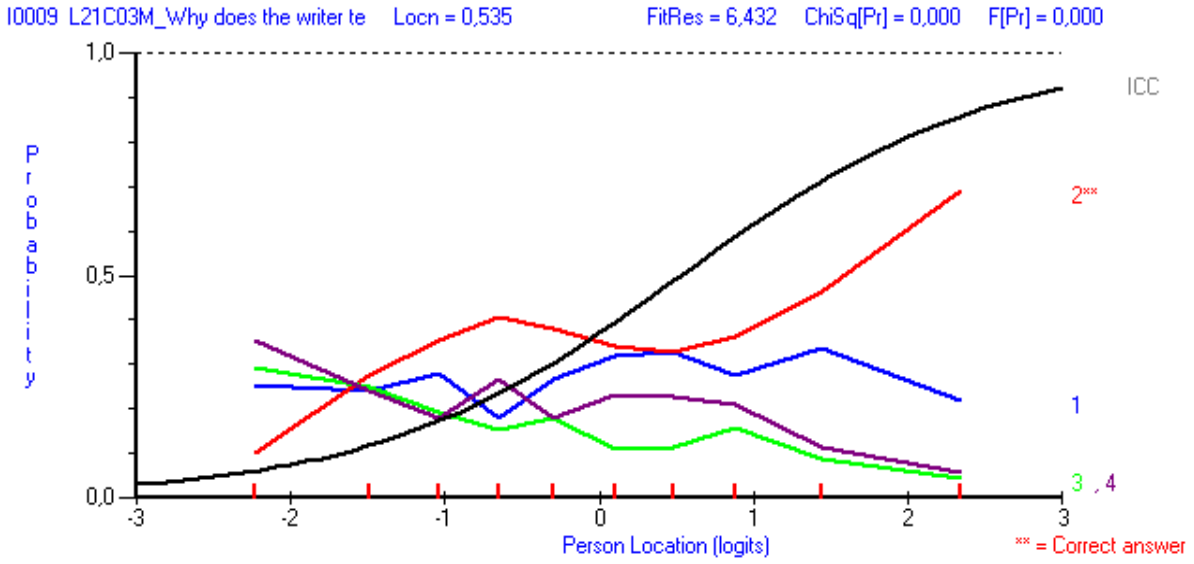

Figure 5.19: Rhino Item 3 Distractor Analysis

Figure 5.19 shows that the learners across the lower and upper class intervals had inconsistent curves across the different distractors. What is more, the chi-square is significant. The significance indicates that the distribution of distractors significantly differs from the expected

model. This finding might be due to the learners between the -3 and -2 person location having approximately a 10% probability of selecting the correct answer, distractor B(2). The learners had approximately a 35% chance of selecting distractor D(4), followed by distractors C(3) and A(1). At the -0.6 person location, the learners seem to have been able to better distinguish between the plausible answers for the correct answer; however, between the 0 and 1 person locations, the learners appeared to be tempted by distractor A(1). At the 2.3 person location, the learners had an approximate 70% chance of selecting the correct answer and an approximate 20% chance of selecting distractor A(1). The text mentioned that the rhino is a very large land animal and that only the elephant is larger (*cf*. Appendix C). From this sentence, the learners were required to evaluate and critique the content to answer the question. Based on the ICC graph as well as the distractor analysis graph, it would appear that the learners struggled with this item. It seems a higher cognitive level was required by these learners to answer this item.

Figure 5.20 shows the differential functioning on *Rhino Item 16*. This item took the form of a constructive response type question. It required learners to *Focus on and Retrieve Explicitly Stated Information* by asking the following: What does the oxpecker do to warn the rhino of danger? The question had a mark allocation of one.

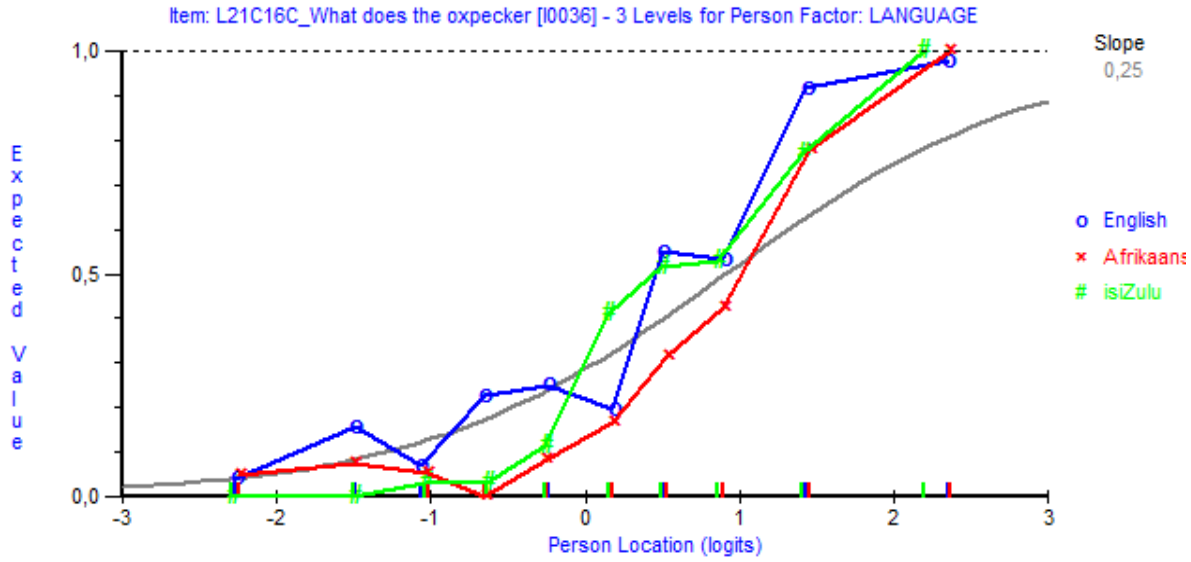

Figure 5.20: Rhino Item 16 Characteristic Curve

Figure 5.20 illustrates the differential functioning of the item across the English, Afrikaans and isiZulu sub-groups. At the -2.2 person location, all three languages had less than 10%

probability of correctly answering the item. The isiZulu sub-group at -1.5 person location had a zero percent chance of answering the item correctly. The Afrikaans sub-group remained the same and the English sub-group had a slight increase in their probability (approx. 15%). The item seemed somewhat difficult for the learners across the lower class interval as none of the language sub-groups had above a 30% chance of correctly answering the item. However, the learners who completed the test in English found the item less difficult than the Afrikaans and isiZulu sub-groups. At the 0.1 person location, both the English and Afrikaans sub-groups had an approximate 20% chance of answering the item correctly, whereas the isiZulu sub-group had a 40% chance. Then, at the 0.5 person location, the Afrikaans sub-group had the lowest probability (approx. 30%) while both English and isiZulu sub-groups had an approximate 50% probability of correctly responding to the item. At the 1.4 person location, all three language sub-groups were above the model curve. At this location, the Afrikaans and isiZulu sub-groups had both an approximate 80% chance of correctly answering the item and the English sub-group had a 90% chance. Between the 2 and 3 person location, the Afrikaans and isiZulu had a 100% probability of correctly responding to the item followed closely by the English sub-group. An acceptable response, for the one mark, would include that the oxpecker bird makes a loud noise. At the lower class interval, the English learners seemed to have found the item less difficult and at the upper class interval, the Afrikaans learners found the item more difficult. It would appear that the item discriminated against the Afrikaans sub-group.

The subsequent section examines the differential item functioning (DIF) of the *Flowers on the Roof* text PIRLS Literacy 2016 text for the learners who wrote the assessment in English, Afrikaans and isiZulu.

5.3.4 Reliability of PIRLS Literacy 2016 Flowers Text

This section provides the summary statistics for the *Flowers on the Roof* text as part of the PIRLS Literacy 2016 assessment. The analysis file *Flowers* was created in order to explore its specific items for differential item functioning (DIF). However, before conducting DIF analysis, the model fit statistics of the *Flowers on the Roof* text were examined. The *Flowers* file only included items from the *Flowers on the Roof* text. It also included the English, Afrikaans and isiZulu learner scores. Table 5.13 presents the power of analysis as well as the person reliability indices for the English, Afrikaans and isiZulu combined analysis. It also shows the information per language.

Table 5.13: Reliability of PIRLS Literacy 2016 Flowers Text

Analysis Name	Power of Analysis of Fit	PerSepIdx with extremes	PerSepIdx no extremes	Coefficient Alpha with extremes	Coefficient Alpha no extremes
Flowers [a]	Reasonable	0.53986	0.53596	N/A	N/A
English Only	Good	0.68384	0.67271	0.76884	0.75525
Afrikaans Only	Good	0.66821	0.65360	0.77425	0.75263
isiZulu Only	Too Low	0.30891	0.24412	0.53673	0.48474

[a] The analysis file *Flowers* included all languages.

The statistics in the above table for the analysis file *Flowers* (combined languages) excludes the Cronbach Alpha (α) as there were missing data present. As such, for each of the languages, the summary statistics were conducted without the missing data. Both the English and Afrikaans language sub-group test can adequately differentiate between different groups of persons (Combrinck, 2019). The α coefficient for both English and Afrikaans sub-groups were above .7 while the isiZulu sub-group had a low α coefficient of .5, which is not acceptable as it is below the acceptable threshold. It also shows a low separation index (Combrinck, 2019). For the isiZulu language sub-group, the low separation index means that the scale is not able to adequately differentiate between groups of persons.

The next section looks at the individual item-fit statistics for the *Flowers on the Roof* text for the South African Grade 4 English, Afrikaans and isiZulu learners.

5.3.5 Individual Item Statistics for Flowers Text

This section provides the individual item-fit statistics of the *Flowers on the Roof* text per item in order of difficulty. Table 5.14 provides evidence whether the items and the persons link to the fit of the model. As mentioned previously (*cf.* 5.3.2), the assumption of the model declares that as the person ability increases, so should the chance of correctly answering the more difficult items (Combrinck, 2019). When there is a lack of fit, it causes a violation of the assumptions. In addition, the chi-square value is provided in order to reveal whether there is invariance across the trait (Pallant & Tennant, 2007). This value defines the goodness-of-fit of the item-trait interaction.

Table 5.14 presents the 13 *Flowers on the Roof* items across the three language sub-groups. The item-fit analysis was conducted to determine whether the null hypothesis of this study could be accepted. The null hypothesis declares that the mean score of the learners, across English, Afrikaans and isiZulu, are equal (μEnglish = μAfrikaans = μisiZulu). If the null hypothesis is rejected, then an alternative hypothesis could be accepted. The alternative hypothesis states that the English mean score is not equal to the Afrikaans nor the isiZulu mean scores (H_a = μEnglish ≠ μAfrikaans ≠ μisiZulu).

Table 5.14: Individual Item-Fit Statistics for Flowers Text Combined Languages

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 2	-1,236	0,083	2,442	11,90	0,219
Flowers Item 4	-0,975	0,080	-3,456*	37,66	0,000**
Flowers Item 3	-0,837	0,082	-1,676	19,41	0,022
Flowers Item 13	-0,547	0,086	2,398	7,79	0,454
Flowers Item 1	-0,338	0,083	-2,742*	28,88	0,001**
Flowers Item 10	-0,260	0,087	1,780	33,17	0,000**
Flowers Item 5	-0,220	0,084	5,727*	52,72	0,000**
Flowers Item 11	-0,128	0,087	2,315	15,69	0,074
Flowers Item 9	0,260	0,063	-4,085*	33,41	0,000**
Flowers Item 7	0,909	0,076	-0,234	14,55	0,104
Flowers Item 12	0,964	0,079	-1,024	11,37	0,251
Flowers Item 6	0,989	0,107	-2,744*	18,05	0,035
Flowers Item 8	1,419	0,120	-3,530*	42,52	0,000**

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.001282 for 13 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

A total of 13 items accompanied the *Flowers on the Roof* text, of which five displayed a misfit that was significant. The five items are categorised by their misfit, namely overfit and underfit. *Flowers Item 5* was the only overfit item, meaning that the item discriminates a great deal between learners. Items 4, 1, 9 and 8 displayed underfit, which means that the items do not adequately discriminate between the less able and more able learners (van Staden, 2018). This finding is valuable; however, to be able to accept or reject the null hypothesis, this study

examined the languages' individual item-fit statistics. Tables 5.15, 5.16 and 5.17 present the item-fit statistics for the English, Afrikaans and isiZulu language sub-groups.

Table 5.15: Individual Item-Fit Statistics for Flowers Text English Only

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 2	-1,384	0,131	0,929	5,46	0,707
Flowers Item 4	-1,108	0,126	-2,598*	22,13	0,005
Flowers Item 3	-0,850	0,127	-1,957	14,28	0,075
Flowers Item 1	-0,663	0,126	-2,058	19,48	0,021
Flowers Item 13	-0,571	0,132	1,251	2,43	0,965
Flowers Item 9	-0,068	0,089	-4,148*	24,05	0,004
Flowers Item 11	-0,021	0,133	1,477	9,90	0,272
Flowers Item 10	0,078	0,136	1,935	9,87	0,274
Flowers Item 5	0,570	0,142	4,654*	60,66	0,000**
Flowers Item 12	0,704	0,105	0,064	11,75	0,163
Flowers Item 6	0,917	0,153	-1,614	11,84	0,159
Flowers Item 7	0,966	0,109	0,009	19,07	0,014
Flowers Item 8	1,430	0,174	-2,601*	23,73	0,003

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000769 for 13 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

The above table (5.15) presents the English sub-group item-fit statistics of which only one item displayed significant misfit. *Flowers Item 5* displayed overfit and over discrimination between learners with less or more ability.

Table 5.16 shows the Afrikaans sub-group item-fit statistics of the *Flowers on the Roof* text.

Table 5.16: Individual Item-Fit Statistics for Flowers Text Afrikaans Only

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 2	-1,220	0,169	2,033	7,96	0,437
Flowers Item 3	-0,682	0,165	0,329	10,25	0,331

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 5	-0,634	0,165	1,368	11,62	0,169
Flowers Item 4	-0,607	0,165	-2,668*	14,66	0,066
Flowers Item 10	-0,553	0,168	0,591	23,44	0,003**
Flowers Item 11	-0,396	0,170	0,817	6,30	0,614
Flowers Item 13	-0,191	0,176	0,546	13,21	0,105
Flowers Item 1	0,031	0,175	-0,954	9,42	0,308
Flowers Item 9	0,562	0,132	-0,545	5,74	0,677
Flowers Item 6	0,620	0,192	-0,388	8,51	0,484
Flowers Item 7	0,753	0,137	-0,111	3,60	0,936
Flowers Item 8	0,980	0,207	-1,365	12,41	0,134
Flowers Item 12	1,336	0,170	-0,810	9,12	0,244

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000769 for 13 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Based on the table above, none of the items had a misfit which was significant. Only one item displayed underfit but it was not significant, namely *Flowers Item 4*. It would appear that the Afrikaans sub-group had a better fit to the model than the English sub-group.

Table 5.17 shows the item-fit statistics for the last of the three languages, namely isiZulu.

Table 5.17: Individual Item-Fit Statistics for Flowers Text isiZulu Only

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 2	-1,171	0,142	2,493	11,09	0,269
Flowers Item 4	-1,165	0,136	-0,580	13,76	0,131
Flowers Item 5	-1,103	0,137	-0,642	11,58	0,238
Flowers Item 3	-1,031	0,140	-1,149	9,06	0,431
Flowers Item 13	-0,903	0,149	2,272	10,05	0,346
Flowers Item 10	-0,647	0,152	-0,281	12,80	0,172
Flowers Item 1	-0,247	0,153	-1,261	13,19	0,154
Flowers Item 11	-0,152	0,162	1,952	17,18	0,046

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 9	0,636	0,150	-1,161	7,05	0,632
Flowers Item 7	0,802	0,167	-0,183	9,68	0,377
Flowers Item 12	1,162	0,200	-0,423	3,86	0,920
Flowers Item 6	1,770	0,297	-1,542	3,82	0,873
Flowers Item 8	2,047	0,329	-1,316	10,18	0,336

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000769 for 13 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

None of the 13 items of the *Flowers on the Roof* text showed a misfit nor significance. It would appear that the isiZulu sub-group had a better fit than the English and Afrikaans sub-groups. Based on Tables 5.15, 5.16 and 5.17, it would appear that the item-fit statistics varied across the three languages. As such, this finding assists in partially rejecting the null hypothesis that stated that the mean scores of the three languages are equal.

The following section examines the differential item functioning (DIF) of *Flowers on the Roof* PIRLS Literacy 2016 text for the learners who wrote the assessment in English, Afrikaans and isiZulu.

5.3.6 Differential Item Functioning for Flowers Text

The items derived from the *Flowers on the Roof* text, which displayed differential item functioning (DIF) across the English, Afrikaans and isiZulu language sub-groups are discussed in this section. As mentioned in Section 5.3.3, an ANOVA test is included in the DIF summary (as shown in Table 5.18). This summary provides evidence as to whether the mean scores of each language group are comparable. In other words, the DIF summary along with the ANOVA test provides a way of testing the null hypothesis[29] where all groups have equal means (Field, 2009; Tabachnick & Fidell, 2007). Table 5.18 presents the DIF analysis summary for the *Flowers on the Roof* text.

[29] The ANOVA also provides the F-ratio and p-value. The former measures two or more quantities that are expected to be the same under the null hypothesis. The latter refers to the statistical model that is given as evidence to whether the null hypothesis can be accepted or not.

Table 5.18: DIF Summary for PIRLS Literacy 2016 Flowers Text

Item	F-ratio	Probability
Flowers Item 2	1,278	0,279
Flowers Item 4	3,444	0,032
Flowers Item 3	0,406	0,666
Flowers Item 13	3,349	0,036
Flowers Item 1	8,644	0,000*
Flowers Item 10	6,661	0,001
Flowers Item 5	30,393	0,000*
Flowers Item 11	2,267	0,104
Flowers Item 9	19,500	0,000*
Flowers Item 7	2,597	0,075
Flowers Item 12	8,305	0,000*
Flowers Item 6	7,270	0,001*
Flowers Item 8	6,621	0,001

*Significant at the 5 percent level (Bonferroni 0.001282)

A summary of the DIF for the *Flowers on the Roof* text, as demonstrated by ANOVA statistics, is provided in the above table. The *p*-value is significant if the *p*-value is smaller than <0.05. Uniform DIF was used for the purposes of this study as it looks at the persons who have the same ability that consistently have a differing probability of correctly responding to an item when compared to other groups (Andrich et al., 2012). The *Flowers on the Roof* text has 13 accompanying items, of which five displayed differential functioning across the English, Afrikaans and isiZulu sub-groups, namely items 1, 5, 9, 12 and 6. The rest of this section focuses on these items by looking at their item characteristic curves (ICC). Where the item is a multiple choice item, the distractor analysis graphs are also presented.

Flowers Item 1 is a multiple choice question that requires learners to *Interpret and Integrate Ideas and Information* by asking: Who is telling the story?

a. a granny
b. a child* (correct answer)
c. a doctor
d. a farmer

Figure 5.21 presents the item characteristic curve for *Flowers Item 1*. This item displayed DIF across the three language sub-groups and has extreme inconsistency across the lower class interval (between -3 and 0). Between the -3 and -2 person locations, all three language sub-groups have less than 10% probability of correctly responding to the item. Extreme inconsistency occurs between the -2 and -1 interval. At the -1.7 person location, the learners who completed the test in English and Afrikaans were above the model curve having an approximate 20% chance of answering the item correctly. The isiZulu sub-group remained below the model curve with an approximate 10% chance. The Afrikaans sub-group at -1.6 decreased to zero percent chance of correctly responding to the item whereas the English and isiZulu sub-groups had an approximate 30% and 10% chance respectively. Then at the -1.2 person location, the isiZulu sub-group had an approximate 30% chance and was above the model curve but both the English and Afrikaans sub-groups were below the model curve. At around the -1 person location, all three languages fell below the model curve. The Afrikaans sub-group at -0.6 remained below the model curve and had an approximate 40% probability of correctly responding to the item, but the same sub-group at -0.1 person location had less than a 30% chance. The English and isiZulu sub-groups at -0.2 person location had an approximate 70% and 50% chance, respectively, to answer the item correctly. However, the learners across the upper class interval were above the model curve with the English sub-group having the highest percent chance at the 0.3 person location. The isiZulu sub-group at 0.8 had a 100% probability of correctly responding to the item. Overall, it would appear that item discriminated against learners at the lower class interval as well as the Afrikaans sub-group at the upper class interval.

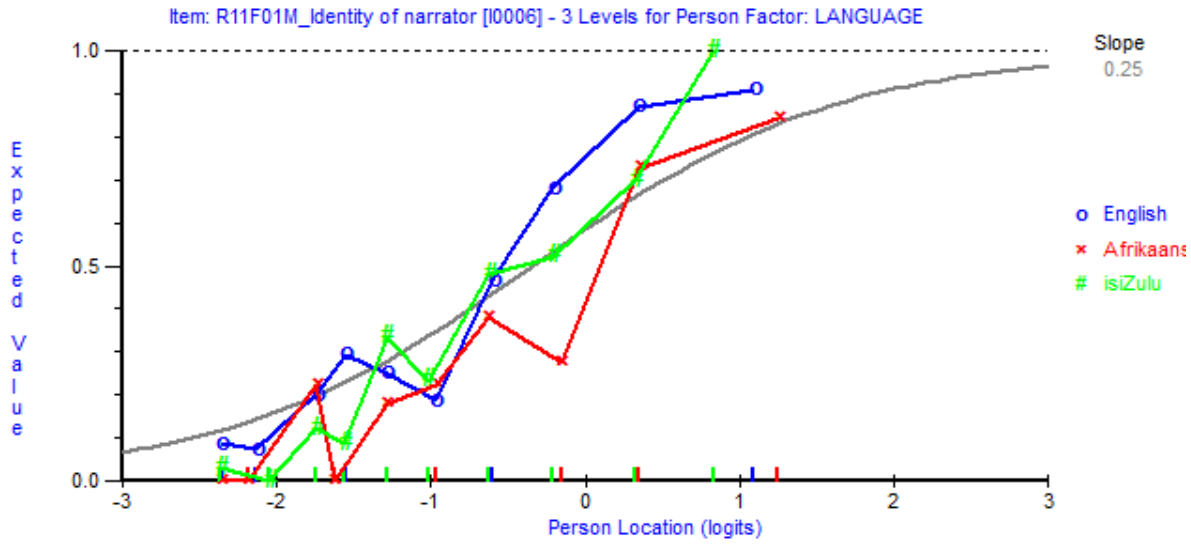

Figure 5.21: Flowers Item 1 Characteristic Curve

The next figure depicts the distractor analysis of *Flowers Item 1*. The correct answer is distractor B(2).

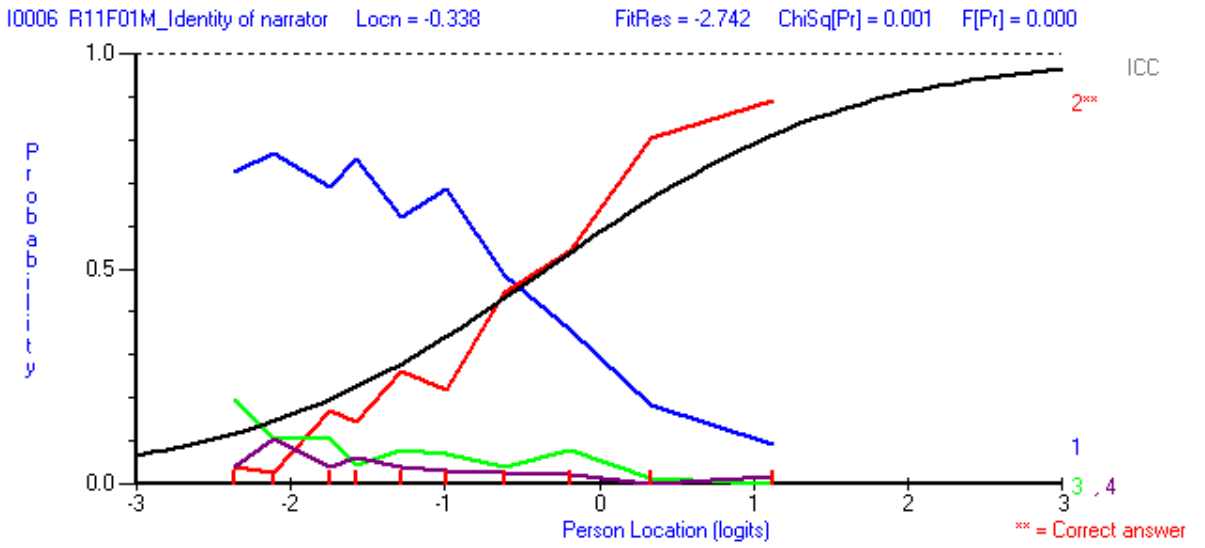

Figure 5.22: Flowers Item 1 Distractor Analysis

The question asked who told the story and learners had four plausible distractors to choose from. It would appear that the learners across the lower class interval (between -3 and 0) were much more likely to select distractor A(1) rather than the correct answer. At the -2.1 person location, the learners had an approximate 80% chance of selecting distractor A(1). The chi-square of this item is significant, indicating that the distribution of distractors differs significantly from the expected model. Learners across the lower class interval may have

thought the granny told the story as it focuses on certain things that she does, but she it is not the narrator. The learners had to make the inference from the text and it would appear that they struggled with the text. Neither distractor C(3) nor D(4) drew sufficient attention from the learners. As these two distractors do not draw enough attention away from the correct distractor, it may be beneficial to rework or replace them with plausible answers that learners may consider. At the -0.6 person location, the learners were able to distinguish between the more plausible distractor, and selecting distractor B(2) as the correct answer.

The next item, *Flowers Item 5*, displayed differential functioning across the three language sub-groups (as illustrated in Figure 5.23). It asked learners to *Make Straightforward Inferences* from the text. The item asked the following: Granny Gunn did not like the walls and windows in her new flat. Why else was she unhappy?

a. She was ill.
b. She missed her cat.
c. She did not like the balcony.
d. She felt homesick.* (correct answer)

Flowers Item 5 showed extreme inconsistency across the three language groups as seen in the figure below. It would appear that the learners who completed the test in English found the item considerably more difficult. Moreover, the English sub-group did not follow the expected model curve. At the -2.3 person location, both the English and Afrikaans sub-group had an approximate 30% probability of correctly responding to the item, whereas the isiZulu sub-group had less than 10% chance. The learners who wrote the test in isiZulu experienced a gradual increase across the lower and upper class interval, with a deviation at -1.2 person location. The isiZulu sub-group found this item less difficult than the English and Afrikaans sub-groups. At the -1.7 person location, the Afrikaans sub-group had zero percent probability of correctly responding to the item, while the English and isiZulu sub-groups had approximately 10% and 20% chance, respectively. At the next person location (-1.6), the Afrikaans sub-group had again an approximate 30% chance of correctly answering the item. Both the Afrikaans and isiZulu sub-groups remained above the expected model curve from -1.6 onwards. The English sub-group remained relatively level across the lower and upper class intervals. At the -0.6 person location, the English sub-group had less than 20% chance to correctly respond to the item, whereas the Afrikaans and isiZulu sub-groups had an approximate 35% and 55% chance of answering the item correctly. At the upper class interval,

specifically at 0.3 person location, the English sub-group had an approximate 20% probability of correctly responding to the item. Both the Afrikaans and isiZulu sub-groups were above the expected model curve and had an approximate 60% and 90% chance of answering the item correctly. The percentage gap between the three languages increased at the last person location for each language. Based on the above findings, it would appear that the item discriminated against the English sub-group. These learners struggled with item 5 and had a lower chance of correctly answering it.

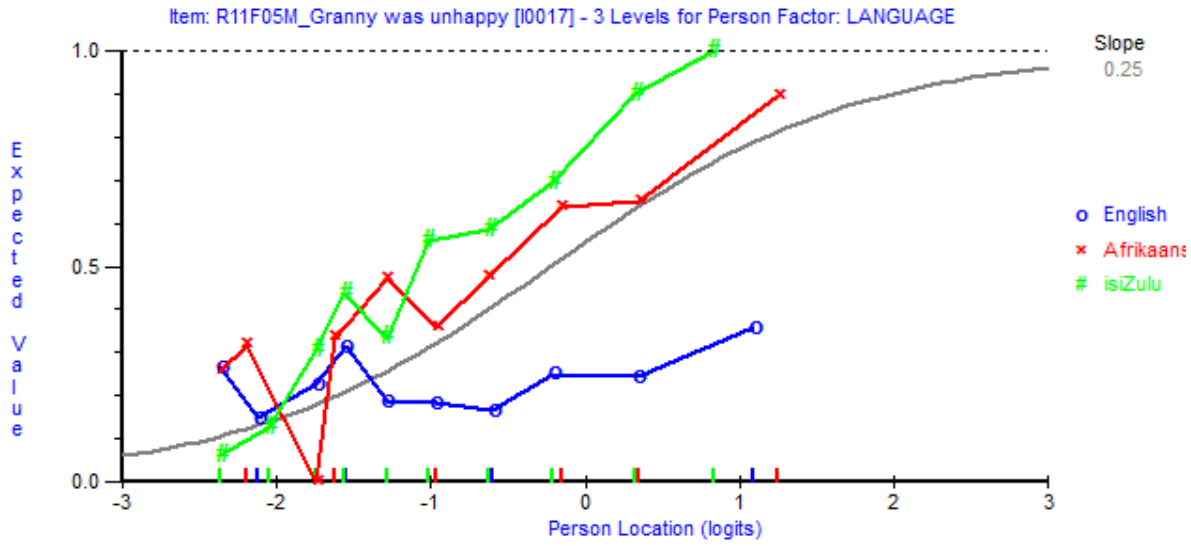

Figure 5.23: Flowers Item 5 Characteristic Curve

The next figure shows the distractor analysis of *Flowers Item 5*. Distractor D(4) was the correct answer.

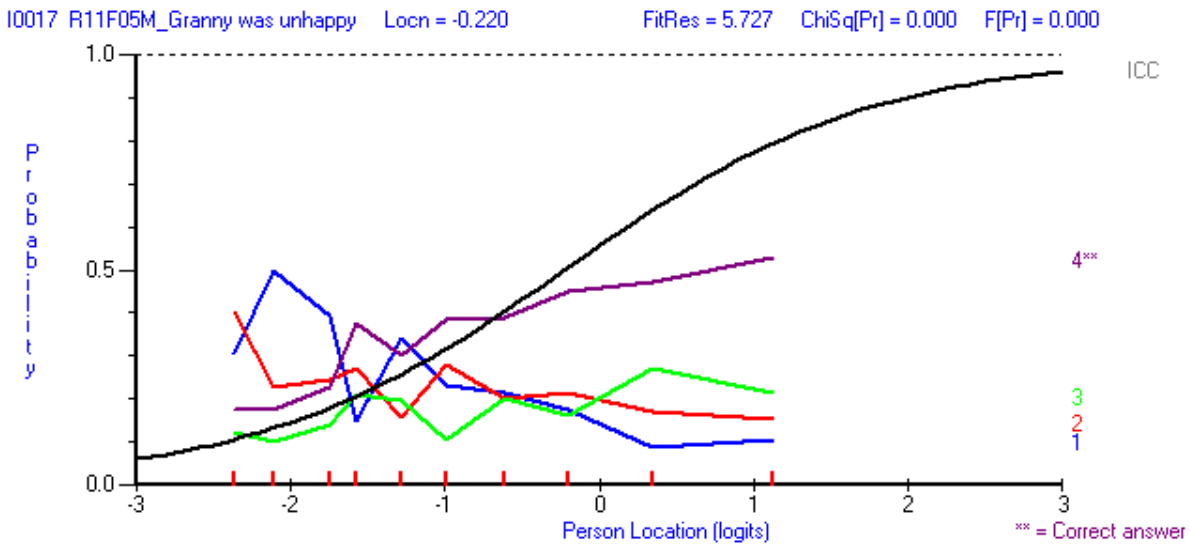

Figure 5.24: Flowers Item 5 Distractor Analysis

Figure 5.24 shows the distractor analysis of *Flowers Item 5*. Although the correct answer is distractor D(4), the learners across the lower class interval were tempted by distractors A(1), B(2) and C(3). The question asked, besides the walls and windows of her new flat, why else Granny Gunn was unhappy. Learners at the -2.3 person location had a 40% probability of selecting distractor B(2) and at the -2.1 person location, they had a 50% chance of selecting distractor A(1). The former mentioned that Granny Gunn missed her cat while the latter stated that she was ill. Both these distractors were plausible for the learners as they inferred that she had missed her cat and it was stated that she was ill and therefore had to move to the city. However, both these answers are not correct, as they were not the reason for her being unhappy. Learners had to infer from the text that because she missed her previous home and her animals, it made her feel homesick. The word '*homesick*' is more opaque than the Afrikaans '*terugverlang plaastoe*' (missed the farm) and isiZulu '*wayekhumbule ekhaya*' (she missed home). At the -0.9 person location, the learners had approximately a 40% chance of selecting distractor D(4) and at the 1.1 person location, they had approximately a 50% chance of selecting it. At the 1.1 person location, the learners had an approximate 20% chance of selecting distractor C(3). Furthermore, the chi-square is significant, signifying that the distribution of distractors is significantly different from the expected model. Learners had difficulty identifying the correct response from the four plausible answers.

The next item that displayed differential functioning across the language sub-groups was *Flowers Item 9* (as illustrated in Figure 5.25). It took the form of a constructed type item for two marks. It required learners to *Focus on and Retrieve Explicitly Stated Information* from

the text. The question asked the following: Write two ways in which Granny Gunn made her new flat feel like home.

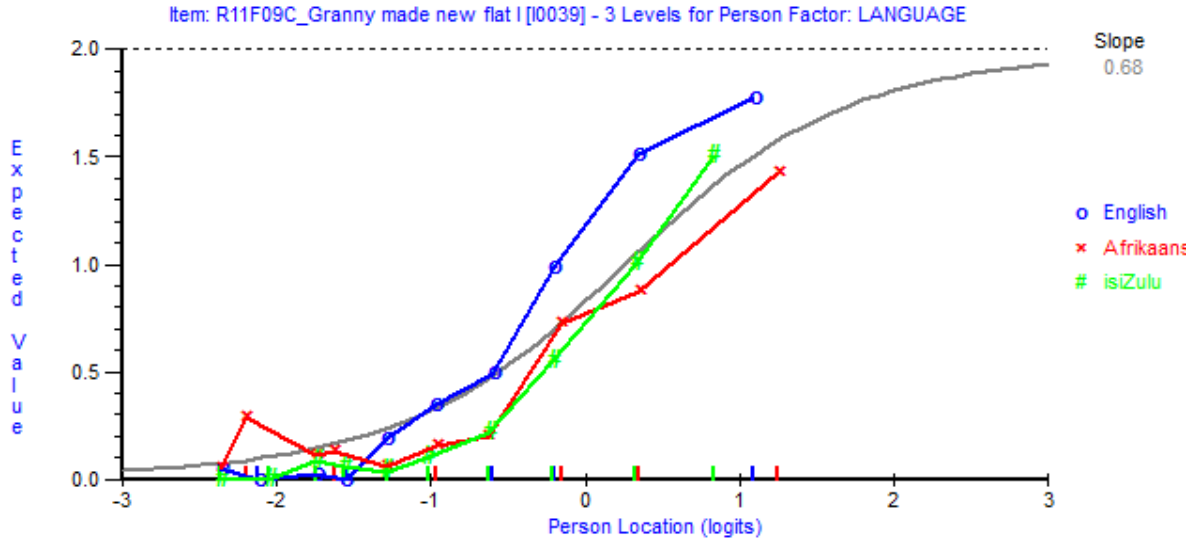

Figure 5.25: Flowers Item 9 Characteristic Curve

In order to obtain two marks, the learner responses should indicate two actions of how Granny Gunn made her new flat feel like home (*cf.* Appendix C). For example, she planted grass/flowers on the roof and brought her cat to her new flat. Another example may include that Granny Gunn brought her hens to her new flat as well as her cat. For one mark, the response would only include one of the above answers. The above figure shows inconsistency across the language sub-groups across the lower and upper class intervals. The Afrikaans and isiZulu sub-groups were more often than not below the expected model curve. The learners across the lower class interval, specifically between -3 and -1, found this item extremely difficult. At the -2.2 person location, the learners who wrote the test in Afrikaans had an approximate 30% chance of obtaining one out of two marks. Only 23% of the learners who completed the test in English obtained full marks for the item, while only 10% and 3% of the Afrikaans and isiZulu sub-groups, respectively, achieved full marks[30]. At the -0.1 person location, the Afrikaans sub-group had an approximate 70% chance of partially getting the item right. At the -0.2 person location, the English and isiZulu sub-groups had an approximate 90% and 60% chance achieving one out of two marks. Learners across the upper class interval seem to have also struggled with the item as the Afrikaans sub-group at the 0.3 person location had less than 90%

[30] *cf.* Table 5.6 on page 129.

chance of obtaining a partial mark while the English sub-group had an approximate 50% probability of obtaining full marks. Based on the ICC graph, it would appear that the item discriminates against the Afrikaans sub-group as a higher cognitive level was required by the Afrikaans learners to answer this item.

The next item, *Flowers Item 12*, displayed differential functioning across the English, Afrikaans and isiZulu learners. It is a constructed type item worth three marks. It required the learners to *Interpret and Integrate Ideas and Information* by asking: What were the little boy's feelings about Granny Gunn when she first moved in and at the end of the story? Use what you have read to describe each feeling and explain why his feelings changed.

Figure 5.26 shows the ICC graph for the *Flowers Item 12*.

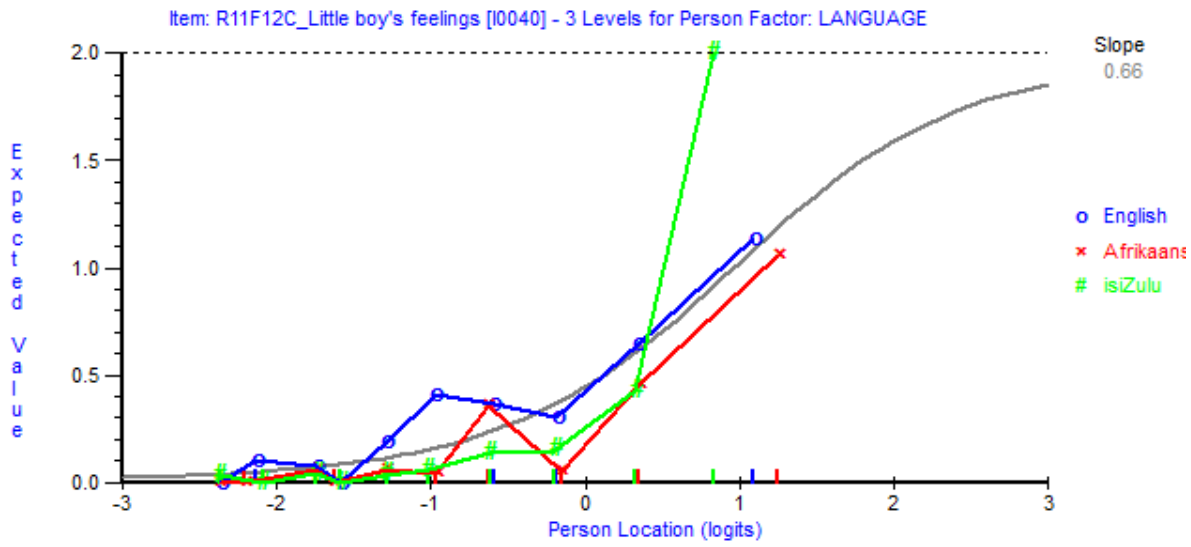

Figure 5.26: Flowers Item 12 Characteristic Curve

Even though the item carries three marks, none of the learners was able to obtain full marks as shown in Table 5.6.[31] As such, the model discarded the three-mark parameter; using only the two-mark parameter. Figure 5.26 shows that *Flowers Item 12* was exceptionally difficult for all learners, particularly for the learners at the lower class interval (between -3 and 0). Little variation occurs -2.3 and -1.5. Learners between the aforementioned person locations had less than 10% chance of partially answering the item correctly – in other words, achieving one

[31] *cf*. Table 5.6 on page 129.

mark. Across the lower class interval, the English sub-group were more often above the expected model curve. The learners who completed the test in either Afrikaans or isiZulu had almost zero percent chance of correctly answering the item, while at the -0.6 person location, the Afrikaans sub-group suddenly experienced an increase in their probability of correctly answering the item. The same sub-group at the -0.1 person location had again almost zero percent chance. The Afrikaans and isiZulu sub-group at the 0.3 person location, had 40% probability of partially correctly responding to the item compared to the English sub-group who had an approximate 60% chance. Peculiarly, the isiZulu sub-group at the 0.8 person location had a 100% chance of correctly responding to the item, in other words, obtaining two marks. It would appear that the item discriminated against the Afrikaans sub-group. These learners found the item more difficult than the other two language sub-groups; the Afrikaans sub-group also remained below the expected model curve. Appendix C shows the extensive scoring guide for this item. Even though the item originally counted out of three marks, this section will only look at the two and one mark scoring guide. For two marks, or satisfactory comprehension, the learners had to provide a response that showed their understanding of the little boy's feelings about Granny Gunn and why his feelings had changed. In order to obtain the two marks, the learners had to incorporate a description of the negative feelings the little boy experienced when the granny first moved into her new flat. Then the learner had to explain or describe that the little boy's feelings had undergone a positive change without providing a reason why the feelings had changed. For limited comprehension, the learners had to demonstrate some understanding of the little boy's feelings by giving either positive or negative feelings regarding the granny moving in next door. As this item was an interpret and integrate type item, it would appear that the learners, especially at the lower class interval, struggled to make the required inferences regarding the little boy's feelings.

The last item that displayed DIF among the three sub-groups is *Flowers Item 6* (as illustrated in Figure 5.27). This item is also a constructed response type question and required learners to *Make Straightforward Inferences*. The question is as follows: Why did Granny Gunn scream when the cat jumped out of the window? This item has a mark allocation of one.

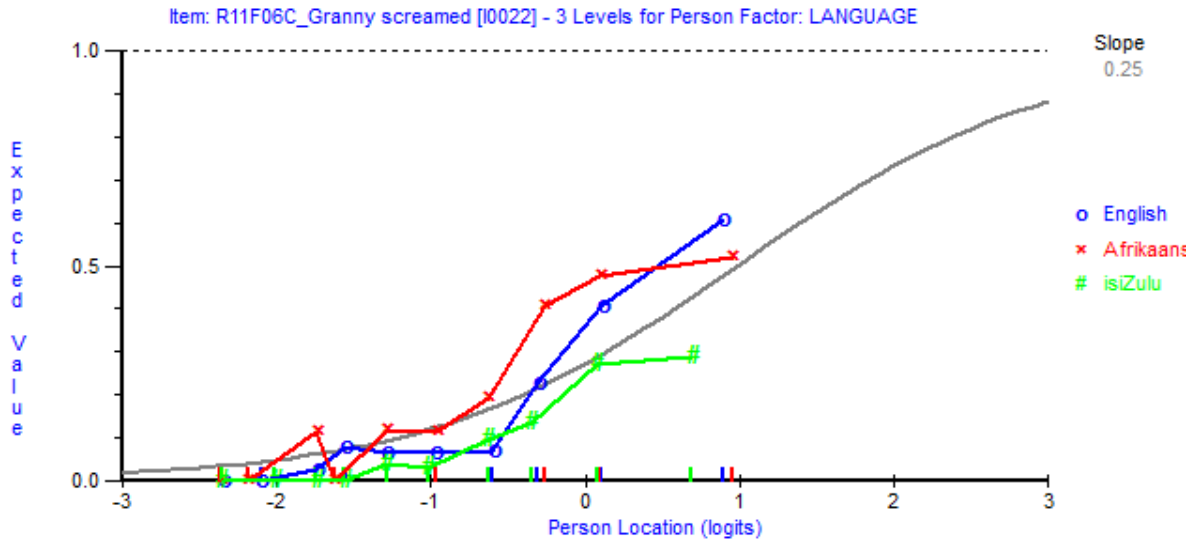

Figure 5.27: Flowers Item 6 Characteristic Curve

Figure 5.27 shows that the learners across the lower class interval found the item extremely difficult. The learners who wrote the test in isiZulu and who were between the -2.3 and -1.5 person locations had zero percent chance of correctly answering this item. The English learners between the same person locations had less than 10% chance. The Afrikaans sub-group experienced inconsistency between the same person locations. The isiZulu sub-group remained below the expected model curve across the lower and upper class intervals. At the -0.6 person location, the English and isiZulu sub-groups had a less than 10% probability of correctly responding to the item, while the Afrikaans sub-group had an approximate 20% chance. From this person location onwards, the Afrikaans item characteristic curve remained above the model curve. At the 0.8 person location, the English sub-group had the highest percent chance of correctly responding to the item. It would seem that the item discriminates against the isiZulu sub-group.

The final section of this chapter provides a brief summary of the PIRLS Literacy 2016 items that displayed differential functioning across the English, Afrikaans and isiZulu subgroups.

5.4 CHAPTER SUMMARY

The following section provides a summary of the items that displayed different item functioning (DIF) across the English, Afrikaans and isiZulu sub-groups. Table 5.19 displays these items by text and item number and is sorted based on text and item number.

Table 5.19: Summary of Problematic Items across PIRLS Literacy 2016 Released Texts

Text Title	Item	Process of Comprehension	Item Type	F-ratio	Probability
The Pearl	Item 2	Make Straightforward Inferences	MC	13,616	0,000*
The Pearl	Item 5	Make Straightforward Inferences	CR	11,554	0,000*
The Pearl	Item 14	Focus on and Retrieve Explicitly Stated Information	MC	19,097	0,000*
African Rhinos and the Oxpecker Birds	Item 2	Evaluate and Critique Content and Textual Elements	MC	8,386	0,000*
African Rhinos and the Oxpecker Birds	Item 3	Evaluate and Critique Content and Textual Elements	MC	9,208	0,000*
African Rhinos and the Oxpecker Birds	Item 4	Focus on and Retrieve Explicitly Stated Information	CR	52,452	0,000*
African Rhinos and the Oxpecker Birds	Item 5	Focus on and Retrieve Explicitly Stated Information	CR	33,405	0,000*
African Rhinos and the Oxpecker Birds	Item 6	Make Straightforward Inferences	MC	13,524	0,000*
African Rhinos and the Oxpecker Birds	Item 10	Focus on and Retrieve Explicitly Stated Information	MC	13,704	0,000*
African Rhinos and the Oxpecker Birds	Item 12	Focus on and Retrieve Explicitly Stated Information	CR	16,706	0,000*
African Rhinos and the Oxpecker Birds	Item 13	Make Straightforward Inferences	MC	23,910	0,000*

Text Title	Item	Process of Comprehension	Item Type	F-ratio	Probability
African Rhinos and the Oxpecker Birds	Item 16	Focus on and Retrieve Explicitly Stated Information	CR	13,884	0,000*
Flowers on the Roof	Item 1	Interpret and Integrate Ideas and Information	MC	8,644	0,000**
Flowers on the Roof	Item 5	Make Straightforward Inferences	MC	30,393	0,000**
Flowers on the Roof	Item 6	Make Straightforward Inferences	CR	7,270	0,001**
Flowers on the Roof	Item 9	Focus on and Retrieve Explicitly Stated Information	CR	19,500	0,000**
Flowers on the Roof	Item 12	Interpret and Integrate Ideas and Information	CR	8,305	0,000**

Note. MC is Multiple Choice Item; while CR is Constructed Response Item.
*Significant at the 5 percent level (Bonferroni 0.000521)
**Significant at the 5 percent level (Bonferroni 0.001282)

Of the 17 items that displayed DIF, almost half (9) were multiple choice (MC) type questions while the remaining items were constructed response (CR) items.

Figure 5.28 indicates the number of DIF items in relation to the *Processes of Comprehension*. Interestingly, the majority (7) of the items were literal type items requiring learners only to access and retrieve information stated in the text, while only four items tested the learners' ability to move beyond the literal interpretation of the text by generating their own answers. These four items tested learners' ability to interpret and integrate ideas from across the text, to evaluate the information contained in the text and as such, construct meaning from what they are reading (McLeod Palane, 2017).

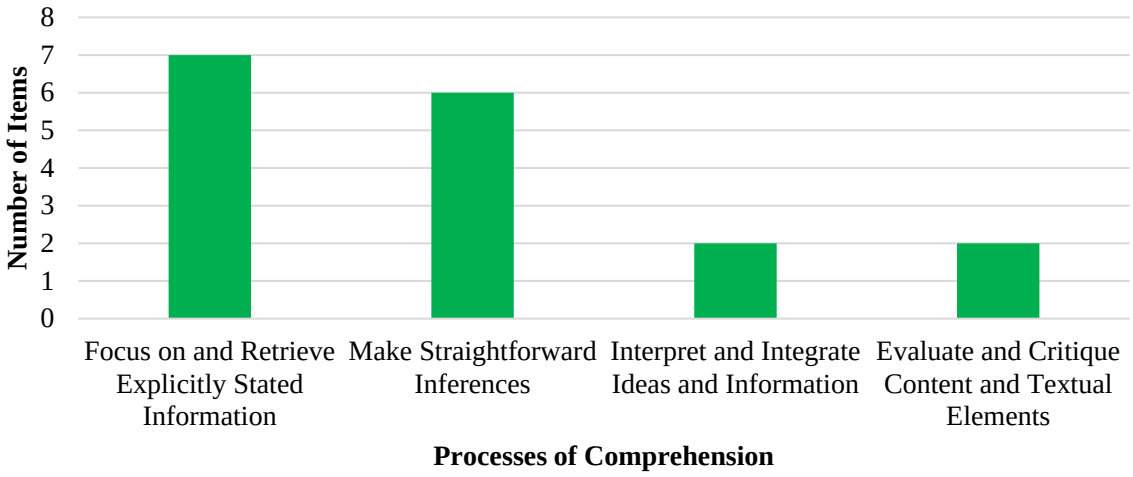

Figure 5.28: PIRLS Literacy 2016 DIF Items per Processes of Comprehension

The next table provides a summary of the number of items that displayed DIF by the total number of items per text.

Table 5.20: Summary of Problematic Items per PIRLS Literacy 2016 Released Text by Total Number of Items

Text Title	Total Number of Items	Total Number of Items Displaying DIF	Percent (%)
The Pearl	15	3	20
African Rhinos and the Oxpecker Birds	17	9	53
Flowers on the Roof	13	5	38

This chapter presented evidence for the sub-question: *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?* This question looked at the differences between the three languages using mean scores, percentages and Rasch measurement theory. Based on evidence shown in this chapter, as well as the summary tables (5.19 and 5.20), it appears that the selected PIRLS Literacy 2016 texts are not equivalent across the three languages. This finding reveals that the PIRLS Literacy 2016 texts lack metric equivalence. Yet, the items did not consistently favour one language above the other; the questions also did not discriminate against only one language. The items that displayed differential functioning showed, in some cases, extreme inconsistencies across the language groups and in some instances, between the lower and upper class intervals. The most difficult

item for the English sub-group was *Flowers Item 5* (Why else was Granny Gunn unhappy?). For the Afrikaans sub-group *Rhino Item 12* (What are these parts of an oxpecker like?) was somewhat more difficult compared to the other two language sub-groups. *Pearl Item 2* (Why are the children eager to touch the pearl?) has extreme inconsistencies for the isiZulu language sub-group. *Flowers Item 12* (What were the little boy's feelings about the granny?) was the most difficult for the learners across the lower class interval. While there may be other reasons for disparities and inconsistencies, there is insufficient evidence from DIF to indicate this.

Chapters Seven and Eight provide evidence of possible cultural and linguistic equivalence based on the evidence from the Rasch analysis as set out in this chapter.

-⌘-

CHAPTER 6

Phase One: PIRLS 2016 Grade 5 Quantitative Data Presentation and Results

6.1 ORIENTATION

This chapter presents the findings for the second sub-question of this study and follows a similar structure to Chapter Five. This phase includes the quantitative component focusing on the South African PIRLS 2016 data. Section 6.2 provides the descriptive results of the secondary analysis, which is followed by the Rasch analyses outputs and discussion (6.3). Section 6.4 provides a summary of the items that displayed differential functioning across the three languages. These items are based on the following released texts: *Macy and the Red Hen*, *The Green Sea Turtle's Journey of a Lifetime* and *Flowers on the Roof*. The goal of this chapter is to answer the second sub-question of the study: *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016*?

6.2 DESCRIPTIVE RESULTS OF SUB-QUESTION 2

This section comprises two sub-sections that provide evidence for answering sub-question 2. During the first sub-section (6.2.1) the South African Grade 5 learner results during PIRLS 2016 are provided. Next, descriptive statistical evidence for addressing sub-question 2 (6.2.2) is provided. These results are discussed in order to assist in addressing the sub-question.

6.2.1 PIRLS 2016 Results

South African Grade 5 learners participated in the PIRLS 2016 study. However, it should be noted that three languages were selected to be part of PIRLS, the results are only for benchmarking purposes and do not represent all Grade 5 learners in South Africa. The sample was, nevertheless, representative of English, Afrikaans and isiZulu by language and province (Howie et al., 2017b). These learners were a year older than the target population defined by the IEA (*cf.* Chapter Two). The reason for selecting these three languages and the age-group was based on the results from the previous rounds of PIRLS in 2006 and again in 2011 (*cf.* Combrinck et al., 2017, p. 32). The 2016 sample included 5 285 Grade 5 learners in 125 schools across the nine provinces (Howie et al., 2017b). This sub-sample was further reduced since this study only analysed the released texts. Since the same learners completed *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime*, the sample consisted of 865 learners; the sample for *Flowers on the Roof* was 863 learners.

Figure 6.1 presents the South African Grade 5 learner results for PIRLS 2016. The PIRLS scale is from 0 to 1 000 with an average of 500 and a standard deviation of 100, which was calculated by using Item Response Theory (IRT) scaling (*cf.* Chapter Two). The South African Grade 5 learner results are depicted relative to the international mean. Overall, the South African Grade 5 learners achieved a score of 406 (SE=6.0) even though they were older than their international peers. Three of the 11 participating countries in PIRLS 2016 obtained scores lower than the international mean. South Africa's Grade 5 learners (English/Afrikaans/isiZulu) obtained the lowest scores, roughly 100 points below the international mean (*cf.* Appendix A).

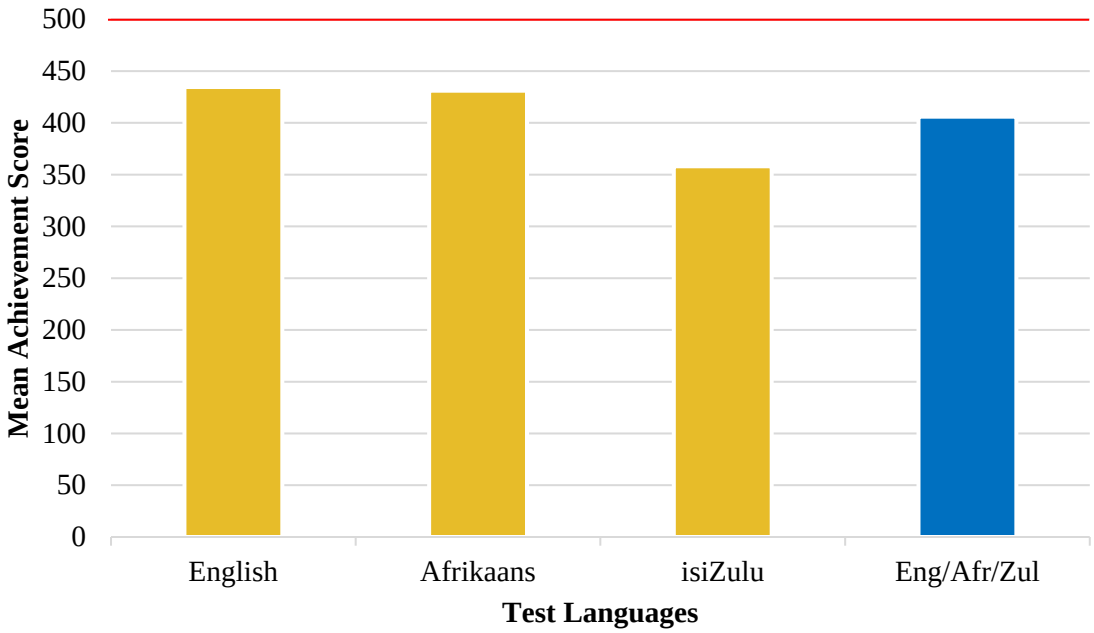

Figure 6.1: South African Grade 5 Learner Achievement across Languages

It is evident that none of the selected languages reached the international centre point, despite South African learners[32] who wrote PIRLS being a year older than their international counterparts at Grade 4 level. The learners who wrote the PIRLS 2016 assessment in English (435 points, SE=11.9) obtained the highest score among the three languages, followed by Afrikaans (431 points, SE=11.6) and isiZulu (358 points, SE=11.5). There appears to be substantial variation between the three languages as there is a 77-point difference between English and isiZulu. This point difference between the highest and lowest performing

[32] See Appendix B for more detailed information regarding the South African Grade 5 learners' results.

languages equates to almost two years of education (Howie et al., 2017b). Table 6.1 depicts the statistical significance in differences across the English, Afrikaans and isiZulu languages.

Table 6.1: Language Comparisons for South African Grade 5 Learner Achievement

Languages	Mean	SE	English	Afrikaans	isiZulu
English	435	11.9		●	▲
Afrikaans	431	11.6	●		▲
isiZulu	358	5.1	▼	▼	

▲ Significantly higher than ▼ Significantly lower than ● Not significantly different
Significance level < 0.05

Note. From *PIRLS 2016 Progress in International Reading Literacy Study 2016 Grade 5 Benchmark Participation: South African Children's Reading Literacy Achievement* (p. 54), by S. Howie et al., 2017b, Centre for Evaluation and Assessment. Copyright 2017 by Centre for Evaluation and Assessment.

Similar to the results presented in Chapter Five, no statistical difference was found between learners writing in English and Afrikaans, although both language groups performed significantly higher than the learners writing in isiZulu. The differences between the language groups may be conflated by several factors; for example, teaching and learning resources, school characteristics and the learners' own reading skills.

Because of the global role of English, the links with colonialism and capitalism, some South African parents seem to elect to send their children to English medium schools (Evans & Cleghorn, 2014) even though they speak a different language at home. This phenomenon also resonates in the PIRLS 2016 findings where only 51% of Grade 5 learners who participated in the study always spoke the language of the test at home. This finding reveals that almost half of the learners who wrote the test in English spoke a different language at home (Howie et al., 2017b). In order to understand this phenomenon, it was found that when combining the data from the learner and parent questionnaires, learners who wrote the test in the same language or a different language showed no statistical difference in their overall achievement scores (Howie et al., 2017b). When the language of the test was the same as the language the learners spoke at home, it appeared they achieved a higher score. For instance, English learners who completed the test in English obtained a score of 509 (SE=9.5), while those who had a different home language but completed the test in English obtained a score of 412 (SE=11.2) (*cf*. Howie et al., 2017b). The next section (6.2.2) gives descriptive evidence for this study's second sub-question.

6.2.2 Descriptive Statistics to Address Sub-Question 2

The second sub-question asks: *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?* This question looked at the items of the three PIRLS 2016 released texts, namely *Macy and the Red Hen*, *The Green Sea Turtle's Journey of a Lifetime* and *Flowers on the Roof*. Table 6.2 provides a breakdown of the number of items as well as maximum points per text.

Table 6.2: Total Items and Maximum Points per PIRLS 2016 Released Text

Text	Total Number of Items	Maximum Points
Macy and the Red Hen	16	19
The Green Sea Turtle's Journey of a Lifetime	16	16
Flowers on the Roof	13	16
Total	45	56

It is important to note that *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime* were both in the Reader booklet. *Flowers on the Roof* was part of other achievement booklets that were not released in the public domain (*cf.* Section 5.2.2). The total number of items across the three texts are 45 with a maximum of 56 points. The overall raw mean score for each of the texts are captured in Table 6.3.

Table 6.3: PIRLS 2016 Raw Mean Score per Text by Language

Text	Overall Mean Score	English Mean Score	Afrikaans Mean Score	isiZulu Mean Score
Macy and the Red Hen (19)	5.24	6.09	6.63	4.08
The Green Sea Turtle's Journey of a Lifetime (16)	4.85	6.06	6.19	3.12
Flowers on the Roof (16)	5.43	6.82	7.27	3.22

The results from the raw scores of the South African Grade 5 students who wrote the PIRLS 2016 released texts indicated that the English and Afrikaans language groups did not differ significantly across the released texts. However, there are noteworthy differences between the learners who wrote the PIRLS 2016 assessment in isiZulu compared to English and Afrikaans. The Grade 5 learners seemed to struggle the most with *The Green Sea Turtle's Journey of a*

Lifetime text, obtaining an overall mean score of 4.85 out of a possible 16 points. However, on closer inspection, the Afrikaans (6.19) learners achieved the highest score on all three texts. The learners who wrote *The Green Sea Turtle's Journey of a Lifetime* in isiZulu obtained the lowest score (3.12).

Tables 6.4, 6.5 and 6.6 present the percentage of learners across English, Afrikaans and isiZulu that obtained individual items correctly by PIRLS text[33]. Table 6.4 shows the percentage of learners who correctly answered the 16 items of *Macy and the Red Hen*. This text is a literary text that focuses on Macy and her struggle to control a bossy red hen. She eventually outsmarts the red hen and ascends to the top of the pecking order.

Table 6.4: Macy and the Red Hen: Number and Percentage of Grade 5 learners who correctly answered items in English, Afrikaans and isiZulu

Item	English			Afrikaans			isiZulu		
No	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
1	340	242	71	166	141	85	332	260	78
2	340	202	59	164	91	55	322	132	41*
3	341	60	18*	168	29	17*	328	21	6*
4	339	21	6*	170	28	16*	327	9	3*
5	336	185	55	164	95	58	330	147	45*
6	338	87	26*	168	57	34*	327	13	4*
7	336	109	32*	166	43	26*	322	84	26*
8	328	33	10*	165	31	19*	319	19	6*
9	333	98	29*	165	36	22*	319	41	13*
10	330	163	49*	165	75	45*	320	170	53
11	332	205	62	166	95	57	315	176	56
12	329	193	59	165	66	40*	314	108	34*
13	317	6	2*	164	2	1*	296	1	0*
14	308	52	17*	160	31	19*	279	10	4*
15	306	128	42*	161	72	45*	281	46	16*

[33] Excludes missing data.

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
16	301	62	21*	156	31	20*	276	25	9*

*Less than 50% of the responses were correct.

Table 6.4 shows the percentage of learners who correctly answered the 16 questions for the text. Based on the above table, it would appear that for all three languages, this particular text was quite difficult as only a few items were correctly answered. Four items (1, 2, 5 and 11) were correctly answered by more than 50% of learners who wrote the test in Afrikaans. Similarly, five items (1, 2, 5, 11 and 12) were correctly answered by more than 50% of the learners who wrote the test in English. However, only three items (1, 10 and 11) were correctly answered by more than 50% of the learners who wrote the test in isiZulu. These figures indicate that most learners, whether it be English, Afrikaans or isiZulu, struggled to correctly answer many of the items from *Macy and the Red Hen*. This table confirms that all three language groups had difficulty in understanding the story and were unable to answer many of the posed items correctly.

Table 6.5 shows the percentage of learners who answered items correctly from the text *The Green Sea Turtle's Journey of a Lifetime*. It is an informational text that follows the lifecycle of the green sea turtle from a hatchling to becoming a juvenile and finally reaching adulthood. The text has accompanying pictures of the green sea turtle during each stage of her life.

Table 6.5: The Green Sea Turtle's Journey of a Lifetime: Number and Percentage of Grade 5 learners who correctly answered items in English, Afrikaans and isiZulu

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
1	336	228	68	170	112	66	331	142	43*
2	335	82	24*	169	42	25*	328	11	3*
3	336	49	15*	169	20	12*	330	12	4*
4	329	58	18*	168	21	13*	321	9	4*
5	334	167	50	169	73	43*	330	121	37*
6	326	126	39*	165	43	26*	318	68	21*
7	317	19	6*	167	11	7*	318	0	0*

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
8	310	89	29*	161	42	26*	296	21	7*
9	313	101	32*	163	49	30*	312	60	19*
10	306	134	44*	163	46	28*	305	36	12*
11	293	26	9*	153	10	7*	284	0	0*
12	296	171	58	161	93	58	298	146	49*
13	294	95	32*	160	40	25*	294	52	18*
14	281	73	26*	154	46	30*	276	45	16*
15	263	118	45*	142	51	36*	263	76	29*
16	261	149	57	144	62	43*	259	100	39*

*Less than 50% of the responses were correct.

Table 6.5 depicts the South African Grade 5 English, Afrikaans and isiZulu learners' performance per item. Very few of the learners who responded to the text in English and Afrikaans correctly answered the items by more than 50%. Only items 1 and 12 for the Afrikaans group and items 1, 5, 12 and 16 were correctly answered by more than 50% of the learners in the English group. None of the items were answered correctly by 50% of the isiZulu learners, although item 12 was answered correctly by 49% and item 1 by 43% of the isiZulu learners. The results indicate that all language groups, especially isiZulu, had difficulty in answering many of the items correctly. Of concern are two items, namely item 7 and 11, where 0% of learners in the isiZulu language group were able to answer these two items. Based on information contained in the table, it can be said that the isiZulu learners had the most difficulty of the three language groups in responding correctly to the items.

Table 6.6 shows the percentage of learners who answered items correctly from the text *Flowers on the Roof* (*cf.* Section 5.2.2 for description).

Table 6.6: Flowers on the Roof: Number and Percentage of Grade 5 learners who correctly answered items in English, Afrikaans and isiZulu

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
1	347	198	57	162	86	53	339	96	28*

Item No	English			Afrikaans			isiZulu		
	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct	N Completed	N Correct	% Correct
2	340	248	73	165	123	75	326	168	52
3	342	217	63	166	102	61	339	139	41*
4	345	212	61	166	106	64	344	169	49*
5	345	85	25*	166	100	60	342	135	39*
6	342	110	32*	164	54	33*	332	29	9*
7	337	36	11*	163	35	21*	330	11	3*
8	337	123	36*	164	75	46*	324	15	5*
9	330	121	37*	165	35	21*	315	15	5*
10	330	245	74	165	117	71	319	50	16*
11	326	156	48*	164	72	44*	327	87	27*
12	315	0	0*	160	0	0*	302	0	0*
13	316	158	50	162	72	44*	316	108	34*

*Less than 50% of the responses were correct.

The results from Table 6.6 indicate that the scores of learners who responded to the text in English and Afrikaans did not differ considerably, but that there is a large difference between them and those completing it in isiZulu. Table 6.6 depicts the percentage of South African Grade 5 learners who correctly answered the items from the *Flowers on the Roof* text. Six items (1, 2, 3, 4, 10, as well as 13) were correctly answered by more than 50% of the learners who responded in English. For the learners who responded to the text in Afrikaans, six items (1, 2, 3, 4, 5 and 10) were correctly answered by more than 50% of the learners. Learners who responded to the text in isiZulu struggled more with this text as 52% of learners could correctly answer item 2 only. Curiously, none of the learners who responded to the text in English, Afrikaans, or isiZulu correctly answered item 12. The next section provides Rasch evidence that partially addresses sub-question 2.

6.3 RASCH EVIDENCE FOR ADDRESSING SUB-QUESTION 2

In this section, the overall Rasch results of *Macy and the Red Hen, The Green Sea Turtle's Journey of a Lifetime* and *Flowers on the Roof* for the Grade 5 English, Afrikaans and isiZulu learners are presented. Section 6.3.1 provides the reliability of *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime* texts. Next, the individual item-fit statistics of these

texts are discussed (6.3.2). Thereafter the differential item functioning (DIF) results are provided as it indicates the possible misfit between the languages (6.3.3). The next section provides the reliability of *Flowers on the Roof* (6.3.4), which is followed by the individual item-fit statistics (6.3.5). Subsequently, the DIF results of the *Flowers on the Roof* items are discussed (6.3.6). The last section presents a summary of the problematic items identified during this study (6.4).

6.3.1 Reliability of PIRLS 2016 Macy and Turtle Texts

Table 6.7 depicts the summary statistics for the passages *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime*. For the PIRLS 2016 analyses, the file *MacyTurtle* was created that contained the items from both *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime* texts as well as the scores from the learners who responded to these texts in English, Afrikaans and isiZulu. The summary statistics, specifically the power of analysis and person reliability indices have been calculated for the combined analysis (English, Afrikaans and isiZulu). Furthermore, these statistics were also calculated for each language.

Table 6.7: Reliability of PIRLS 2016 Macy and Turtle Texts

Analysis Name	Power of Analysis of Fit	PerSepIdx with extremes	PerSepIdx no extremes	Coefficient Alpha with extremes	Coefficient Alpha no extremes
MacyTurtle [a]	Good	0.80928	0.81236	N/A	N/A
English Only	Excellent	0.89566	0.89566	0.91563	0.91563
Afrikaans Only	Excellent	0.86592	0.86590	0.89422	0.89422
isiZulu Only	Reasonable	0.60621	0.60621	0.67907	0.67907

[a] The analysis file *MacyTurtle* included all languages.

The Cronbach Alpha (α) for *MacyTurtle* (combined languages) could not be calculated as missing data was present. As such, the summary statistics were run for each language separately across the two texts. Both the English and Afrikaans groups' tests are good as the test can strongly differentiate between different groups of persons (Combrinck, 2019). The α coefficient for both English and Afrikaans groups are above .90 and .80, respectively. However, the isiZulu group only had an α coefficient of .60. During the analysis of the isiZulu only data, two items were removed as they did not fit the model. These items included *Turtle Item 7* and *Turtle Item 11*. Both these items had zero persons for some categories and would

appear to have been very difficult for the learners. The next section looks at the individual item-fit statistics for the two texts.

6.3.2 Individual Item-fit Statistics for Macy and Turtle Texts

This section presents the item location information for the overall sample per item in order of difficulty as drawn from individual item-fit statistics. As such, it gives information of items and persons and how they link to the fit of the model. The assumption of this model is that, as the ability of the persons increase, so should the probability of correctly responding to the more difficult items (Combrinck, 2019). When there is a lack of fit, it means that this assumption is violated. The item-fit statistics also provide the chi-square values which reveal the invariance across the trait (Pallant & Tennant, 2007). In other words, the chi-square assesses whether hierarchical ordering of items is consistent across increasing levels of the trait (Combrinck, 2019), thus encompassing the property of invariance. In sum, the use of chi-square is to determine the goodness-of-fit between the observed and expected values (item-trait interaction).

For the purposes of this study, Table 6.8 includes the items for both *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime* across the three languages. The reason for selecting all three languages for this particular item-fit statistic is to test whether the null hypothesis posed is correct. The null hypothesis for this study states that the raw mean score of the learners who completed the PIRLS 2016 assessment in English, Afrikaans and isiZulu are equal (μEnglish = μAfrikaans = μisiZulu). If the null hypothesis is rejected, then the alternative hypothesis is accepted – the alternative hypothesis states that the English mean score is not equal to the Afrikaans mean score nor to the isiZulu mean score (H_a = μEnglish ≠ μAfrikaans ≠ μisiZulu).

Table 6.8: Individual Item-Fit Statistics for Macy and Turtle Texts Combined Languages

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Macy Item 1	-2.530	0.087	-0.233	16.34	0.060
Macy Item 11	-1.636	0.079	-1.979	17.88	0.037
Turtle Item 1	-1.579	0.078	-1.880	19.59	0.021
Turtle Item 12	-1.387	0.081	-0.125	6.00	0.739
Macy Item 2	-1.232	0.078	-1.235	14.92	0.093

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Macy Item 5	-1.173	0.078	4.049*	19.12	0.024
Macy Item 10	-1.163	0.078	-0.285	7.83	0.551
Turtle Item 16	-1.031	0.087	2.485	8.35	0.499
Macy Item 12	-0.909	0.079	1.668	11.97	0.215
Turtle Item 5	-0.757	0.079	4.995*	40.77	0.000**
Turtle Item 15	-0.519	0.091	1.531	12.88	0.168
Macy Item 15	-0.268	0.087	-4.490*	41.96	0.000**
Macy Item 6	-0.186	0.056	-1.785	6.12	0.728
Turtle Item 2	-0.171	0.058	-0.168	4.23	0.896
Turtle Item 6	-0.088	0.086	-4.486*	32.26	0.000**
Macy Item 7	0.000	0.087	4.182*	33.70	0.000**
Turtle Item 10	0.004	0.089	-3.651*	24.38	0.004
Turtle Item 9	0.081	0.090	3.890*	58.22	0.000**
Turtle Item 13	0.183	0.094	2.875*	23.29	0.006
Turtle Item 14	0.333	0.100	-0.500	25.16	0.003
Macy Item 9	0.411	0.094	3.009*	32.13	0.000**
Turtle Item 3	0.543	0.068	-1.974	9.73	0.373
Turtle Item 8	0.640	0.101	-4.541*	45.49	0.000**
Macy Item 16	0.964	0.111	-3.178*	22.42	0.008
Turtle Item 11	1.124	0.072	-3.124*	20.75	0.014
Macy Item 3	1.182	0.111	-1.855	10.32	0.325
Macy Item 13	1.303	0.070	-0.553	32.39	0.000
Macy Item 14	1.331	0.121	-2.084	12.38	0.193
Macy Item 8	1.489	0.123	-1.744	11.71	0.230
Turtle Item 7	1.548	0.097	-1.552	7.35	0.601
Turtle Item 4	1.604	0.127	-3.361*	30.65	0.000
Macy Item 4	1.891	0.137	0.268	33.88	0.000**

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000313 for 32 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Of the 32 items, seven items have misfits that are significant where the fit residuals were either above 2.5 or below -2.5:

- *Macy Item 15*, *Turtle Item 6* and *Turtle Item 8* displayed overfit which means that these items discriminate too much.
- In addition, the following items displayed underfit: *Turtle Item 5*, *Macy Item 7*, *Turtle Item 9* and *Macy Item 9*.

Underfit means that the information yielded from the correct responses to these items could not discriminate adequately between learners who are less able and those who are more able (van Staden, 2018). However, as the hypothesis of this study states that all three languages should have equal mean scores on the PIRLS 2016 assessment, it was necessary to investigate each language's individual item-fit statistics. Tables 6.9, 6.10 and 6.11 depict the item-fit statistics for each language.

Table 6.9: Individual Item-Fit Statistics for Macy and Turtle Texts English Only

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Macy Item 1	-2.011	0.135	-0.706	9.66	0.379
Turtle Item 1	-1.887	0.133	-1.309	7.08	0.629
Macy Item 11	-1.568	0.131	-2.036	15.69	0.074
Macy Item 2	-1.345	0.127	-1.007	16.67	0.054
Macy Item 12	-1.259	0.129	0.568	14.90	0.094
Turtle Item 12	-1.259	0.136	-0.378	14.53	0.105
Turtle Item 16	-1.245	0.146	0.770	8.38	0.497
Macy Item 5	-1.104	0.127	1.545	6.63	0.676
Macy Item 10	-0.839	0.127	-1.044	13.06	0.160
Turtle Item 5	-0.714	0.127	5.203*	62.10	0.000**
Turtle Item 15	-0.585	0.146	0.384	16.84	0.051
Turtle Item 10	-0.535	0.134	-2.356	10.91	0.282
Macy Item 15	-0.349	0.135	-1.113	11.52	0.242
Turtle Item 6	-0.269	0.132	-4.012*	18.93	0.026
Turtle Item 2	-0.248	0.090	0.596	19.83	0.019
Macy Item 6	-0.022	0.087	-0.479	10.94	0.280

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Turtle Item 13	0.134	0.144	2.767*	17.95	0.036
Turtle Item 9	0.140	0.139	1.067	12.36	0.194
Macy Item 7	0.147	0.134	3.174*	25.41	0.003
Macy Item 9	0.283	0.137	3.062*	31.12	0.000**
Turtle Item 8	0.439	0.145	-3.126*	24.91	0.003
Turtle Item 3	0.530	0.098	-1.590	9.05	0.432
Turtle Item 14	0.556	0.156	-0.659	8.79	0.457
Turtle Item 11	0.954	0.088	-1.672	12.46	0.189
Macy Item 16	0.964	0.160	-2.475	10.88	0.284
Macy Item 3	1.185	0.158	-0.724	10.87	0.285
Macy Item 14	1.269	0.167	-0.425	3.09	0.961
Turtle Item 4	1.343	0.167	-2.505*	24.19	0.004
Macy Item 13	1.462	0.101	-0.461	18.04	0.035
Turtle Item 7	1.497	0.125	-1.283	3.20	0.956
Macy Item 8	1.905	0.193	-0.929	2.96	0.966
Macy Item 4	2.429	0.225	0.787	9.09	0.429

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000313 for 32 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Table 6.9 shows the item-fit statistics for the English language subset where only two of the 32 items have significant misfit, namely *Turtle Item 5* and *Macy Item 9*. Both these items displayed underfit. As such, these two items could not adequately discriminate between more and less able learners.

Table 6.10 presents the Afrikaans language item-fit statistics.

Table 6.10: Individual Item-Fit Statistics for Macy and Turtle Texts Afrikaans Only

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Macy Item 1	-2.866	0.226	-0.274	14.47	0.107
Turtle Item 1	-1.793	0.180	-0.665	15.04	0.090

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Turtle Item 12	-1.322	0.179	-0.415	6.26	0.713
Macy Item 11	-1.317	0.176	-1.224	10.82	0.288
Macy Item 2	-1.188	0.175	-0.733	5.95	0.745
Macy Item 5	-1.128	0.176	2.402	19.36	0.022
Macy Item 6	-0.745	0.113	-1.049	5.16	0.820
Macy Item 10	-0.700	0.176	-1.556	7.60	0.575
Macy Item 15	-0.696	0.178	-2.625*	19.87	0.019
Turtle Item 16	-0.598	0.189	2.959*	12.11	0.207
Turtle Item 5	-0.567	0.175	1.287	11.84	0.223
Turtle Item 2	-0.400	0.118	1.781	7.88	0.546
Macy Item 12	-0.379	0.180	1.915	13.63	0.136
Turtle Item 15	-0.234	0.196	0.545	5.13	0.823
Turtle Item 9	0.119	0.191	2.075	43.32	0.000**
Turtle Item 14	0.215	0.199	0.747	15.61	0.076
Turtle Item 10	0.232	0.195	-1.496	9.81	0.366
Turtle Item 6	0.410	0.200	-3.519*	21.82	0.009
Turtle Item 8	0.420	0.202	-2.076	14.86	0.095
Macy Item 7	0.434	0.199	2.332	16.63	0.055
Turtle Item 13	0.459	0.204	0.860	7.58	0.577
Turtle Item 3	0.613	0.144	-1.742	4.73	0.857
Macy Item 9	0.655	0.208	1.649	13.78	0.130
Macy Item 8	0.944	0.221	-0.733	8.82	0.454
Macy Item 14	0.950	0.223	-1.121	11.01	0.275
Macy Item 16	0.956	0.225	-1.148	7.62	0.573
Macy Item 4	1.008	0.222	0.502	31.70	0.000**
Macy Item 3	1.041	0.225	-1.024	7.00	0.637
Turtle Item 11	1.186	0.146	-2.551*	11.13	0.267
Macy Item 13	1.219	0.136	0.250	15.24	0.085
Turtle Item 7	1.469	0.191	-0.027	11.09	0.270

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Turtle Item 4	1.603	0.260	-0.857	10.34	0.324

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000313 for 32 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Interestingly, none of the items that showed misfit were significant. Overall, the Afrikaans language subset had a better fit than those of the English or isiZulu language groups.

Table 6.11 depicts the isiZulu language item-fit statistics.

Table 6.11: Individual Item-Fit Statistics for Macy and Turtle Texts isiZulu Only

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Macy Item 1	-3.228	0.138	-0.056	14.36	0.110
Macy Item 11	-2.165	0.121	-0.776	14.71	0.099
Macy Item 10	-2.027	0.119	-0.731	10.80	0.290
Turtle Item 12	-1.845	0.123	0.355	10.07	0.345
Macy Item 5	-1.557	0.118	3.995*	17.96	0.036
Turtle Item 1	-1.512	0.118	0.729	7.91	0.543
Macy Item 2	-1.451	0.120	0.748	7.32	0.604
Turtle Item 16	-1.390	0.135	1.409	8.64	0.471
Turtle Item 5	-1.222	0.121	0.295	7.81	0.553
Macy Item 12	-1.159	0.125	-0.125	3.77	0.926
Turtle Item 15	-0.922	0.142	1.709	9.33	0.407
Macy Item 7	-0.733	0.132	0.307	7.83	0.552
Turtle Item 6	-0.437	0.141	-1.240	8.24	0.511
Turtle Item 9	-0.311	0.146	3.408*	50.90	0.000**
Turtle Item 13	-0.245	0.155	0.930	8.52	0.483
Turtle Item 14	-0.152	0.166	-1.196	18.49	0.030
Macy Item 15	-0.087	0.165	-2.850*	18.18	0.033
Turtle Item 2	-0.014	0.111	-0.605	7.77	0.557
Macy Item 6	-0.006	0.114	-1.260	5.00	0.834

Item	Difficulty	SE	Fit residual	Chi-Square	Probability
Turtle Item 3	0.139	0.125	-0.714	9.19	0.420
Macy Item 9	0.148	0.166	0.955	15.51	0.078
Turtle Item 10	0.350	0.179	-1.232	7.80	0.555
Macy Item 13	0.632	0.141	-1.208	5.82	0.758
Macy Item 16	0.692	0.215	-1.731	10.36	0.322
Turtle Item 8	0.930	0.225	-1.405	8.27	0.508
Macy Item 3	1.029	0.222	-1.125	6.17	0.723
Macy Item 8	1.124	0.236	-1.662	14.72	0.099
Macy Item 14	1.712	0.324	-1.276	10.37	0.322
Turtle Item 4	1.915	0.327	-1.324	3.89	0.919
Macy Item 4	1.931	0.325	-0.546	12.36	0.194
Turtle Item 7	4.548	0.311	-0.384	7.89	0.545
Turtle Item 11	5.315	0.694	-0.247	6.61	0.678

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000313 for 32 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Only one item showed significant misfit. *Turtle Item 9* displayed underfit which means that information from only one item out of the pool could not adequately discriminate between the less and more able learners. Based on the information from Tables 6.9, 6.10 and 6.11, it is evident that the item-fit statistics varied across the languages, meaning that the mean scores were not equal between English, Afrikaans and isiZulu. This finding partially rejects the null hypothesis of this study for the PIRLS 2016 assessments.

The next section focuses on the differential item functioning (DIF) of the *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime* texts across learners who responded to the texts in English, Afrikaans and isiZulu.

6.3.3 Differential Item Functioning for Macy and Turtle Texts

Possible item bias may occur when the item functions differently for different groups (*cf.* Bond & Fox, 2015; Sandilans et al., 2013). Rasch analysis can be used to identify such bias. The DIF analysis summary of the data, which includes an ANOVA test, is shown in Table 6.12. The

ANOVA was conducted to compare the mean scores of three independent groups, namely English, Afrikaans and isiZulu and to test a null hypothesis where all group means are equal (Field, 2009; Tabachnick & Fidell, 2007). The ANOVA output gives both the *F*-ratio and *p*-value. The *F*-ratio is a measure where two or more quantities are expected to be equal under the null hypothesis – the quantities include the English, Afrikaans and isiZulu mean scores. The *p*-value (or the probability value) is the statistical model that is used to provide evidence against the null hypothesis. The ANOVA output thus gives values that inform whether the differences between the mean scores of the groups are statistically significant. Table 6.12 presents the DIF summary for the two PIRLS 2016 released reading texts.

Table 6.12: DIF Summary for PIRLS 2016 Macy and Turtle Texts

Item	F-ratio	Probability
Macy Item 1	12.735	0.000*
Macy Item 11	3.951	0.020
Turtle Item 1	8.251	0.000*
Turtle Item 12	1.159	0.314
Macy Item 2	0.580	0.560
Macy Item 5	0.672	0.511
Macy Item 10	18.808	0.000*
Turtle Item 16	3.118	0.045
Macy Item 12	9.875	0.000*
Turtle Item 5	1.396	0.248
Turtle Item 15	1.360	0.257
Macy Item 15	9.018	0.000*
Macy Item 6	21.706	0.000*
Turtle Item 2	4.582	0.011
Turtle Item 6	5.118	0.006
Macy Item 7	5.683	0.004
Turtle Item 10	18.525	0.000*
Turtle Item 9	0.727	0.484
Turtle Item 13	1.419	0.243

Item	F-ratio	Probability
Turtle Item 14	1.787	0.168
Macy Item 9	1.122	0.326
Turtle Item 3	1.804	0.165
Turtle Item 8	6.177	0.002
Macy Item 16	0.062	0.940
Turtle Item 11	9.567	0.000*
Macy Item 3	0.293	0.746
Macy Item 13	9.602	0.000*
Macy Item 14	3.647	0.027
Macy Item 8	6.331	0.002
Turtle Item 7	0.460	0.632
Turtle Item 4	4.273	0.014
Macy Item 4	10.656	0.000*

**Significant at the 5 percent level (Bonferroni 0.000521)*

Table 6.12 provides a summary of the DIF results for two PIRLS 2016 texts, namely *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime*, as indicated by ANOVA statistics. Significant results which have a small *p*-value (less than <0.05) are reported in terms of uniform DIF. Uniform DIF includes persons who have the same underlying ability that have a consistently different probability of correctly responding to an item when they are compared to another/different group(s) (Andrich et al., 2012). Table 6.12 shows that approximately one-third of the items display differential item functions indicating that these items worked differently across the language groups. These items include: *Macy Item 1*, *Turtle Item 1*, *Macy Item 10*, *Macy Item 12*, *Macy Item 15*, *Macy Item 6*, *Turtle Item 10*, *Turtle Item 11*, *Macy Item 13* and *Macy Item 4*. As such, all of the items displaying DIF are analysed further in terms of their item characteristic curves (ICC).

Item characteristic curves are graphically presented below for each item displaying DIF. These items are then described with reference to the ICC. In some instances, when the items take the form of multiple choice, the distractor analysis graphs are also presented. The ICC graphs

illustrate the differential functioning across the lower and/or upper class interval[34] for each item.

Item 12 from *Macy and the Red Hen* takes the form of a multiple choice question. It requires learners to *Interpret and Integrate Ideas and Information* by asking: Macy "hit the wings with her hands and pushed the wings away." What does Macy want the hen to think?

a. that Macy is saving the hen* (correct answer)
b. that Macy is angry with the hen
c. that Macy is terrified of the owl
d. that Macy is playing with the owl

Figure 6.2 indicates that the item was considerably more difficult for the learners tested in Afrikaans (red). The learners tested in Afrikaans almost consistently fell below the IRT model (grey) whereas the learners tested in English (blue) were mostly above the model curve. The Afrikaans language sub-group at -3 person location, had a zero percent chance of correctly answering the item, whereas the isiZulu and English language sub-groups had an approximate 10% and 15% chance, respectively. Some inconsistency takes places across the language sub-groups between the -2 and -1 person locations with the English sub-group having a 50% chance of correctly responding to the item. The isiZulu sub-group at between -1 and 0 had almost a 90% probability of correctly responding to the item. The Afrikaans language sub-group at the higher class interval (between 0 and 2) had a lower characteristic curve than the model. It would appear that for the Afrikaans sub-group, a higher cognitive level was required by learners in order to answer the item.

[34] Lower and upper class intervals look at the spread of observations between certain points on the x-axis of the graph (Van Staden, 2018).

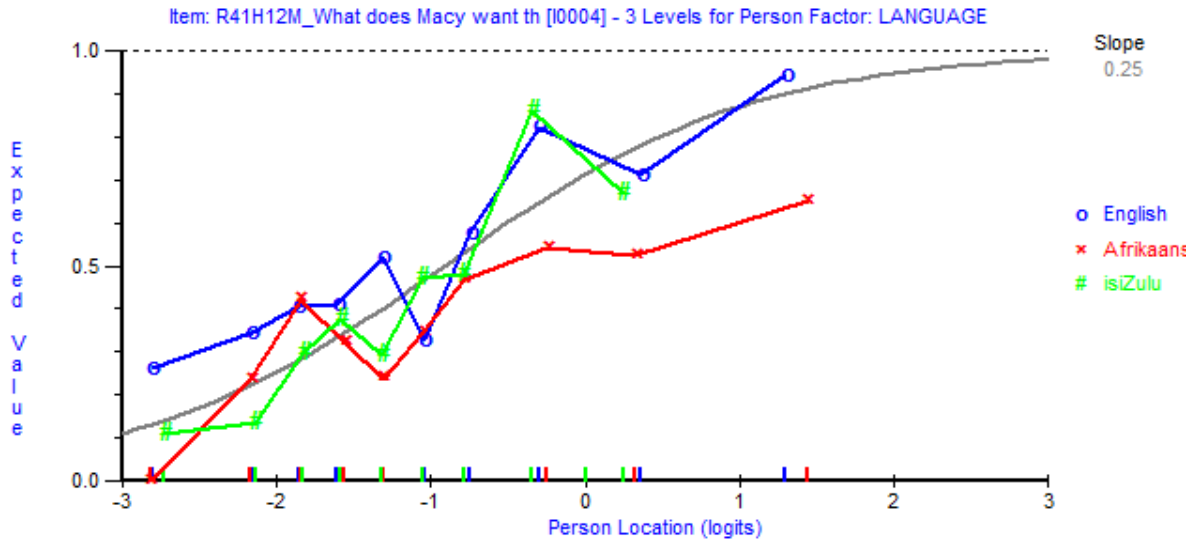

Figure 6.2: Macy Item 12 Characteristic Curve

Figure 6.3 illustrates distractor analysis for *Macy Item 12*. The correct answer for *Macy Item 12* is distractor A(1), but it seems that distractor B(2) was too tempting as learners from -3 to -1 and had approximately a 40% chance of selecting it.

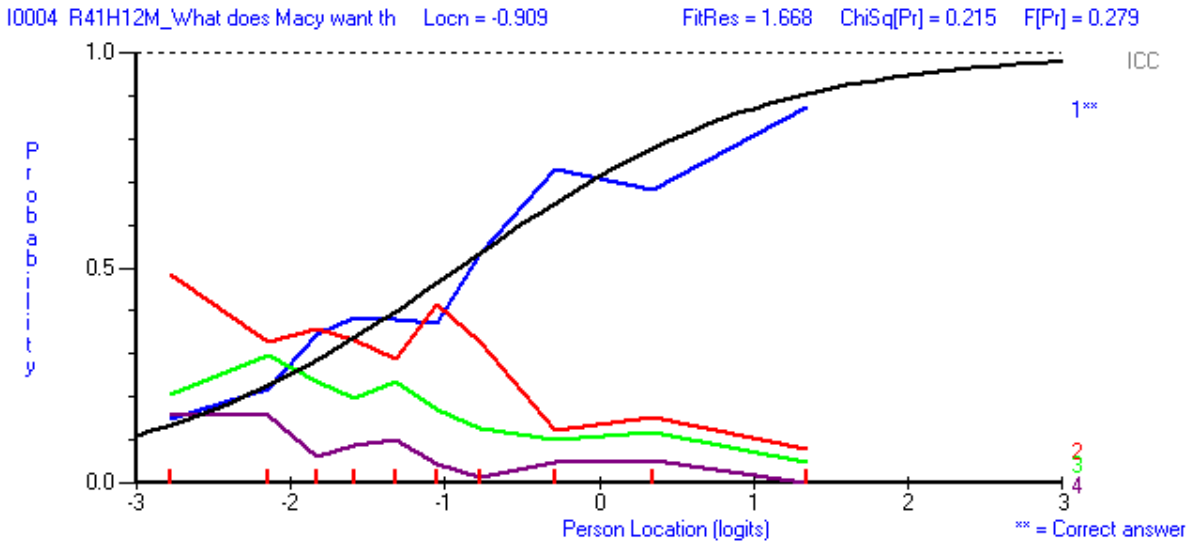

Figure 6.3: Macy Item 12 Distractor Analysis

Distractor B(2) seemed a plausible answer to the question as learners may have accessed information from the text where it indicated that Macy was frustrated with the hen. Distractor D(4) did not draw much attention as learners may have ascertained from the text that Macy did not play with the owl. Since distractor D(4) did not draw the learners' attention, it would be ideal to replace it with a different distractor.

The next item that indicated DIF is *Macy Item 1* (as illustrated in Figure 6.4). This item stems from the *Macy and the Red Hen* text and requires learners to *Focus on and Retrieve Explicitly Stated Information*. The item asks: What is Macy doing at the start of the story?

a. catching a hen
b. feeding the hens* (correct answer)
c. looking for eggs
d. collecting feathers

Figure 6.4 shows that the learners who responded to the text in English at the lower class interval were below the model curve. Although this item was easier, learners who responded to the item in Afrikaans and isiZulu were from the lower class interval and had a more than 50% chance to correctly respond to the item. Learners completing the item in Afrikaans, between the ability levels from 0 onwards, had a 100% chance of correctly responding to the items. This item was overall, more difficult for the learners who completed the test in English.

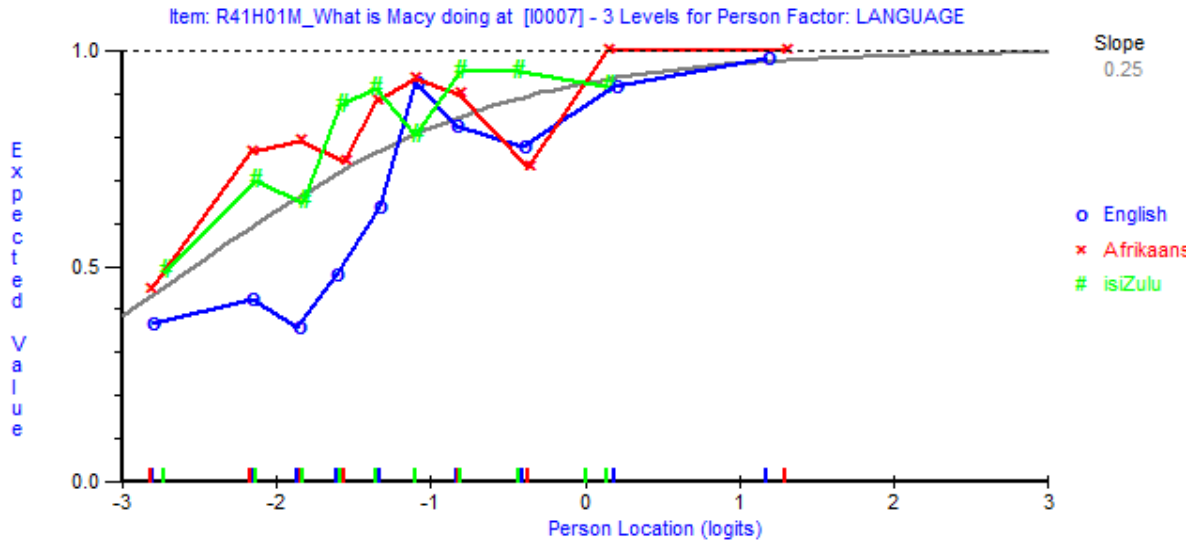

Figure 6.4: Macy Item 1 Characteristic Curve

The distractor analysis of *Macy Item 1* is shown in Figure 6.5. It would appear that, overall, the learners were able to identify the correct distractor.

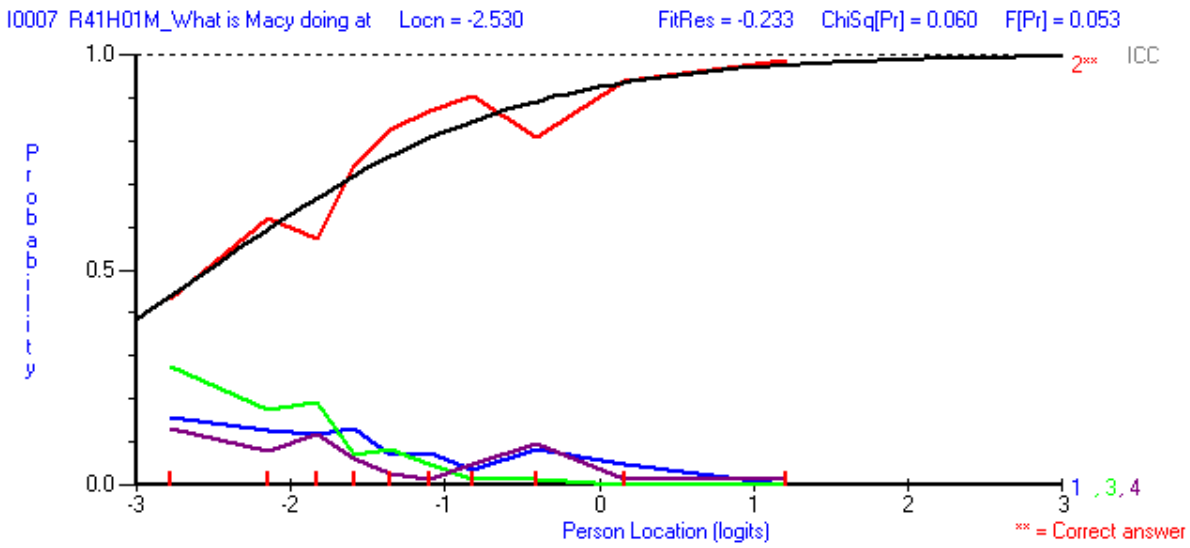

Figure 6.5: Macy Item 1 Distractor Analysis

None of the other distractors attracted much attention from the learners indicating that the question was perhaps too easy or that the distractors were not difficult enough or did not make sense. Since distractors A(1), C(3) and D(4) did not draw the learners' attention, these distractors could be refined or replaced as the current distractor analysis does not provide much clarity on the item.

Figure 6.6 depicts item 10 from the *Macy and the Red Hen* text. The item requires learners to *Focus on and Retrieve Explicitly Stated Information* from the text and asks the following question: How does Macy get her idea?

a. Macy's brother, Sam, tell her the plan.
b. Macy sees an owl catch a mouse.* (correct answer)
c. Macy's Dad tells her about owls.
d. Macy sees wire and some white cloth.

Figure 6.6 indicates that *Macy Item 10* was considerably easier for learners who responded in isiZulu as their person locations were above the model curve. The item characteristic curve for the learners who completed the test in Afrikaans and English was inconsistent in the lower class interval (between -3 and 0). The learners who responded to the item in Afrikaans were more often than not below the model curve with learners between the person location of -3 and -2 had less than a 10% chance of answering the item correctly. The learners who completed the test in English with person location between 1 and 2, had a roughly 90% chance of correctly

responding to the item. The results from Figure 5.6 can be explained as item discrimination towards those learners who completed the test in Afrikaans and English in the lower class interval.

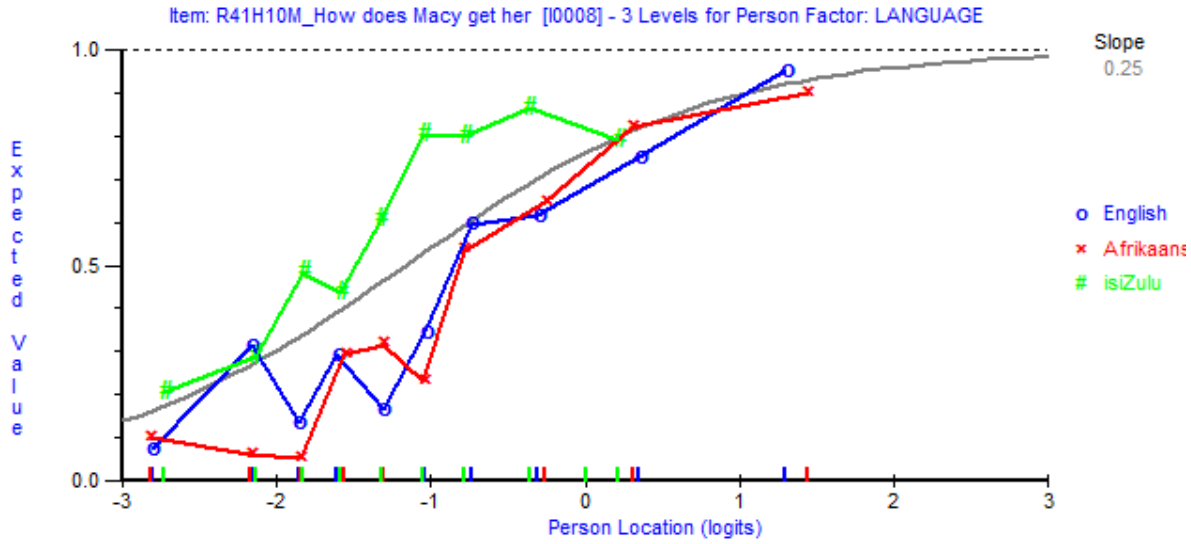

Figure 6.6: Macy Item 10 Characteristic Curve

The distractor analysis of *Macy Item 10* is depicted in Figure 6.7. The correct response is distractor B(2).

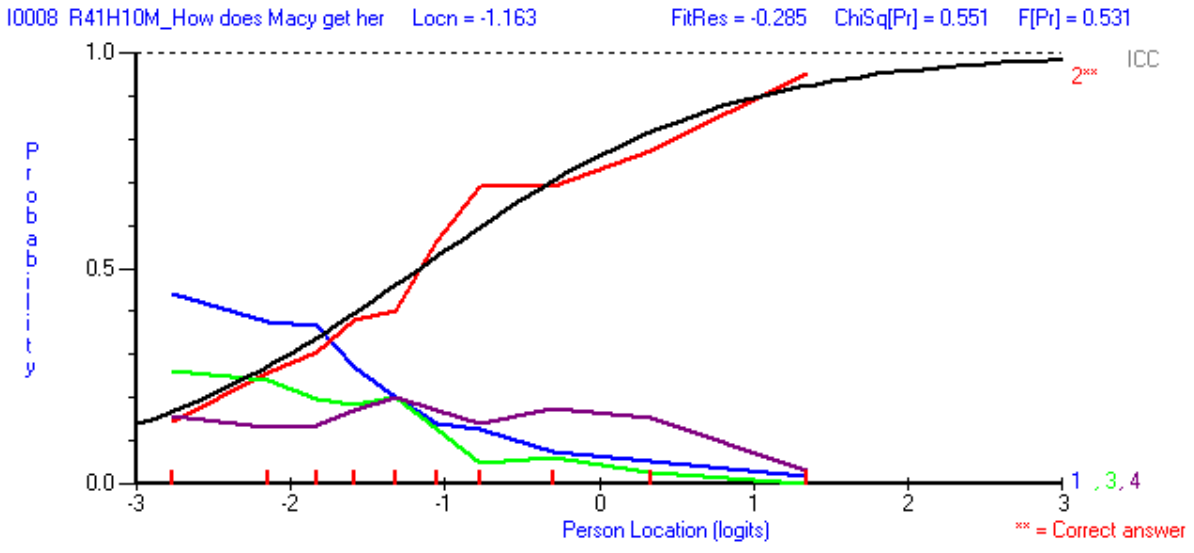

Figure 6.7: Macy Item 10 Distractor Analysis

The above figure indicates that learners from person location -1.3 onwards had a higher probability of selecting the correct answer. Distractor A(1) was tempting for learners between person locations -3 and -1.8. Moreover, these learners had an approximate 40% chance of

selecting it. The reason for selecting this distractor may be partially due to the information contained in the text – it says that "she explained her plan to her brother Sam" [35] but only after the owl swooping down to catch the mouse gave her the idea to mimic it. This sentence may have caused confusion for the learners at the -3 to -1.8 person locations.

The next item originates from *The Green Sea Turtle's Journey of a Lifetime* text. Figure 6.8 presents the characteristic curve for *Turtle Item 1*, which required learners to *Make Straightforward Inferences*. The item is as follows: What is the first section "Out From the Sand" about?

a. what different sea turtles look like
b. how sea turtles learn to swim
c. what sea turtles like to eat
d. how sea turtles' eggs hatch* (correct answer)

Figure 6.8 indicates that the item characteristic curve for learners who completed the test in Afrikaans and isiZulu was inconsistent in the lower class interval (between -3 and 0). *Turtle Item 1* was considerably more difficult for those learners who completed the test in isiZulu. These learners' person locations were mostly below the model curve. The characteristic curve for the English language group mostly followed the model curve with learners between 0 and 1.2 having a more than 90% chance of correctly responding to the item. There is a sharp rise in the characteristic curve for the Afrikaans language group between -1.8 and -1.3 person locations, having a less than 20% chance of correctly answering the item to almost 80%. However, a drop in the characteristic curve is shown between -1.3 and -1.1 person locations meeting the expected model curve. The information contained in this figure gives evidence that there is item discrimination against Afrikaans and isiZulu learners in the lower class intervals (that is, between -2 and 0).

[35] *cf*. Page 5, paragraph 1 of the *Macy and the Red Hen* text.

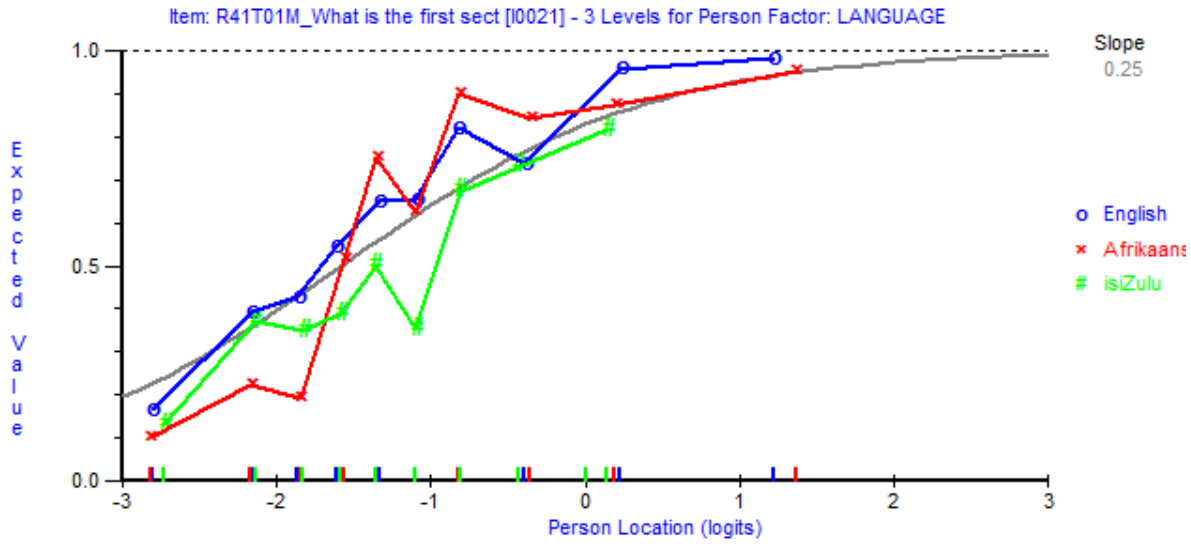

Figure 6.8: Turtle Item 1 Characteristic Curve

The distractor analysis of *Turtle Item 1* is shown in Figure 6.9 and the correct answer for this item is distractor D(4).

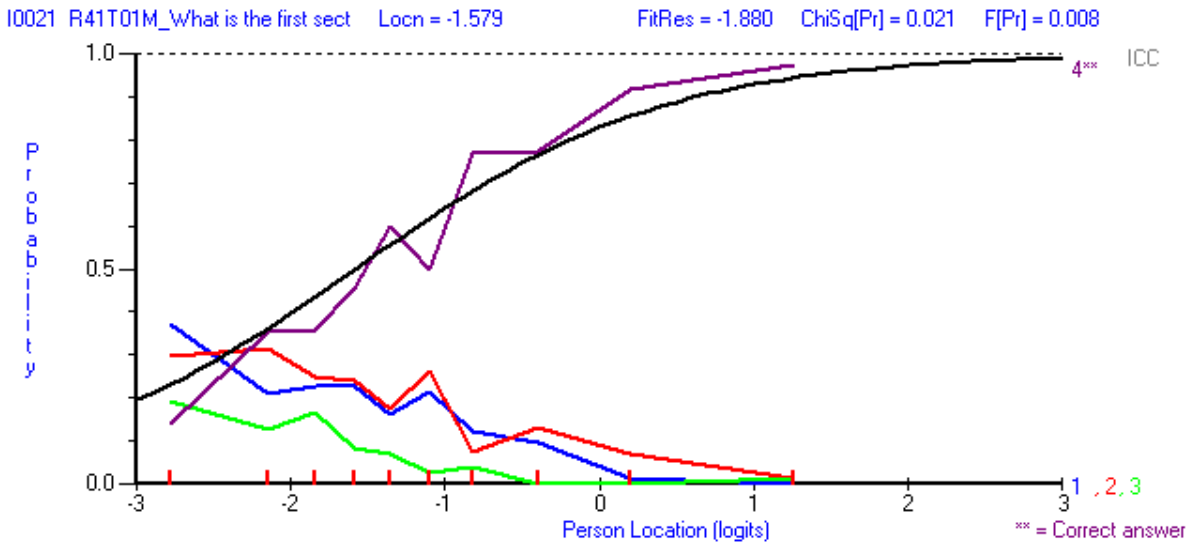

Figure 6.9: Turtle Item 1 Distractor Analysis

Figure 6.9 indicates that learners in the lower class interval, between -3 and 0, had inconsistent characteristic curves across all the distractors. Moreover, the chi-square is significant which indicates that the distribution of distractors is significantly different from the expected model. This deviation from the expected model could be partly due to learners at the -2.7 person location who had an approximate 40% chance of selecting distractor A(1) and a 30% chance of selecting distractor B(2). These learners may have recalled what they had read throughout

the entire text rather than focusing on the specific section referred to in this item. However, from -1.8 onwards, the learners were more able to identify the correct answer from the pool of distractors. Very few learners selected distractor C(3), which included what sea turtles eat.

The next item, which displayed DIF across the three languages, is *Macy Item 4* (as illustrated in Figure 6.10). This item is a constructed response type item, consisting of 1 mark that required learners to *Interpret and Integrate Ideas and Information*. The item asked: Why does the red hen play tricks on Macy?

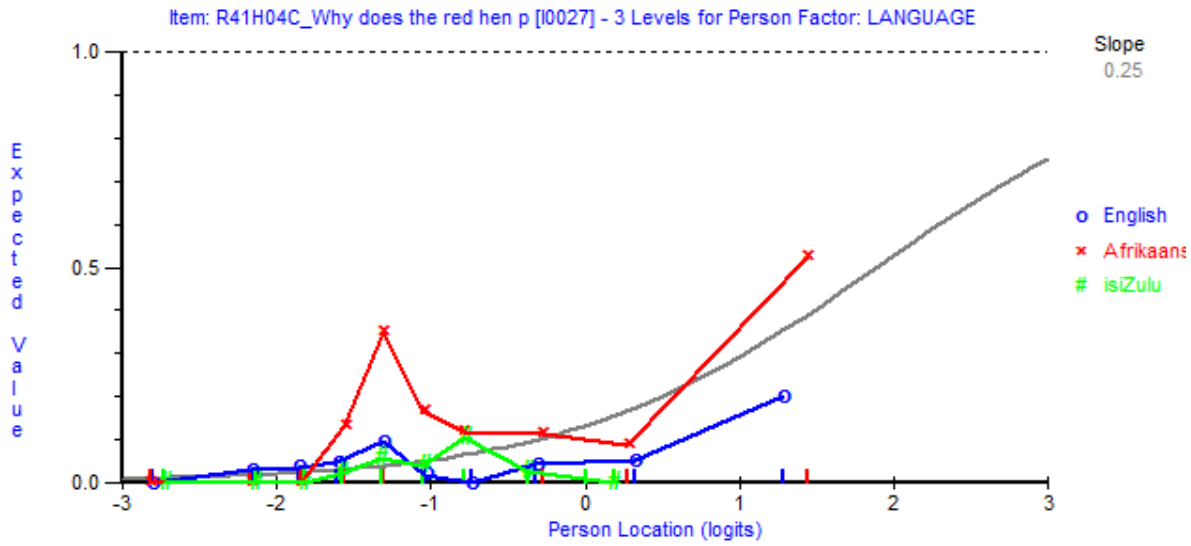

Figure 6.10: Macy Item 4 Characteristic Curve

Figure 6.10 indicates that the item was extremely difficult across the three language sub-groups on the lower class interval (between -3 and 0), except for the anomaly for those learners who completed the test in Afrikaans at -1.3 who had an approximate 40% chance of answering the item correctly. This item seemed to tax the learners' cognitive abilities as those who completed the test in English and Afrikaans, with higher abilities (between 1 and 2), had an approximate 20% and 50% chance respectively, to correctly answer the question. An acceptable response would refer to the hen being at the top of the pecking order or being the boss (*cf.* Appendix C). This figure provides evidence that this item was very difficult for all language sub-groups as very few learners had a chance to correctly respond to the item. However, as commented above, some of the learners who completed the test in Afrikaans had a higher chance of correctly answering this item than the other two language sub-groups.

The next figure illustrates the characteristic curves for *Macy Item 15* (as presented in Figure 6.11). This item also displayed differential functioning across the language sub-groups. *Macy Item 15* is a constructed response type item with a total mark allocation of one. This item required learners to *Interpret and Integrate Ideas and Information* and asked the following question: What do you think the red hen will do next time Macy puts the hens in their cage?

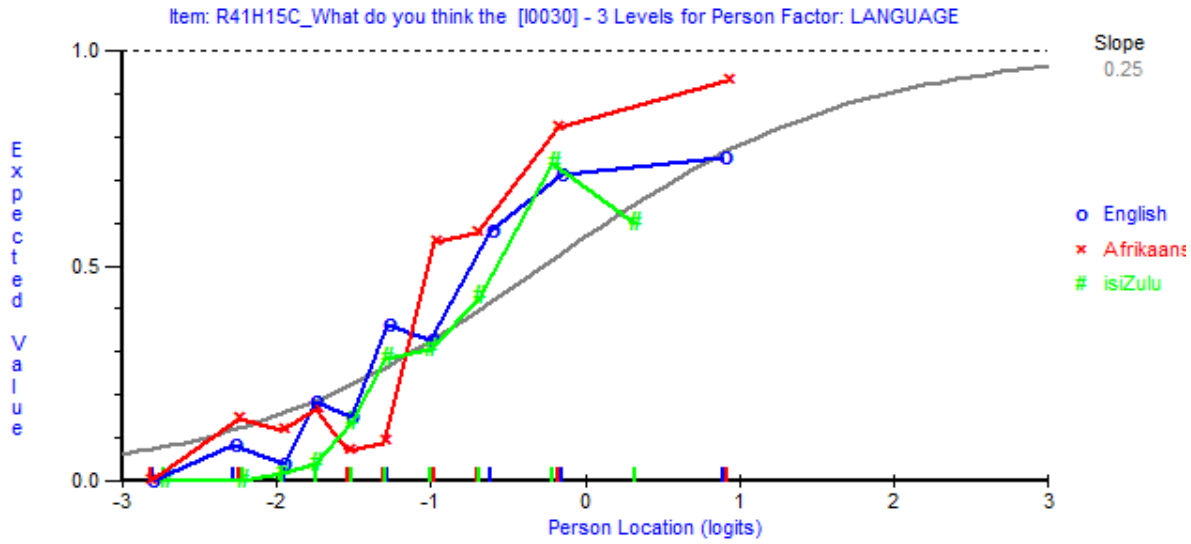

Figure 6.11: Macy Item 15 Characteristic Curve

All three language sub-groups followed a similar pattern to the model curve. Yet, much inconsistency is found in the lower class interval, specifically between -3 and -1. All three language sub-groups at the -3 person location had zero percent chance of correctly responding to the item. For this item, the learners obtained a mark if their answer related to the hen behaving herself (*cf.* Appendix C). Learners who completed the test in isiZulu, who are between the person locations of -2 and -1, experienced a steady increase in chance to complete the item correctly. At the -1.7 person location, these learners had a less than 10% chance of correctly answering the item, and at -1.5, -1.2 and -1 the learners had an approximate 15, 30 and 30 percent chance, respectively, of correctly responding to the item. Between the -1 and 0 person location, all three language sub-groups' characteristic curves were above that of the model curve; however, the learners who completed the test in Afrikaans had the highest chance of getting the item correct. These learners at the -0.9 had an approximate 60% chance of answering the item correctly. At the higher class internal (0 to 1), the learners who completed the test in isiZulu had an approximate 60% chance of correctly responding to the item, whereas the English sub-group had a 70% chance and those learners who completed the test in

Afrikaans, had the highest chance (90%) of correctly responding to the item. Both the English and the isiZulu subgroups between the 0 and 1 person location fell below the model curve.

Figure 6.12 shows the characteristic curves for *Turtle Item 10* as it displayed differential functioning across the three languages. This item takes the form of a constructed response type item for one mark. *Turtle Item 10* entails learners to *Focus on and Retrieve Explicitly Stated Information* from the text. It posed the following question: Why does a sea turtle's body fat become green?

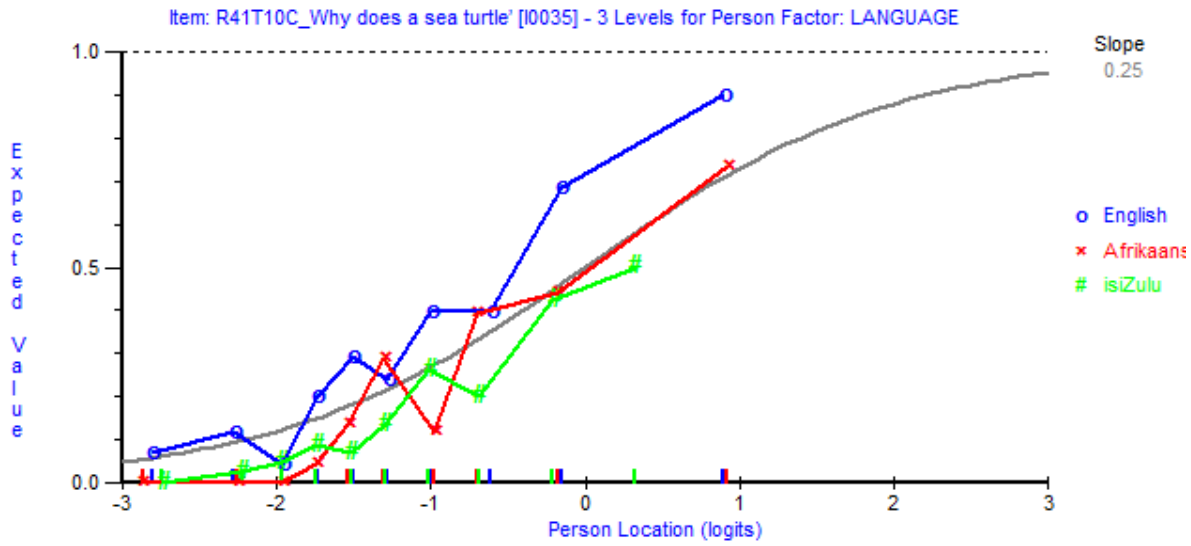

Figure 6.12: Turtle Item 10 Characteristic Curve

For this item, learners had to find evidence from the text about why the green sea turtle's body fat becomes green – because of eating green foods such as algae or turtle grass (*cf.* Appendix C). *Turtle Item 10* appears to have been more difficult for the Afrikaans and isiZulu language sub-groups. Moreover, there is some inconsistency across all three language sub-groups at the lower class interval (between -3 and 0). Both Afrikaans and isiZulu language sub-groups at -3 had zero percent chance of correctly responding to the item. However, the English sub-group had a very low chance (approximately 8%) but was above the model curve. The Afrikaans language sub-group had a very poor chance of correctly answering the item from -3 to -2. From -2 to -1, these learners' chance of correctly answering the item grew to roughly 30% then experienced a drop at -1 where they had about a 10% chance. Overall, the learners who completed the test in English had a higher chance of correctly responding to the item and their characteristic curve was more often than not above the model curve. These learners at 0.9 had

an approximate 90% chance of correctly responding to the item, whereas the Afrikaans language sub-group had an approximate 70% chance of answering the item correctly.

Figure 6.13 displays the differential functioning for *Macy Item 6* across the three language sub-groups. This constructed response type item consisted of two marks and was a *Focus on and Retrieve Explicitly Stated Information* type item. This item asked the following question: Macy wants the red hen to go into the cage. What are two things Macy does that do not work?

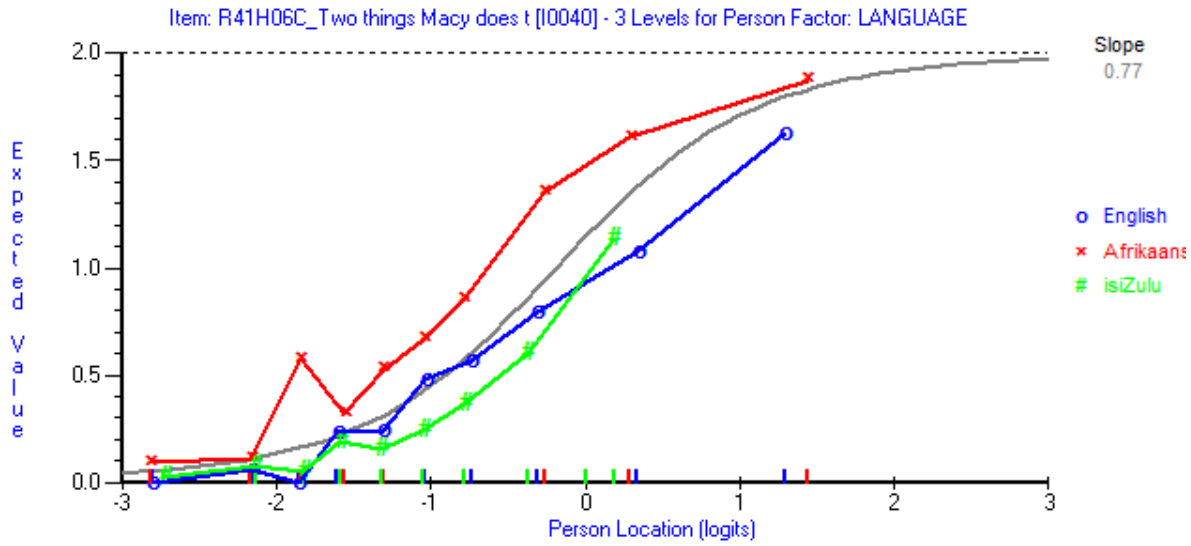

Figure 6.13: Macy Item 6 Characteristic Curve

Figure 6.13 shows the characteristic curves for the three language sub-groups. As evidenced in the above figure, the three characteristic curves follow that of the model curve. In order to obtain full marks, learners had to include two of the following answers: picking up the red hen; luring the hen into the cage; making a noise; or chasing the red hen (*cf.* Appendix C). This item was particularly difficult for all three language sub-groups at the lower class interval (-3 to 0). The Afrikaans sub-group's characteristic curve was consistently above that of the model curve suggesting that the Afrikaans sub-group found this item to be less difficult compared to the other two language sub-groups. The Afrikaans sub-group between the 1 and 2 person location had an approximate 90% chance of obtaining full marks for the item. This item discriminated against the English and isiZulu language sub-groups as the item did not function similarly to the Afrikaans sub-group.

Macy Item 13 (as illustrated in Figure 6.14) was a constructed response type item that measured the learners' ability to *Interpret and Integrate Ideas and Information* by asking: You learn what

Macy is like from the things she does. Describe what Macy is like and give two examples from the story that show this. The item consisted three marks.

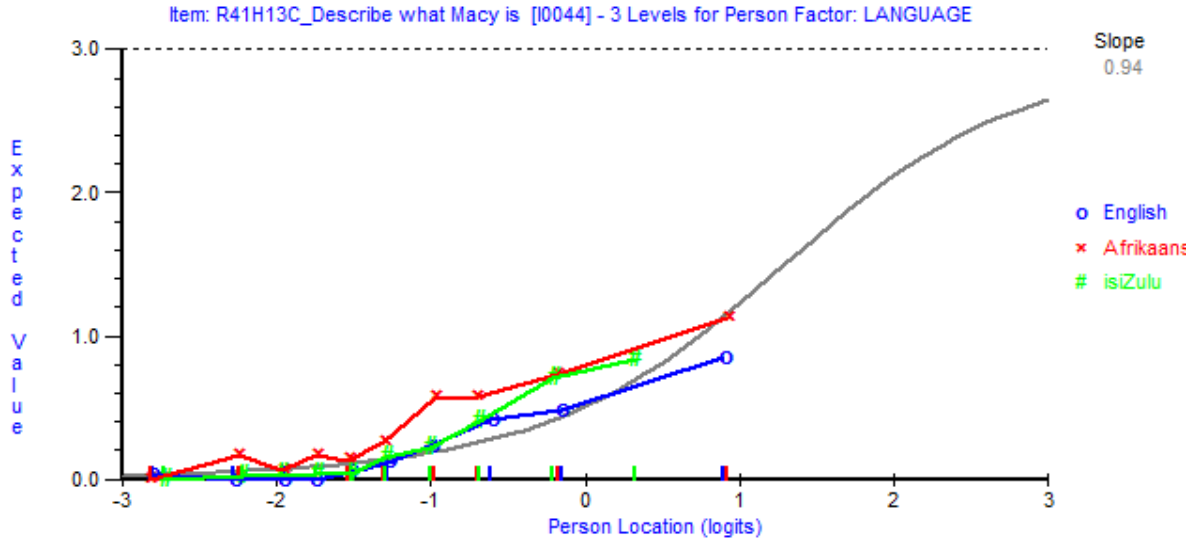

Figure 6.14: Macy Item 13 Characteristic Curve

Figure 6.14 depicts the ICC for Macy Item 13 and reveals that this item was particularly difficult for all three language sub-groups at the lower class interval (between -3 and 0). This item has the highest mark allocation for the *Macy and the Red Hen* text. In order to obtain the three marks, learners had to apply in-depth comprehension skills and provide a response that shows what Macy is like (a personal trait or attitude) which is supported by two different examples or by providing two traits (or attitudes) and an example for each (*cf.* Appendix C). Table 6.14 showed that very few learners were able to obtain full marks. Only 2% of the learners who completed the test in English obtained full marks, whereas 1% of those in Afrikaans and zero percent of those who completed the test in isiZulu received full marks. Even though this item was extremely difficult, the learners who completed the test in Afrikaans, between the -1 and 1 person locations, had a higher chance of responding correctly to the item. This item does not provide much clarity other than the fact that the learners from all three language sub-groups found this item taxing. It may be that it was difficult for the learners to infer from the text the traits, attitudes or feelings Macy exhibited or to provide examples based on the information contained in the text to support their answers. In order to obtain one mark, learners had to show minimal comprehension by giving a response that lists at least one of Macy's traits with a vague example or without an example, such as 'Macy is clever'.

Turtle Item 11 (as illustrated in Figures 6.15), a constructed response type item, required learners to complete a table with some missing information. This item required learners to *Interpret and Integrate Ideas and Information* from the text about the turtle's size and food at each stage of its life and consisted three marks. The item asked: What information does the article provide about the sea turtle's size and food at each stage of its life? Complete the table below. Three have been done for you (see Appendix C for the complete example of this item).

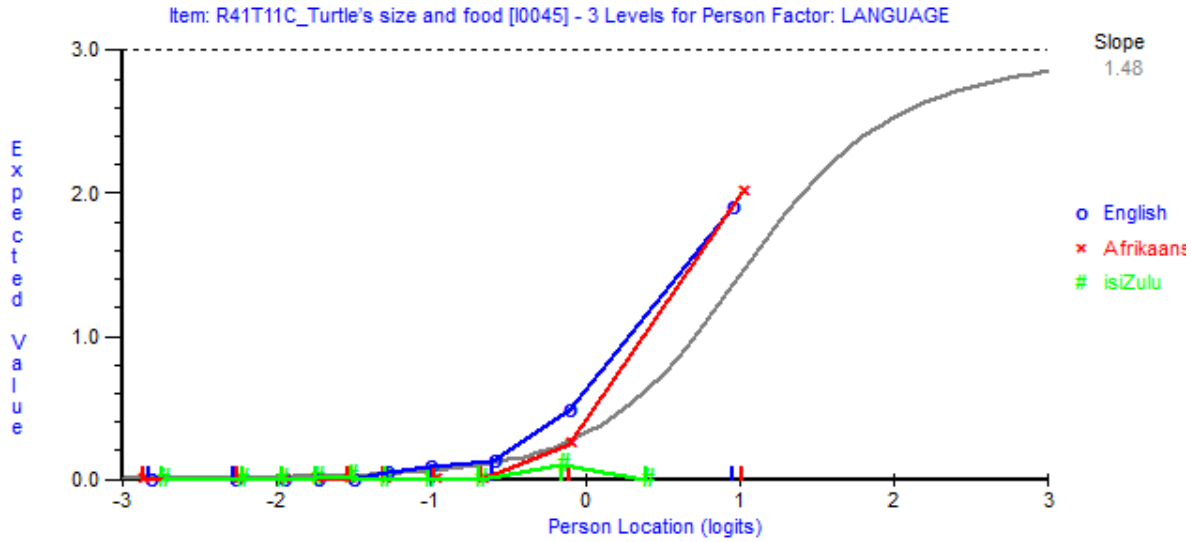

Figure 6.15: Turtle Item 11 Characteristic Curve

Turtle Item 11 was extremely difficult for all three language sub-groups as these learners had zero percent chance of correctly responding to the item from -3 to -1 person locations (lower class interval). Yet, the learners completing the test in either English or Afrikaans and who had abilities between the 0 and 1 person location, had a higher chance of partially answering the item correctly. In order to score full marks for this item, learners had to enter the missing information in the table, which included five pieces of information. These include the size of the egg, size of the hatchling, food that the hatchling consumes, food that the juvenile consumes and the size of the adult green sea turtle. These answers are outlined in Appendix C. Table 6.5 depicts the percentage of learners who were able to correctly answer this item. Only nine percent and seven percent of the learners who completed the test in English and Afrikaans were able to obtain full marks for this item, whereas zero percent of the isiZulu sub-group were able to score full marks. Learners were able to get two marks if they completed four out of the five missing pieces and for one mark, learners were required to complete three out of the five missing pieces. This item was extremely difficult for all three the language sub-groups and did

not yield more information other than the difficulty. The underlying problem with this particular item might be the format of the item. It takes the form of a table and even though South African Grade 5 learners are taught how to complete tables, it would appear that doing so is nonetheless difficult for learners. This item requires learners to infer from the text but the information is contained within the text under the relevant headings.

6.3.4 Reliability of PIRLS 2016 Flowers Text

Table 6.13 presents the summary statistics for *Flowers on the Roof*. For this analysis, the file *Flowers* was created that only included the items from the *Flowers on the Roof* text. The analysis also included the scores from the learners who responded to this text in English, Afrikaans and isiZulu. The power of analysis and the person reliability indices for the combined analysis (English, Afrikaans and isiZulu) were calculated. In addition, these statistics were also calculated separately for each language.

Table 6.13: Reliability of PIRLS 2016 Flowers Text

Analysis Name	Power of Analysis of Fit	PerSepIdx with extremes	PerSepIdx no extremes	Coefficient Alpha with extremes	Coefficient Alpha no extremes
Flowers [a]	Good	0.70348	0.69179	N/A	N/A
English Only	Good	0.76953	0.74659	0.80979	0.78492
Afrikaans Only	Good	0.75546	0.73877	0.79799	0.78122
isiZulu Only	Low	0.40309	0.36615	0.56818	0.53067

[a] The analysis file *Flowers* included all languages.

The above statistics for *Flowers* (combined languages) do not include the Cronbach Alpha (α) as missing values were present. Therefore, the summary statistics were run for each language separately. The English and Afrikaans language sub-groups' tests can adequately differentiate between different groups of persons (Combrinck, 2019). The α coefficient for the English sub-group is above .80 whereas the Afrikaans sub-group is slightly below .80. Notably the isiZulu sub-language group had an α coefficient of only .5, which is below the accepted range and also shows a low separation index (*PerSepInx*) (Combrinck, 2019). This means that the scale cannot adequately differentiate between groups of persons. The following sections explore the individual item-fit statistics of the *Flowers on the Roof* text.

6.3.5 Individual Item-fit Statistics for Flowers Text

Table 6.14 shows the individual item-fit statistics for the overall sample per item in order of difficulty. This table provides information of the items as well as the persons and how they link to the fit of the model. As mentioned in Section 6.3.4, the assumption of the model states that as the person ability increases, so should the probability of correctly responding to the more difficult items (Combrinck, 2019). If there is a lack of fit, it results in the violation of the assumption. Table 5.14 also provides the chi-square values and discloses the invariance across the trait (Pallant & Tennant, 2007). This value determines the goodness-of-fit of the item-trait interaction.

Table 6.14 depicts the 13 items of the *Flowers on the Roof* text across English, Afrikaans and isiZulu languages. In order to determine whether the null hypothesis for this study can be rejected, the item-fit statistics is examined. The null hypothesis states that the mean score of the learners who completed the *Flowers on the Roof* text as part of the PIRLS 2016 assessment, across the three languages, are equal (μEnglish = μAfrikaans = μisiZulu). However, if the null hypothesis is rejected, then the alternative hypothesis could be accepted – the alternative hypothesis states that the English mean score is not equal to that of the Afrikaans nor the isiZulu mean scores (H_a = μEnglish ≠ μAfrikaans ≠ μisiZulu).

Table 6.14: Individual Item-Fit Statistics for Flowers Text Combined Languages

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 2	-1.316	0.083	2.422	11.86	0.221
Flowers Item 4	-1.042	0.080	-0.734	17.94	0.036
Flowers Item 3	-0.906	0.080	-1.654	31.37	0.000**
Flowers Item 10	-0.717	0.082	-1.821	20.61	0.014
Flowers Item 1	-0.414	0.080	-4.197*	39.18	0.000**
Flowers Item 13	-0.223	0.084	4.952*	57.06	0.000**
Flowers Item 11	-0.014	0.083	3.838*	18.87	0.026
Flowers Item 5	0.027	0.082	4.915*	45.77	0.000**
Flowers Item 9	0.309	0.058	-4.052*	26.69	0.002
Flowers Item 8	0.857	0.092	-5.560*	65.68	0.000**
Flowers Item 6	0.972	0.093	-2.427	16.59	0.056

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 7	1.170	0.068	0.768	18.44	0.030
Flowers Item 12	1.296	0.072	-2.041	10.22	0.333

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000769 for 13 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

The *Flowers on the Roof* text had 13 accompanying items, of which four displayed a significant misfit. These items can be categorised by their misfit. Items 1 and 8 displayed overfit and as such, discriminate too much against the learners. Items 5 and 13 display underfit and as a result, the information gained from these items indicate that these items are discriminating too little between more able and less able learners (van Staden, 2018). In order to make an informed decision regarding the null hypothesis, it is necessary to examine each language sub-group's individual item-fit statistics. The next three tables (6.15, 6.16 and 6.17) shows the item-fit statistics for each language.

Table 6.15: Individual Item-Fit Statistics for Flowers Text English Only

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 10	-1.519	0.146	-0.596	8.12	0.522
Flowers Item 2	-1.279	0.137	1.496	15.60	0.076
Flowers Item 3	-0.952	0.132	-1.965	17.91	0.036
Flowers Item 4	-0.802	0.130	-1.157	14.40	0.109
Flowers Item 1	-0.595	0.128	-1.735	21.62	0.010
Flowers Item 13	-0.081	0.132	2.664*	20.67	0.014
Flowers Item 9	-0.044	0.088	-3.262*	19.88	0.019
Flowers Item 11	0.050	0.130	2.860*	10.83	0.288
Flowers Item 8	0.730	0.133	-3.507*	30.74	0.000**
Flowers Item 6	0.864	0.134	-1.086	6.12	0.728
Flowers Item 12	0.890	0.096	0.158	2.27	0.986
Flowers Item 5	1.304	0.142	1.409	11.58	0.238
Flowers Item 7	1.435	0.099	1.780	22.78	0.007

*Fit residuals are indicated if above +2.5 or below -2.5

**Bonferroni adjustment is 0.000769 for 13 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Table 6.15 shows the English sub-group's item fit statistics. Only one of the 13 items displayed a significant misfit, namely *Flowers Item 8*. This item displayed overfit and over discriminates against the learners with different abilities.

Table 6.16 displays the Afrikaans sub-group item fit statistics for the *Flowers on the Roof* text.

Table 6.16: Individual Item-Fit Statistics for Flowers Text Afrikaans Only

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 2	-1.333	0.200	1.625	19.42	0.022
Flowers Item 10	-1.221	0.196	0.878	7.88	0.546
Flowers Item 4	-0.862	0.188	-0.550	9.72	0.374
Flowers Item 3	-0.713	0.186	0.762	13.13	0.157
Flowers Item 5	-0.599	0.184	1.181	30.35	0.000**
Flowers Item 1	-0.229	0.183	-2.916*	14.76	0.098
Flowers Item 8	0.107	0.182	-1.976	17.59	0.040
Flowers Item 13	0.175	0.183	0.885	13.06	0.160
Flowers Item 11	0.251	0.183	-0.631	5.11	0.824
Flowers Item 9	0.780	0.127	-0.924	8.88	0.449
Flowers Item 7	0.824	0.126	0.296	7.47	0.589
Flowers Item 6	0.886	0.191	-0.636	15.93	0.068
Flowers Item 12	1.933	0.164	-0.075	11.42	0.248

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000769 for 13 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

The information contained in Table 6.16 reveals that only one item, *Flowers Item 1*, had a misfit but it is not significant. Based on this information, it could be said that the Afrikaans sub-group has a better fit than the English sub-group.

Table 6.17 shows the item-fit statistics of the isiZulu sub-group.

Table 6.17: Individual Item-Fit Statistics for Flowers Text isiZulu Only

Item	Difficulty	SE	Fit Residual	Chi-Square	Probability
Flowers Item 2	-1.483	0.123	1.471	4.64	0.864
Flowers Item 4	-1.468	0.120	-1.128	13.94	0.124
Flowers Item 3	-1.077	0.121	-1.503	19.13	0.024
Flowers Item 5	-1.006	0.122	-0.709	9.49	0.394
Flowers Item 13	-0.719	0.129	3.660*	21.71	0.010
Flowers Item 1	-0.428	0.129	-1.928	12.07	0.210
Flowers Item 11	-0.353	0.133	3.077*	20.58	0.015
Flowers Item 10	0.308	0.154	0.370	16.47	0.058
Flowers Item 9	0.411	0.113	-2.264	10.21	0.334
Flowers Item 7	0.749	0.133	-0.846	4.70	0.860
Flowers Item 6	1.123	0.192	-1.253	7.86	0.548
Flowers Item 12	1.967	0.296	-1.435	5.79	0.671
Flowers Item 8	1.973	0.266	-1.967	13.51	0.141

*Fit residuals are indicated if above +2.5 or below -2.5.

**Bonferroni adjustment is 0.000769 for 13 items. All items smaller than the Bonferroni adjustment are highlighted and are significant.

Table 6.17 presents the item-fit statistics of the isiZulu language sub-group. Of the 13 items, none of the items that displayed misfit was significant. The two items in question are *Flowers Item 11* and *Flowers Item 13*. As with the Afrikaans sub-group, the isiZulu sub-group had a better fit than the English sub-group. It is evident, based on the information from Tables 6.15, 6.16 and 6.17 that the item-fit statistics varied across the languages. This finding partly assists in rejecting the null hypothesis, which stated that the mean scores of the languages are equal. The next section specifically looks at the items that functioned differently across English, Afrikaans and isiZulu.

6.3.6 Differential Item Functioning for Flowers Text

This section focuses on the items from the *Flowers on the Roof* text that displayed differential functioning across the language sub-groups (*cf.* Section 6.3.3). Table 6.18 provides an opportunity to compare the mean scores of English, Afrikaans and isiZulu languages.

Moreover, it provides a means to test the null hypothesis[36] where all groups have the same or equal means (Field, 2009; Tabachnick & Fidell, 2007).

Table 6.18: DIF Summary for PIRLS 2016 Flowers Text

Item	F-ratio	Probability
Flowers Item 4	4.888	0.008
Flowers Item 11	1.955	0.142
Flowers Item 1	2.777	0.063
Flowers Item 13	6.994	0.001*
Flowers Item 2	0.026	0.974
Flowers Item 3	0.558	0.572
Flowers Item 5	77.451	0.000*
Flowers Item 6	2.139	0.119
Flowers Item 8	38.581	0.000*
Flowers Item 10	74.900	0.000*
Flowers Item 7	6.800	0.001*
Flowers Item 9	24.729	0.000*
Flowers Item 12	38.084	0.000*

*Significant at the 5 percent level (Bonferroni 0.001282)

The above table provides a summary of the DIF for the *Flowers on the Roof* text as proven by the ANOVA statistics, and indicates those with a *p*-value smaller than <0.05 as significant. The specific DIF referred to above, is uniform DIF, which includes persons with the same ability that have a consistently different probability of correctly answering an item in comparison to other groups (Andrich et al., 2012). Of the 13 items, seven displayed differential functioning across the three languages. These include items 13, 5, 8, 10, 7, 9 and 12. The remainder of this section provides the graphical presentation of the item characteristic curves (ICC) of the aforementioned items. Of the seven DIF items, only two were multiple choice questions. These items are accompanied by distractor analysis graphs. The remainder of the items took the form of constructed response type questions.

[36] The ANOVA also provides the F-ratio and p-value. The former measures two or more quantities that are expected to be the same under the null hypothesis. The latter refers to the statistical model that is given as evidence to whether the null hypothesis can be accepted or not.

Item 13 of the *Flowers on the Roof* text is a multiple choice question and required learners to *Evaluate and Critique Content and Textual Elements* by asking: Which of the following might you learn from this story?

a. Old people will never be happy if they change where they live.
b. You can make a new place feel like home if you bring familiar things with you.* (correct answer)
c. You can get used to living with animals, even though they are noisy.
d. Children and old people do not make good friends.

Figure 6.16 (following page) reveals that the item had substantial inconsistency across the three language sub-groups, particularly at the lower case interval (between -3 and 0). The item was somewhat more difficult for the learners tested in English (blue) and Afrikaans (red). The learners tested in Afrikaans frequently fell below the model curve, except at -1.3 and -0.9 person locations. For these locations, the Afrikaans sub-group had approximately a 40% and 60%, chance, respectively, of correctly responding to the item. At the higher class interval (between 0 and 3), the ICC for the Afrikaans and English sub-groups were almost identical, leading to the conclusion that for both these language groups at the higher class interval, the item functioned similarly to one another. However, for the learners who completed the test in isiZulu and were at the higher class interval, this item seemed to be quite easy in comparison to the other two languages sub-groups. The isiZulu sub-group at 0.9 and 1.2 had a 100% probability of correctly responding to the item. Based on the information contained in this figure, it would appear that the item discriminated against those learners who completed the test in either English or Afrikaans.

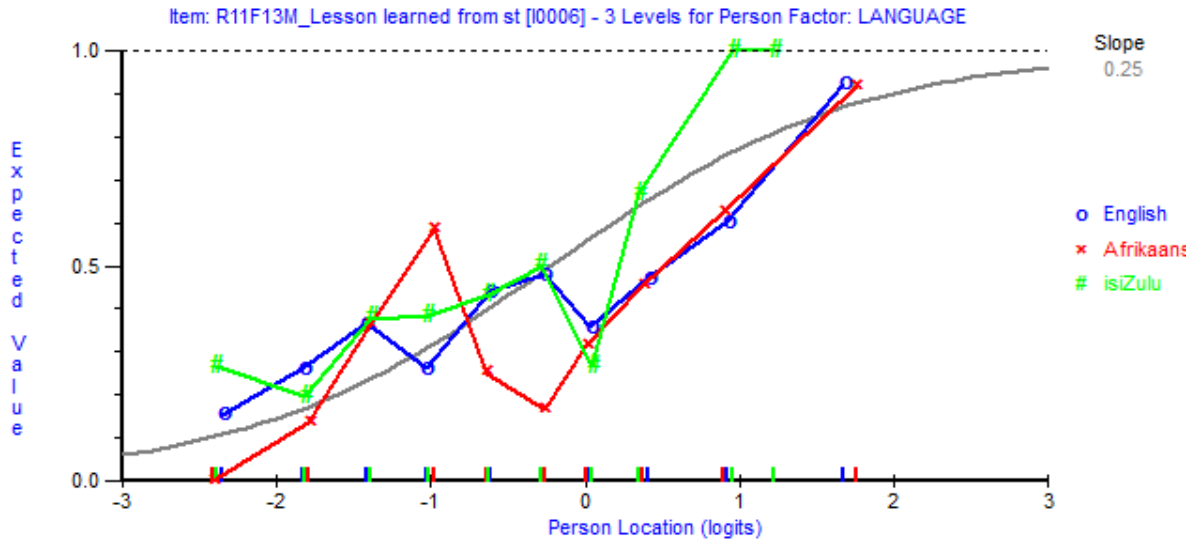

Figure 6.16: Flowers Item 13 Characteristic Curve

The distractor analysis of *Flowers Item 13* is shown in Figure 6.17. The correct response is distractor B(2).

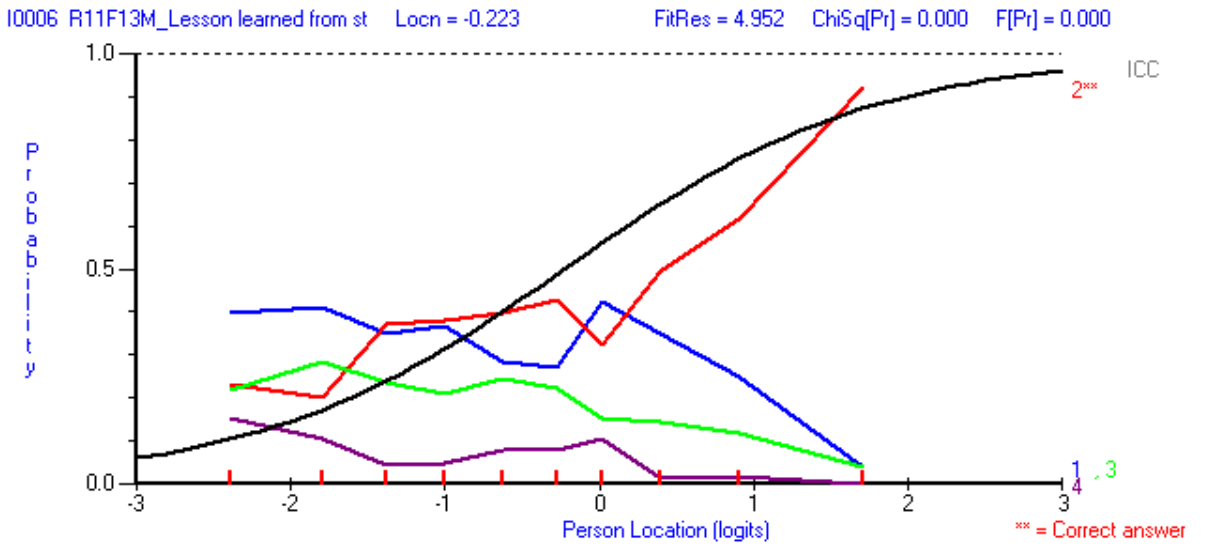

Figure 6.17: Flowers Item 13 Distractor Analysis

Figure 6.17 shows that across the lower class interval, from -3 to 0, the learners had difficulty selecting the correct answer. The chi-square is significant which provides evidence that the distribution of distractors is significantly different from the expected model. Even though the correct answer was distractor B(2), the learners had about a 40% chance of selecting it. For the same learners, distractor A(1) was very tempting and distractor C(3) also attracted some attention from the learners. However, distractor D(4) was clearly negated as a plausible answer

as very few learners seemed to have selected it. Item 13 required learners to access higher order cognitive skills to answer this question. The correct answer referred to making a new place feel like home, whereas distractor A(1) stated that old people would not be happy if they moved. This distractor may have been alluring to the lower ability persons as the text indicates that Granny Gunn was unhappy when she had to move closer to the city and away from her farm. The learners at the higher class internal, between 0 and 3, determined that the only plausible answer could be distractor B(2), as the story focused on making the necessary changes to make a new house a home.

Flowers Item 5 was the next item that displayed DIF across the three languages (as illustrated in Figure 6.18). Item 5 is a multiple choice question and entails learners having to *Make Straightforward Inferences*. The item asked the following: Granny Gunn did not like the walls and windows in her new flat. Why else was she unhappy?

a. She was ill.
b. She missed her cat.
c. She did not like the balcony.
d. She felt homesick.* (correct answer)

Item 5 from the *Flowers on the Roof* text, brings to light the anomalous ICC of the three languages. For both locations, whether at lower or upper class interval, the ICC showed considerable inconsistency for each language. It would appear that the item was exceptionally more difficult for the learners who completed the test in English compared to the other two languages – this finding is similar to the Grade 4 results (*cf.* Figure 5.23). The ICC curve of the English sub-group was just above the model curve at -2.3 where these learners had approximately 15% chance of correctly responding to the item. However, the learners who completed the test in Afrikaans, at -2.4 had the highest probability (50%) of correctly responding to the item. The isiZulu sub-group had a steady increase in the probability of getting the item right at the lower class interval (-3 to 0). At person locations 0.3 and 1.3, these learners had a 100% chance of correctly responding to the item. It would appear that this item was relatively easy for the Afrikaans and isiZulu sub-groups at the higher class interval. Furthermore, the item seems to discriminate against the learners who completed the test in English as this group found the item particularly difficult and had less probability of correctly answering the item.

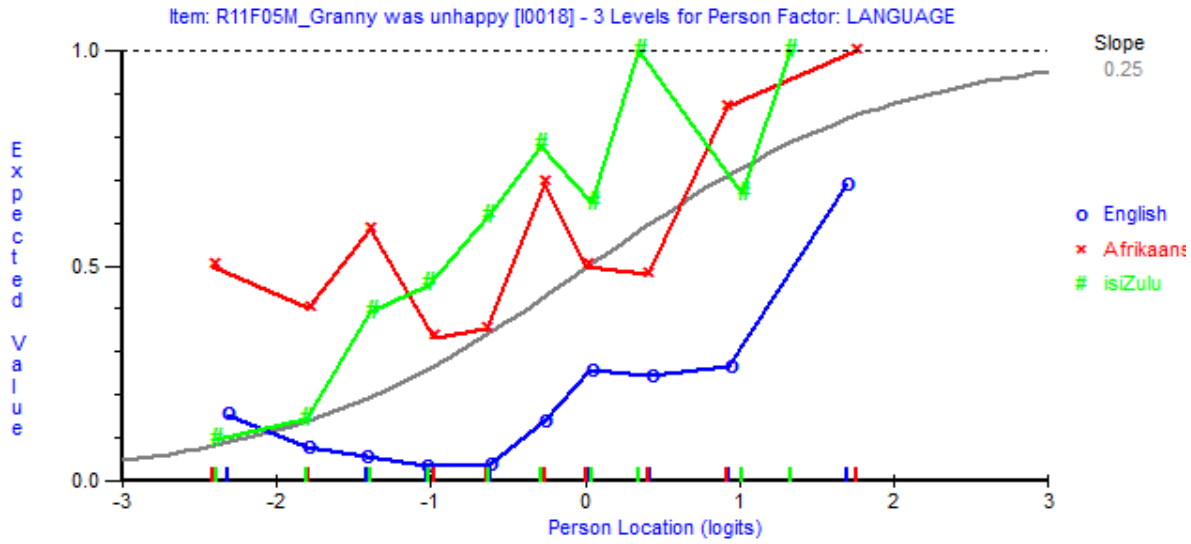

Figure 6.18: Flowers Item 5 Characteristic Curve

As this item takes the form of a multiple choice item, the next figure shows the distractor analysis result.

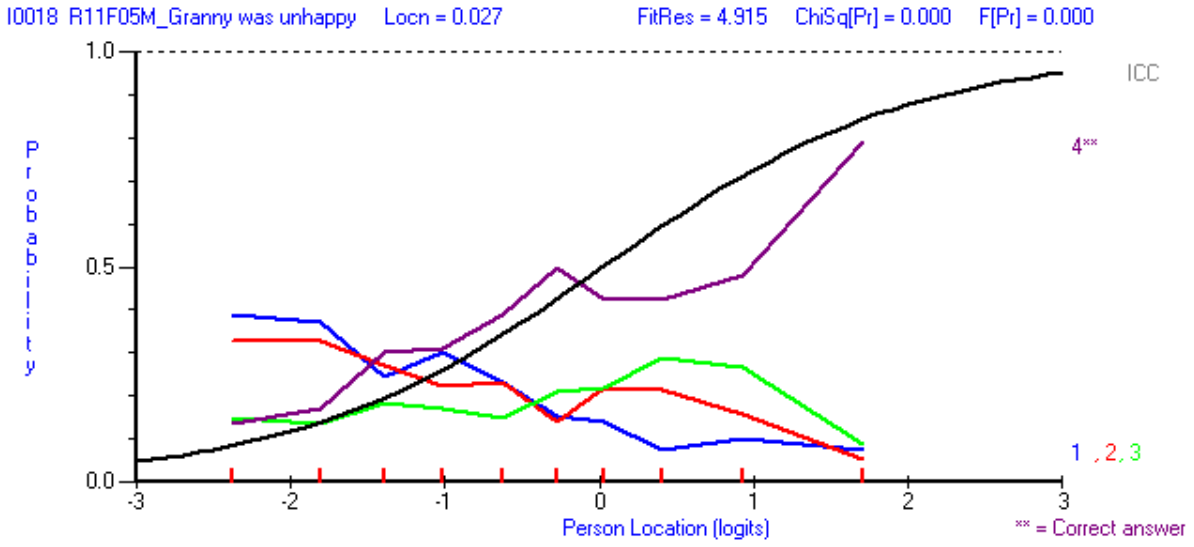

Figure 6.19: Flowers Item 5 Distractor Analysis

Figure 6.19 depicts the distractor analysis of *Flowers Item 5*. The correct answer is distractor D(4). At the lower class interval, specifically -3 to -1, the learners had difficulty in selecting the correct option. Learners at the -2.3 person location appear to have been tempted by distractor A(1) and B(2) rather than distractor C(3) or D(4). The chi-square is significant and as such, offers evidence that the distribution of distractors is statistically significantly different to the expected model. At the -2.3 person location, the learners had approximately a 10%

chance of selecting the correct distractor. Roughly at about the -1 person location, the learners seemed to eliminate distractors A(1), B(2) and C(3) as plausible answers. However, between the 0 and 1 higher class interval, an increase is observed where learners suddenly selected distractor C(3) as a plausible answer and thereafter, distractor D(4) was seen by the learners as the only plausible answer. At 1.7 person location, the learners had approximately a 80% probability of selecting distractor D(4). Distractors A(1), B(2) and C(3) reflected why she moved, one of the animals she missed and the balcony was actually something Granny Gunn enjoyed respectively. The item tested learners' ability to make straightforward inferences. The correct distractor, D(4), in English, assumes that learners know the meaning of homesick; learners had to infer from the text that she felt homesick from the way she described her new home.

Flowers Item 8 (as illustrated in Figure 6.20) shows the differential functioning across the three languages. This item is a constructed response type item with a mark allocation of one. The item requires learners to *Make Straightforward Inferences*, by asking: Find the part of the story next to this picture[37] of Granny Gunn: Why did Granny Gunn wink and grin at the little boy?

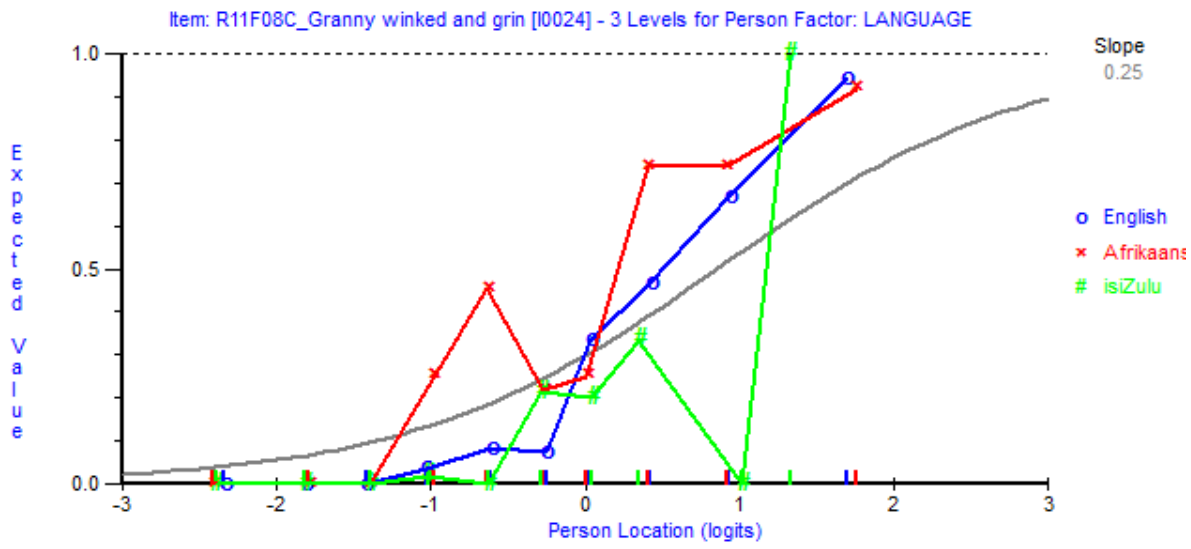

Figure 6.20: Flowers Item 8 Characteristic Curve

Figure 6.20 reveals that there is inconsistency across the languages. Neither the Afrikaans sub-group nor the isiZulu sub-group follow the model curve. All three languages between -3 and

[37] A picture of Granny Gunn winking was provided.

-1 had had a zero percent probability of answering the item correctly. This finding indicates that the item was particularly difficult for learners at the lower class interval. However, the Afrikaans sub-group had a higher probability of correctly answering this item at the -0.9 and -0.6 person locations and surpasses that of the expected model. What is more, both the English and Afrikaans sub-groups at the higher class interval (between 0.4 and 1.7) had higher chances of correctly responding to the item. Specifically, the English sub-group has a steady increase whilst the Afrikaans sub-group had a less consistent increase in their probability of correctly answering the item. In addition, the learners who completed the assessment in isiZulu at the 1 person location, had zero percent chance of answering the item correctly, whereas at the next person location (1.3), the learners had a 100% chance of responding correctly to the item. In order to obtain a mark, learners had to go back to the text where the picture of Granny Gunn winking is found, and read the sentence that follows. An acceptable answer would include learners providing an inference on why Granny Gunn winked. Learners could give an answer that demonstrates their understanding by indicating that Granny had a plan or an idea. Learners may also have given an answer that shows that Granny Gunn agrees with the boy's idea (*cf.* Appendix C). It would appear that this item discriminated against the -1 and 1 learner ability groups.

Figure 6.21 presents the DIF of *Flowers Item 10* across the English, Afrikaans and isiZulu language sub-groups. This item required that learners *Make Straightforward Inferences* from the *Flowers on the Roof* text by asking the following question: At the end of the story, how did Granny Gunn feel about her new home?

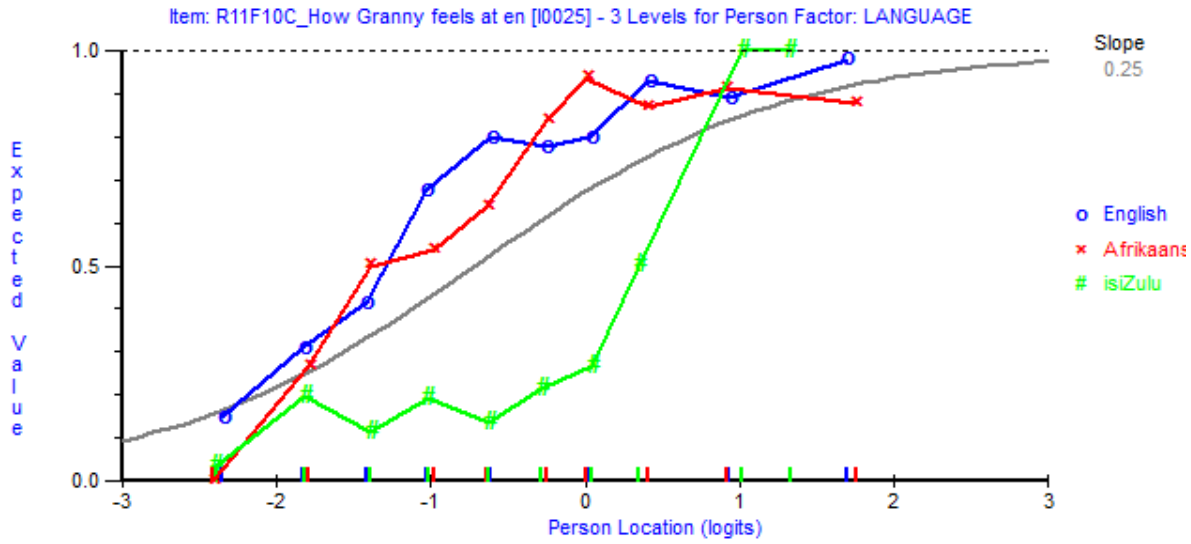

Figure 6.21: Flowers Item 10 Characteristic Curve

For *Flowers Item 10*, the learners had to provide a response (inference) regarding Granny Gunn's feelings at the end of the story for 1 mark. The model curve suggests that the item was balanced in terms of difficulty for the lower and upper class intervals. However, by looking at the ICC for the isiZulu language sub-group, the item was relatively more difficult for these learners at the lower class interval than for the English or Afrikaans sub-groups. Even though both Afrikaans and isiZulu subgroups between -3 and -2 had almost no chance of correctly responding to the item, the Afrikaans sub-group had a higher probability of answering the item correctly from that point on, whereas the isiZulu sub-group remained below the model curve. Both the English and Afrikaans sub-groups followed a similar pattern to the model curve, though the Afrikaans sub-group at 1.7 unexpectedly dropped below the model curve and had an approximate 90% chance of answering the item correctly. The isiZulu sub-group at person locations 1 and 1.3 had a 100% probability of correctly responding to the item. In order to get the one mark, the learners had to provide an answer that involves Granny Gunn's feelings at the end of the story – that she had a positive feeling toward her new home. See Appendix C for example answers for this item from the scoring guide.

The next item displaying differential functioning across the languages is *Flowers Item 7* (as illustrated in Figure 6.22). This item required learners to *Interpret and Integrate Ideas and Information* from the text in order to answer the following question: When Granny Gunn was on the balcony, she crouched down so that she could not see any of the rooftops – only the mountains and the sky. Why did she do this?

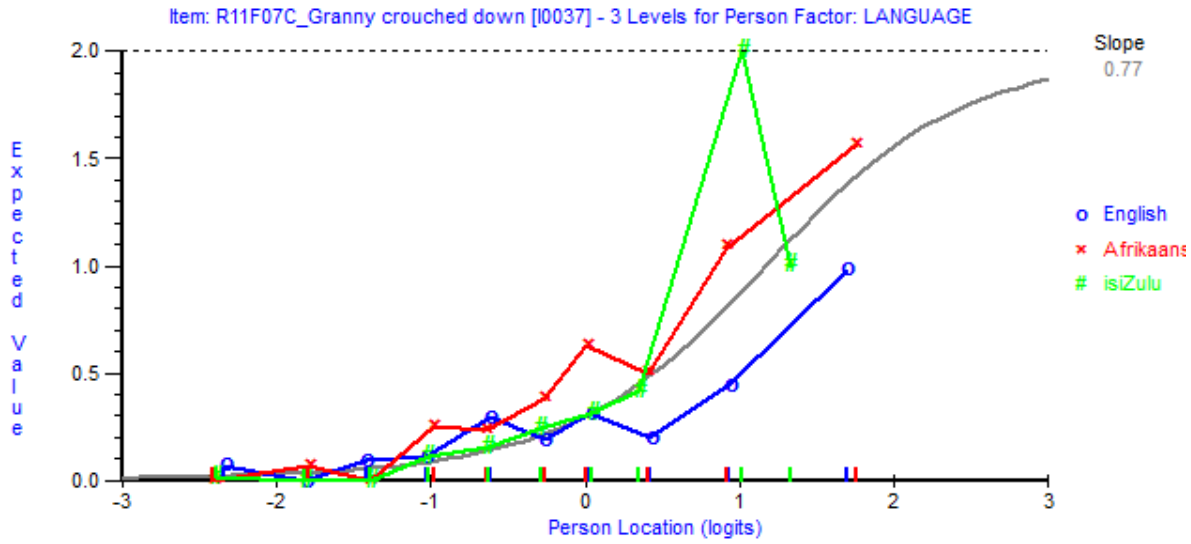

Figure 6.22: Flowers Item 7 Characteristic Curve

Flowers Item 7 is a constructed response type item and consisted of two marks. For learners to obtain full marks, they had to show complete comprehension of what they were asked in relation to the text. Learners had to provide a response that integrates ideas from the text in order to interpret Granny Gunn's feelings about the mountains and the sky. Learners could focus on Granny Gunn's view from her new home's balcony and that it reminded her of her previous home in the country, or that Granny Gunn could see the countryside from her new home. For one mark, learners had to provide a response that partially reflects Granny Gunn's feelings about the mountain and sky. Based on the graphical information provided in Figure 6.22, it would appear that this item was extremely difficult across all languages at the lower class interval. Even though these learners found the item difficult, it would seem that the language sub-groups follow roughly a similar pattern to the model curve. Yet, from .4 person location onwards, the English sub-group fell below the expected model curve indicating that for these learners, the item was more difficult than for the Afrikaans or isiZulu sub-groups. Moreover, the isiZulu sub-group at 1 had a 100% probability of obtaining the two marks. Neither of the two other languages had a 100% chance of obtaining full marks for this item. Nevertheless, the isiZulu group at 1.3 had a large decrease in obtaining the full marks compared to those learners at 1 person location. Overall, this item appears to be difficult, especially for the learners at the lower class interval and for learners who completed the test in English.

Flowers Item 9 is the next item that displayed differential functioning across the three language sub-groups. This item was pinned at the *Focus on and Retrieve Explicitly Stated Information*

of the processes of comprehension. This item asked the following question: Write two ways in which Granny Gunn made her new flat feel like home.

Figure 6.23: Flowers Item 9 Characteristic Curve

Flowers Item 9 is constructed response type item with a mark allocation of two. Learners were awarded full marks if their responses showed complete comprehension of Granny Gunn's actions to make her new flat feel like home (*cf.* Appendix C). These learners' responses had to indicate at least two actions taken by Granny Gunn to make her flat feel like home. For one mark, learners had to provide a response with only one action. The actions taken by Granny Gunn are explicitly reported in the text; learners need only to find them. Figure 6.23 displays the ICC for this item and it would appear that the English sub-group followed the expected model curve. The item was extremely difficult for all language sub-groups between the -3 and -1 person locations, with learners having a less than 50% chance of obtaining at least one mark. It would appear that learners at the lower class interval struggled with this item. The English sub-group at the higher class interval, had a higher chance of correctly answering the item correctly and it exceeds that of the model curve. As such, this item was likely easier for these learners. Furthermore, the learners who completed the test in Afrikaans seemed to have more difficulty in correctly answering this item, which implies that there is more discrimination against this language sub-group.

Figure 6.24 presents the differential functioning of item 12 from the *Flowers on the Roof* text. This item is a constructed response type item for three marks. It asked learners to *Interpret and Integrate Ideas and Information* by answering the following question: What were the little

boy's feelings about Granny Gunn when she first moved in and at the end of the story? Use what you have read to describe each feeling and explain why his feelings changed.

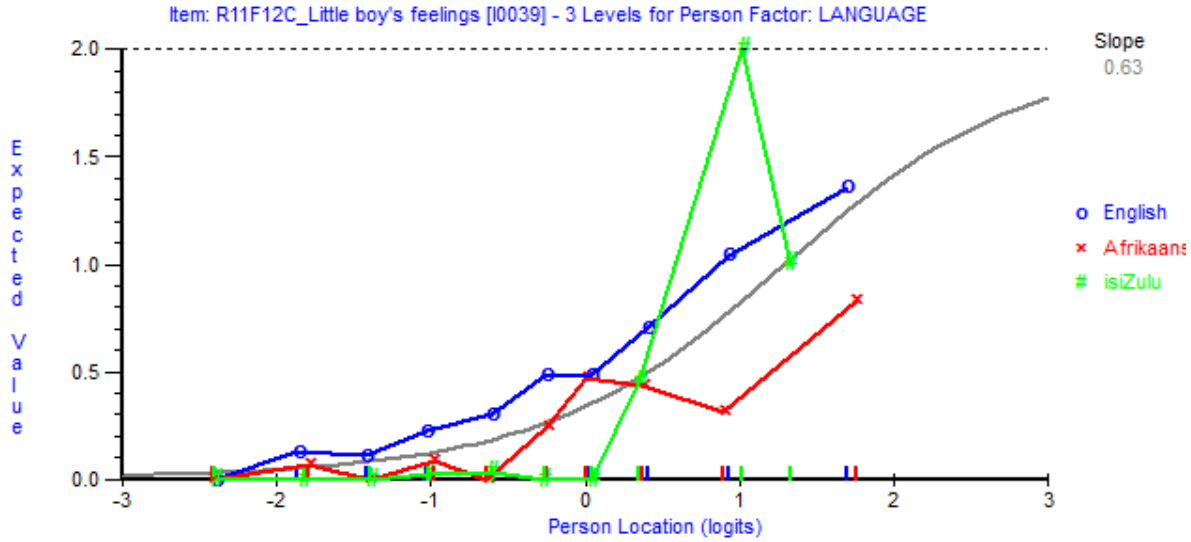

Figure 6.24: Flowers Item 12 Characteristic Curve

It should be noted that none of the learners obtained full marks for this item, as evidenced in Table 6.6. As none of the learners obtained full marks, the model discarded the three-mark parameter for a two-mark parameter. Based on Figure 6.24, it is evident that this item was extremely difficult for the learners, especially for those at the lower class interval (between -3 and 0). Some variation also occurs for the learners at the lower class interval as the English sub-group ICC were above that of the model curve and isiZulu sub-group had consistently almost zero percent probability of correctly responding to the item. Nevertheless, the isiZulu sub-group at the 1 person location had the highest probability of correctly responding to the item across the language sub-groups. These learners had a 100% probability of attaining two marks for the item. Yet, the isiZulu sub-group suddenly dropped to a 50% chance of attaining the two marks at the 1.3 person location. The English sub-group ICC remained consistent to the model curve, which indicates that there is some discrimination toward those completing the test in either Afrikaans or isiZulu, with the Afrikaans sub-group finding this item more difficult than the other two languages. The scoring guide for this item showed extensive examples and guidance for extensive comprehension assigning three out of three marks. However, as none of the three language sub-groups were able to obtain full marks, the focus is on the two mark scoring guide. In order to achieve two marks, learners had to provide a response that demonstrates their comprehension of the little boy's feelings about Granny Gunn as well as why these had changed. An accepted response would have to include a description of the little

boy's negative feelings when Granny Gunn first moved in. Thereafter, the answer could also include a change to positive feelings by the end of the story without providing a reason as to why the feelings had changed from negative to positive (*cf.* Appendix C). For one mark, the learners had to show some comprehension of the little boy's feelings by providing either the negative or positive feelings regarding Granny Gunn. It would appear that the learners across the language sub-groups were not able to infer and make interpretations from the text regarding the little boy's feelings toward Granny Gunn or how they had changed through the story.

The final section of this chapter provides a brief summary of items that displayed differential functioning across the English, Afrikaans and isiZulu subgroups for the three released PIRLS 2016 texts.

6.4 CHAPTER SUMMARY

This section provides a brief overview of the items that showed differential item functioning (DIF) across the three language. Table 6.19 presents these items in accordance with the text and item number.

Table 6.19: Summary of Problematic Items across PIRLS 2016 Released Texts

Text Title	Item	Process of Comprehension	Item Type	F-ratio	Probability
Macy and the Red Hen	Item 1	Focus on and Retrieve Explicitly Stated Information	MC	12.735	0.000*
Macy and the Red Hen	Item 4	Interpret and Integrate Ideas and Information	CR	10.656	0.000*
Macy and the Red Hen	Item 6	Focus on and Retrieve Explicitly Stated Information	CR	21.706	0.000*
Macy and the Red Hen	Item 10	Focus on and Retrieve Explicitly Stated Information	MC	18.808	0.000*
Macy and the Red Hen	Item 12	Interpret and Integrate Ideas and Information	MC	9.875	0.000*
Macy and the Red Hen	Item 13	Interpret and Integrate Ideas and Information	CR	9.602	0.000*

Text Title	Item	Process of Comprehension	Item Type	F-ratio	Probability
Macy and the Red Hen	Item 15	Interpret and Integrate Ideas and Information	CR	9.018	0.000*
The Green Sea Turtle's Journey of a Lifetime	Item 1	Make Straightforward Inferences	MC	8.251	0.000*
The Green Sea Turtle's Journey of a Lifetime	Item 10	Focus on and Retrieve Explicitly Stated Information	CR	18.525	0.000*
The Green Sea Turtle's Journey of a Lifetime	Item 11	Interpret and Integrate Ideas and Information	CR	9.567	0.000*
Flowers on the Roof	Item 5	Make Straightforward Inferences	MC	77.451	0.000**
Flowers on the Roof	Item 7	Interpret and Integrate Ideas and Information	CR	6.800	0.001**
Flowers on the Roof	Item 8	Make Straightforward Inferences	CR	38.581	0.000**
Flowers on the Roof	Item 9	Focus on and Retrieve Explicitly Stated Information	CR	24.729	0.000**
Flowers on the Roof	Item 10	Make Straightforward Inferences	CR	74.900	0.000**
Flowers on the Roof	Item 12	Interpret and Integrate Ideas and Information	CR	38.084	0.000**
Flowers on the Roof	Item 13	Evaluate and Critique Content and Textual Elements	MC	6.994	0.001**

*Significant at the 5 percent level (Bonferroni 0.000521)

**Significant at the 5 percent level (Bonferroni 0.001282)

A total of 17 items functioned differently across the three languages. Of the 17 items, six were MC and 11 were CR type questions. Figure 6.25 shows the number of DIF items per process of comprehension. These findings are somewhat different compared to the Grade 4 items, where problematic items were mostly related to the ability to focus on and retrieve explicitly

stated information. The Grade 5 learners particularly struggled to interpret and integrate items in PIRLS 2016 as well as with the 5 DIF items that aimed to test whether learners could access and retrieve explicitly stated information from the text.

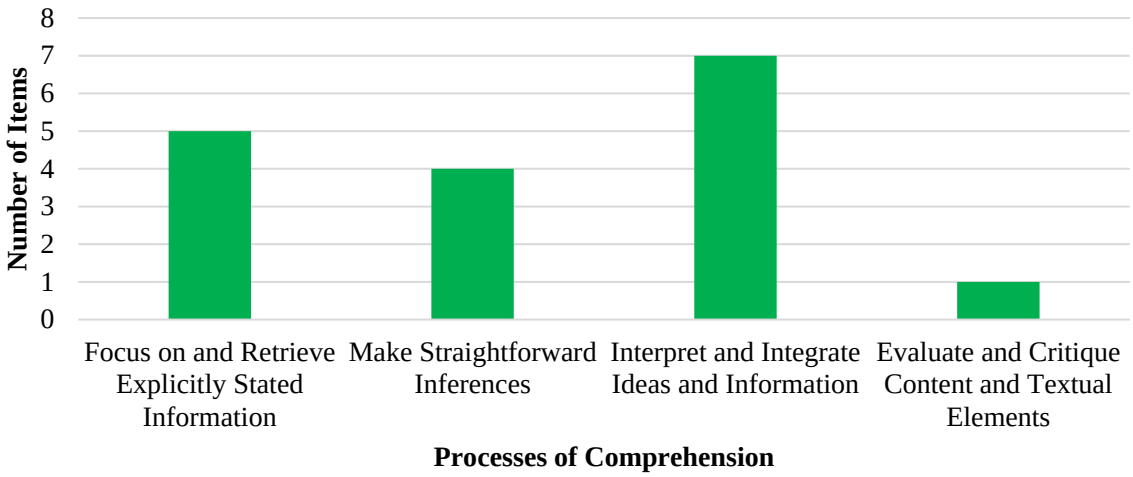

Figure 6.25: PIRLS 2016 DIF Items per Processes of Comprehension

The following table provides a summary of the items which displayed DIF in relation to the number of items per text.

Table 6.20: Summary of Problematic Items per PIRLS 2016 Released Text by Total Number of Items

Text Title	Total Number of Items	Total Number of Items Displaying DIF	Percent (%)
Macy and the Red Hen	16	7	44
The Green Sea Turtle's Journey of a Lifetime	16	3	19
Flowers on the Roof	13	7	54

Based on the evidence presented in this chapter, the following sub-question is addressed: *How does average learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016?* The core of this question looked at the differences, if any, between the three languages by using means, percentages as well as Rasch analysis.

From the evidence shown, it is clear that based on metric equivalence, the tests were not equivalent across the three languages. However, the items did not universally favour one

language nor did the texts discriminate universally against a particular language group. All the items which displayed DIF had inconsistencies across the three languages, for example, *Flowers Item 5* (Why else was Granny Gunn unhappy?) was much less difficult for the isiZulu language sub-group. *Macy Item 6* (What are two things Macy does that do not work?) was relatively easier for the learners who completed the test in Afrikaans and *Turtle Item 10* (Why does a sea turtle's body fat become green?) was easier for the English sub-group. Some items such as *Macy Item 13* (Describe what Macy is like and give two examples from the story that show this) were considerably difficult for all three language sub-groups at the lower case interval as almost none of these learners had a good chance of correctly responding to the item.

Based on the findings from the Rasch analyses, more detailed, qualitative attention is paid to the texts and their items across the different languages. These can be found in Chapter Seven and Eight of this thesis.

-⌘-

CHAPTER 7

Phase Two: Literary Texts Qualitative Data Presentation and Results

7.1 ORIENTATION

In this chapter, the findings for the qualitative Phase Two are presented. It specifically looks at the PIRLS Literacy 2016 and PIRLS 2016 literary or narrative texts. The findings in this chapter complement and extend the results from the Rasch analysis presented in Chapters Five and Six. This chapter deals with the findings for the research sub-question 3 and 4 for the study, namely:

3. *To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?*

4. *How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?*

The aim of this chapter is to present a multipronged analysis of the narrative texts which were used during the PIRLS 2016 assessments. The data presented in this chapter is a consolidation of the findings gathered from various analyses: passage mapping of each text, item difficulties, text analysis, workshop notes and questionnaires on the equivalence of the translated texts, answered by Afrikaans and isiZulu experts (*cf*. Section 7.2).

Reading literacy is associated with the different reasons why people read, such as for their own enjoyment, interest and learning. When children start to read, their main reason for doing so often focuses on reading a text that tells a story, for example a storybook, and informational texts that deals with facts about the world (Mullis & Martin, 2015). The PIRLS framework, based on the main reasons why children read, have two main purposes for reading, namely *for literary experience* and *to acquire and use information*. When children go to school, they are required to read in order to learn new things in different subjects at school (Duke & Carlisle, 2011; Spaull, 2016), meaning that the learners use their reading skill to gain new information across the curriculum (Millin, 2015) and ultimately progress to the next grade.

7.2 MULTIPRONGED ANALYSIS OF THE LITERARY TEXTS

Since this study wishes to determine whether the PIRLS Literacy 2016 and PIRLS 2016 released texts are equivalent across English, Afrikaans and isiZulu, a multipronged analysis of

the texts was required. As such, this study considered whether the PIRLS assessment was suitable and relevant to South African learners and also judged the textual coherence of the PIRLS assessments. Textual coherence is an umbrella term used within this study that refers to the 'togetherness' in a text looking at the features of each text, whether the text makes sense, as well as the logical flow of each text (*cf.* Menzel et al., 2017). In addition, textual coherence includes the mental representation of the text which could either be a *text-based* or a *situation-based* representation of what readers build up while they read (*cf.* Kintsch & Rawson, 2005; Van Dijk & Kintsch, 1983).

Based on the above, the qualitative analysis is divided into *macro*, *meso* and *micro* levels of equivalence.

Macro level

For each of the selected narrative and informational texts, an overview is presented. The overview of the narrative texts includes the identification of the genre, discussion of the explicit and abstract themes and the exploration of the main characters, which includes their traits and functions. In terms of the informational texts, the *macro* level included the main purpose of the text as well as the different organisational elements used throughout the text. At this level, the layout of the text as well as the length of the text is considered.

Meso Level

The *meso* level comprises the different narrative events as indicated by the IEA's passage mapping for each literary/narrative text. The plot of each text, which follows a problem and solution structure, is also provided. The *meso* level of the informational texts differed slightly to the narrative texts. The informational texts made use of organisational elements and headings to aid in structuring the text. In addition to the abovementioned aspects, the question types as well as the comprehension processes are presented. Permission was obtained from the IEA to acquire and make use of the PIRLS Literacy 2016 and PIRLS 2016 passage maps[38].

The four *Processes of Comprehension*, as discussed in Chapter Two, is provided for each of the narrative texts, namely focus on and retrieve explicitly stated information, make straightforward inferences, interpret and integrate ideas and information and evaluate and

[38] These were obtained for the five released texts.

critique content and textual elements (Mullis & Martin, 2015). These question types are examined per text across languages in order to indicate similarities or differences between the languages, specifically to see whether there were any changes in question type for the translated versions of the text. Such changes signal comprehension process non-equivalence. If a difference is found between the source text and the translated text, then a possible explanation is given regarding the change in processes. This undertaking was informed by asking experts to read the text and relevant items, and indicate whether changes occurred.

Micro level

This level focuses on specific aspects such as vocabulary, key expressions and rhetorical devices used in the text, average word length and words per sentence. In terms of vocabulary, consideration was given to the occurrence of low frequency words and whether Grade 4 and 5 learners would have struggled with some of the words in the text. Readability measures were used to determine the readability of the ST as these measures are not yet available for Afrikaans nor isiZulu. The average word length and words per sentence was calculated for each language. The macro, meso and micro levels of equivalence were included for each text as a guide to check the fidelity of the translated texts.

In an attempt to maintain readability, the three levels of equivalence have been incorporated under the relevant sub-headings discussed throughout the chapter. In addition to determining these levels of equivalence, I sought the assistance of Afrikaans and isiZulu language experts to provide their professional opinions on the selected texts.

Professional Opinion on Translated Texts

I served as an expert for the Afrikaans translated texts, as I am fluent in both Afrikaans and English. Moreover, my involvement with PIRLS started during the PIRLS 2011 cycle where I assisted with the test instruments. In the last cycle of PIRLS in 2016, I managed the questionnaire development and refinement and assisted with the test instruments. During PIRLS 2016, I was required to deal with translations, which included meeting the translators to discuss the translated versions and completing National Adaptation Forms (NAFs). As part of the team that conducted the study in South Africa, I also participated in scoring training with the IEA and scored Afrikaans booklets during PIRLS 2016. I was also co-author of the national PIRLS 2016 and PIRLS Literacy 2016 reports (Howie et al., 2017a) as well as the PIRLS 2016 South Africa Encyclopedia (Howie et al., 2017) that is available in the public domain.

As indicated in Chapter Four, a second researcher, who is an Afrikaans home language speaker, also assisted with the Afrikaans texts and has experience in psychometrics. For isiZulu, three scholars were asked to evaluate the isiZulu translated texts. The first isiZulu expert has substantial knowledge in the writing of isiZulu stories and teaches isiZulu to university students. The second isiZulu expert was a Grade 4 isiZulu language teacher at the time of this study while the third isiZulu expert is a professor and specialist in isiZulu linguistics. The experts provided evidence of possible non-equivalence in the translations, indicating specific items that might have functioned differently across the languages due to low frequency words being used or other aspects, such as cultural differences. Some of the examples from the experts were added with the intention of providing evidence of irregularities. Experts were asked to complete questionnaires and workshops were held in order to discuss possible translation non-equivalence.

Per-item Analysis for Each Text

Chapters Five and Six provided evidence of differential item functioning (DIF) across the selected texts. In Chapters Seven and Eight, specific attention is paid to the items that displayed non-equivalence among the different languages. For each narrative and informational text, the percentage of items correct as well as the zero percentage for each item is presented. In order to further explore the items that showed DIF, I examined a small selection[39] of learner responses[40] for each of the texts in order to further investigate possible reasons for the non-equivalence.

The next sections provide an overview of the three released literary texts used during PIRLS Literacy 2016 and PIRLS 2016, namely *The Pearl*, *Flowers on the Roof*, and *Macy and the Red Hen*. These literary texts are from the same genre, realistic fiction, and focus on the main characters who face a problem and by the end of the story find a resolution. Each of these texts take place in different settings, for instance *The Pearl* takes place at a seaside town, *Flowers on the Roof* takes place in the countryside and a flat in town and *Macy and the Red Hen* takes place at her parents' house.

[39] The Centre for Evaluation and Assessment's (CEA) data manager randomly selected the booklets.
[40] Constructed Response (CR) items only.

Orthographies of Languages

Because this study focuses on three different languages spoken in South Africa, it is important to briefly explain the orthographical differences between these languages. African languages in South Africa are agglutinative languages that belong to the same family of Southern African Bantu languages[41] (van der Merwe & Le Roux, 2014). The Nguni languages within this larger family have a rich and complex morphology (Keet & Khumalo, 2017a) with transparent, conjunctive orthographies, meaning that morphemes that occur as prefixes, infixes and suffixes to stems are written together as a single word unit, rather than as separate words (Truddell & Schroeder, 2007). The letters and words used in isiZulu have a "fairly straightforward one-to-one relationship with the sounds they represent" (van Rooy & Pretorius, 2013, p. 282). While Afrikaans and isiZulu orthographies are transparent, English has an opaque orthography, and as a result of the conjunctive orthography, isiZulu words are generally longer (they comprise multiple morphemes attached to a stem) and texts are typically shorter and denser due to increased word length. Even though the text length is shorter, it requires more time to read (Land, 2015). Moreover, as readers find the level of text difficulty increasing, their reading rate is slowed (Land, 2015).

However, this study does not specifically look into the structure of each of these languages but rather whether there are occurrences of non-equivalence amongst the three languages on the selected PIRLS Literacy 2016 and PIRLS 2016 texts and items.

Section 7.3 presents the analysis of *The Pearl*, followed by *Flowers on the Roof* (7.4) and *Macy and the Red Hen* (7.5). The same structure was applied in analysing all the texts. Each text is discussed in terms of the genre, plot, setting, theme and main characters. Furthermore, each section provides the word count, key vocabulary and number of items across the different *Processes of Comprehension*. The last section (7.6) provides the conclusion of this chapter.

7.3 ANALYSIS OF *THE PEARL*

This section provides an overview of the text and composition of *The Pearl* (7.3.1) and is followed by a summary of the different reading comprehension processes (7.3.2) that illustrates the item difficulty of each item of *The Pearl*. Thereafter, the professional opinions of the quality

[41] The Nguni group comprises isiZulu, isiXhosa, siSwati and isiNdebele.

and veracity of the translated texts and items are presented in sub-section 7.3.3. This section specifically focuses on issues that are possibly problematic. In sub-section 7.3.4 a per-item analysis of *The Pearl* is provided with examples of Grade 4 learners' responses in English, Afrikaans, and isiZulu (*cf.* Appendix D for the source text (ST) and the Afrikaans and isiZulu versions). The above structure will apply, except where otherwise indicated, for each of the texts. A total of 832[42,43] English, Afrikaans and isiZulu Grade 4 learners completed *The Pearl*.

7.3.1 Overview of The Pearl

The Pearl is a realistic fiction literary text used for the PIRLS Literacy 2016 cycle. The story is about a young boy named Reuben who wanted to become a wealthy pearl merchant after his friend, Josh, had found a pearl while they were swimming. After finding the pearl, Josh generously gave it to his friend Reuben who then spent his time learning about pearls. He later became a very successful pearl merchant. One day, on returning to his seaside town, he wanted to spoil Josh with an extravagant gift. However, Josh was not interested and suggested to Reuben that his wealth could rather be used for the benefit of the seaside town.

The setting shifts from a seaside town to a great city where pearls are bought and sold and the different locations Reuben visited during his time as a pearl merchant. The explicit theme for this story involves Reuben who has followed his dream of becoming a pearl merchant but, later in the story, his old friend reminds him that there is more to life than being wealthy. The abstract theme is that happiness does not stem from the personal accumulation of wealth but from sharing it to the benefit of others. Table 7.1 presents the characters' traits, attributes and functions in *The Pearl*.

Table 7.1: Character Description of The Pearl

Character Name	Traits or Attributes	Functions
Reuben	He is a dedicated and driven person who is also ambitious. As	Reuben shows the boundaries of money (wealth) and material success. He became a very successful pearl merchant but felt something

[42] For the PIRLS Literacy 2016 and PIRLS 2016 assessments, learners received 40 minutes to read and answer questions based on the text. See Chapter Two for more details.
[43] Because of the matrix design used for the 16 booklets (including the Reader), the learners did not all receive the same booklet with the same texts. See Chapter Two for more details.

Character Name	Traits or Attributes	Functions
	the story progresses, he becomes successful and later, thankful.	missing in his life and went back to visit his old village and friend, Josh.
Josh	He is portrayed as a kind, unmaterialistic and generous character.	Josh leads a simple life and shows that it is important to be kind, to think of the community first before yourself. Reuben wanted to bestow gifts on his old friend but Josh showed him that there are different ways of showing appreciation other than money.

There are various types of literary or rhetorical devices that can be used in a story. Examples of these include metaphor which is a figure of speech that compares two things to attribute a particular quality to the first thing, for example, 'life is a journey'. In *The Pearl* three different literary devices were used, namely suspense, imagery and dialogue.

- *Suspense* is a device used to build on certain aspects of human cognition such as expectation, anticipation and predication (Literacy Devices Editors, 2019). It gives the reader a feeling of 'waiting' for the outcome of an event, leaving the reader wanting to know what happens next. Also commonly known as a 'cliff-hanger'.
- *Imagery* is a literary device that is used when the author wants to create a 'mental image' for the reader by using words, phrases or expressions (Hall, 1994; Literary Devices Editors, 2019). Imagery is a strong device as it enables the reader to visualise the text. Imagery can be invoked by using other literary devices such as metaphors or similes. Imagery is not limited to only visual perceptions, but also other senses such as tactile and auditory senses.
- *Dialogue* is a technique used in stories or text where authors have two or more characters who speak (Literary Devices Editors, 2019). This 'conversation' makes the story more realistic, familiar and contains features of Basic Interpersonal Communicative Skills (BICS) (Cummins, 2008), which make the reading of the text easier. Dialogue can be written either in direct or indirect speech. Two types of dialogue exist, namely inner and outer dialogue. The latter is used most often in texts whereas inner dialogue refers to characters who speak to themselves or are thinking.

Table 7.2 provides examples of these literary devices across the three languages for *The Pearl*.

Table 7.2: Literary Devices for The Pearl across Languages

Literary Devices	English	Afrikaans	isiZulu
Suspense	"But who would dare to ask the one big question…"	"Maar wie sal dit waag om die een groot vraag te vra…"	"Kodwa ubani owayengalokotha abuza umbuzo obalulekile..."
Imagery	"…some silvery pale, others glowing rosy pink"	"…party silweragtig lig, ander wat rosig pienk gloei"	"...amanye ayesasiliva ngokuphaphathekile, amanye ephinki ngokucwebezelayo"
Dialogue	Between Reuben, Josh, and the other children	Tussen Reuben, Josh en die ander kinders	Ngaphakathi kuka-Reuben, u-Josh, kanye nalezi ezinye izingane

In terms of the literary devices used in this text, both translated texts, Afrikaans and isiZulu, made use of the same literary devices as presented in Table 7.2. This finding shows that even though the texts were translated from the source text (ST) (English), the languages were able to keep the specific literacy devices and not deviate from the ST. Moreover, the specific meanings conveyed by each of these literacy devices were the same across the languages.

The plot is a typical problem or conflict, followed by a resolution. The problem that occurs in *The Pearl* is that even though Reuben becomes a wealthy pearl merchant, he is generally unhappy with how his life has turned out. The resolution is that when Reuben returns home and wants to repay Josh for that very first pearl he had given to Reuben when they were children, Josh proposes a more altruistic way of using the money. A list of the major events during *The Pearl* is presented below:

- Children play together down by the sea.
- A beautiful pearl is found by Josh.
- The children gather around Josh to look at the beautiful pearl.
- Reuben asks Josh whether he can have the pearl.
- Even though all the children say the pearl belongs to Josh, Josh decides to give the pearl to Reuben.
- From that day, Reuben starts spending all his time reading and learning about pearls.

- Reuben decides that he will be a pearl merchant when he grows up.
- Reuben travels to lands where pearls are fished from the sea.
- Reuben trades smaller pearls for larger pearls.
- Reuben becomes a wealthy and successful pearl merchant.
- Reuben eventually realises that he is not happy and recalls happier days at the seaside town where he grew up.
- Reuben thought of his friend Josh and how kind he was (gave Reuben his first pearl).
- Reuben returns to the seaside town.
- Reuben sees his friend Josh playing with his own children by the sea.
- Josh and Reuben talk for hours.
- Reuben has had enough of the city (and of trading pearls).
- Reuben wants to repay Josh for his generosity so many years ago.
- Josh thanks Reuben but suggests that he should rather share his money with everyone so they can continue with their lives.
- Reuben is amazed that Josh did not want anything for himself, he then remembers that all his wealth did not make him happy.
- Reuben smiles and agrees to Josh's plan to share his wealth with everyone at the seaside town (IEA, n.d.-a).

The major events describe key aspects that happened during the story, focusing on Reuben and Josh. These major events occur in each language. Some of the vocabulary and expressions used in the text were highlighted by the IEA as key to the story. The key vocabulary and expressions are selected based on criteria set out by the IEA and are captured in Tables 7.3 and 7.5.

Table 7.3: Key Vocabulary for The Pearl across Languages

IEA Vocabulary	Frequency English	Frequency Afrikaans	Frequency isiZulu
oysters	1	1	2
eager	1	1	1
gleaming	1	2	1
shellfish	1 (oyster, shellfish, oysters)	1	1
merchant	0 (trader, traded, traders)	1	1

IEA Vocabulary	Frequency English	Frequency Afrikaans	Frequency isiZulu
wealthy [a]	2 (rich, wealthy, rich, wealth)	3 (ryk, ryk, ryk, rykdom)	1 (zicebile, ecebile, wayecebile, ingcebo enkhulu)
seaside	2	1	2
generous	1	1	1
generosity	1	1	1

[a] 'wealthy' was selected by the IEA, however as it can take different forms such as 'wealth' or a synonym such as 'rich', a decision was made to add these to the table to show how the three languages used the term 'wealthy'.

Table 7.3 indicates the key vocabulary and the frequency in which these words occurred across the three languages. These words were selected as they indicate specific emotions, actions, behaviours, descriptions and objects of *The Pearl*. Interestingly, the word '*pearl*' was not selected by the IEA as a key word. The word '*merchant*' was selected as a key word yet in the ST, '*traders*' was used. In the Afrikaans version, '*handelaar*' (merchant/trader) and in the isiZulu version, '*umthengisi*' (merchant), was used. The word '*wealthy*' was used only once in the ST, but it also included '*rich*' and '*wealth*'. In the Afrikaans version, '*ryk*' (wealthy) was used three times and included the word '*rykdom*' (wealth). The word '*ryk*' is a homonym as it has more than one meaning, has similar spelling and is pronounced the same (Merriam-Webster, 2020). For the isiZulu version, it becomes somewhat complicated to find the precise key word as the structure of the language is different to English and Afrikaans, due to the agglutinative nature of the language, key words do not occur as individual words (as in English and Afrikaans) but are embedded in complex morpho-grammatical structures. The stem in isiZulu, '*-cebile*' occurred three times in the text.

For instance, the word commonly used for oysters in isiZulu is '*kokhwathu*'; however, when oysters were first introduced in the text a different word was used. The sentence read as follows: *Wafunda indlela amapharele akhula ngayo ngaphakathi kwembaza (oyster), uhlobo lwembaza (shellfish) ehlala olwandle* (He learned how pearls grow inside oysters, a kind of shellfish that lives in the sea). The word was used in a locative construction '*kwembaza*'; it consists of the root word, '*mbaza*'; and the locative prefix '*kwe'* which means near or next to an object or place (Wilkes & Nkosi, 2010). Because of the nature of the structure of isiZulu, it is important

to first find the root word as the word could look very different depending on the prefixes or subject concords used.

LexTutor VocabProfile software tool was used to analyse the vocabulary used in a corpus and in Table 7.4, the English word profile of *The Pearl* is presented.

Table 7.4: Word Profile of The Pearl

Word frequency level (English) [a]	Occurrence	% of Tokens	Cumulative Tokens (%)	Examples
K1	481	89.2	89.2	children, sea, reading
K2	18	3.3	92.5	dive, shell, trade
K3	8	1.5	94.0	generous, nets, ocean
K4	14	2.6	96.6	pearl, pearls, gleaming
K6	2	0.4	97	oyster, oysters
Off list	0	0	≈100	

[a] K refers to word frequency level per thousand words, with K1 being the 1,000 most frequently used words in English, K2 being the next 1,000 most frequent words, and so on. High frequency words range from K1 to K3; Mid frequency words range between the 3,000 and 9,000 range (K3 to K9); Low frequency words are beyond the K9 (*cf.* Schmitt & Schmitt, 2014; Stoffelsma, 2019).

The word frequency level focuses on the lexical threshold, which means that the percentage of words used in the text that must be known by the reader in order for them to understand the text (Laufer, 1989). The lexical threshold is divided into two aspects, namely lexical coverage and receptive vocabulary knowledge. The former is the number of words assumed to be known to the learner in the text whereas the latter is the vocabulary knowledge the learner needs in order to attain the lexical coverage (Masrai, 2019). If, for example, at least 95% of the words used in the text falls within the K1 level, it is fairly easy to read by beginner readers. When at least 95% of the words fall between then K1 and K3 levels, in other words between the 1 000- and 3 000-word frequency levels, the words should be known by Grade 4 learners. In other words, the number of K-levels needed to reach 95% could serve as an indicator of the difficulty of the text (Laufer, 2013; Laufer & Ravenhorst–Kalovski, 2010).

The word profile of *The Pearl* confirms research done on vocabulary and its association with reading comprehension (*cf.* Laufer, 1989; Nguyen & Nation, 2011; Schmitt, Cobb, Horst & Schmitt, 2017). Even though South African Grade 4 learners might not know a lot about pearls,

the word '*pearl*'/'*pearls*' is repeated 12 times in the ST text (13 times in Afrikaans and 14 times in isiZulu), which should give enough exposure for them to be able to learn the word. As *The Pearl* is a literary text, it contains rich vocabulary and thus includes words from the mid frequency levels. *The Pearl* balances the lower, middle and high frequency words well, as most of the words contained in this text fall within the K1 to K3 frequency ranges.

The next table presents key expressions used in *The Pearl*, as identified by the IEA.

Table 7.5: Key Expressions for The Pearl across Languages

English	Afrikaans	isiZulu
"And so he was"	"En hy word toe een"	"Nangempela waba nguye"
"…all carry on enjoying our lives"	"...almal voortgaan om ons lewens te geniet"	"...sonke singaqhubeka sijabulela ukuphila kwethu"

The two key expressions as identified by the IEA (n.d.-a) were found in both the translated versions of the ST. The expressions did not change or lose their meaning during the versioning of the English expressions.

The next table unpacks the text per language and presents the readability analysis of *The Pearl* for each of the selected languages, which includes the Flesch reading ease (RE) and grade level for English texts. The Flesch RE index is based on sources used in the American schooling context and was developed to measure the difficulty of a text. The higher the RE index, the easier the text is purported to be, thus it serves as an indicator of the text difficulty (Eleyan, Othman & Eleyan, 2020; Jindal & MacDermid, 2017).

Table 7.6: Readability of The Pearl per Language

Readability Features	English	Afrikaans	isiZulu
No of words	528	539	361
Pages/paragraphs	6	6	6
Sentences per paragraph	8.6	8.3	8.5
Words per sentence	10.1	10.7	7.0
Mean word length	4.0	3.8	7.3
Passives [a]	1.9%		
Flesch reading ease [b]	92.5		

Readability Features	English	Afrikaans	isiZulu
Grade level [c]	2.8		

[a, b, c] These features can only be calculated for English.

The word count was included to give an overview of the length of the text across the languages. As expected, the English and Afrikaans texts were closer in terms of word count compared to isiZulu. In their work, Forrin, Risko and Smilek (2019) found that if a text is lengthy, the test takers may experience 'mind-wandering'. Mind-wandering occurs when the test takers' attention is not focused on the reading text but rather on unrelated things (Thomson, Besner & Smilek, 2013), meaning that mind-wandering happens when the reader's attention is not captured by the text. It should be noted that Forrin et al. (2019) made use of Canadian university students as the participants in their study.

Furthermore, Hasbrouck and Tindal (2017) have done extensive research on oral reading fluency to establish fluency norms in English in terms of number of words correct per minute (WCPM). According to these authors, Grade 4 learners at the 50th percentile should be able to read 112 WCPM by the middle of the year and 123 WCPM by the end of the year. This finding suggests that average Grade 4 English learners should in principle be able to read a narrative text of 536 words fairly comfortably within 5 to 6 minutes. However, because *The Pearl* is divided into smaller sections[44] followed by a few questions, it may take longer as the learners would have to read a section of the text and answer the questions based on that section before continuing with the next part of the text (*cf.* Appendix D).

The number of sentences per paragraph was similar between the three languages. However, the number of words per sentence differed – the isiZulu text had an average of 7.0 words per sentence compared to the 10.7 and 10.1 for the Afrikaans and English texts, respectively. Another difference observed is based on the mean word length. The isiZulu (7.3) text had a higher mean word length compared to English (4.0) and Afrikaans (3.8). The differences observed for the isiZulu text reflect the conjunctive orthography. isiZulu texts typically have

[44] For the PIRLS Literacy 2016 and PIRLS 2016 assessments, learners received 40 minutes to read and answer questions based on the text. See Chapter Two for more details.

longer words and lower word counts because of their conjunctive orthography, where a whole sentence can consist of one orthographic word.

7.3.2 Processes of Comprehension of The Pearl Items

This section provides the item difficulty for *The Pearl* across English, Afrikaans and isiZulu and partially provides evidence of whether non-equivalence is present in the text. *The Pearl* has a total of 15 question items, of which most items only count one point. The maximum number of points for *The Pearl* is 18. Of the 15 items, eight items were constructed response (CR) and seven were multiple choice (MC) items. The CR items account for 11 points whilst MC items only covered seven points. Table 7.7 presents the item difficulty of *The Pearl* per language by looking at the four *Processes of Comprehension*.

Table 7.7: Item Difficulty of The Pearl across Languages

Processes of Comprehension	English Items	Afrikaans Items	isiZulu Items
Focus on and Retrieve Explicitly Stated Information	3, 4, 6, 7, 8, 10, 12, 13, 14	3, 4, 6, 7, 8, 10, 12, 13, 14	3, 4, 6, 7, 8, 10, 12, 13, 14
Make Straightforward Inferences	1, 2, 5, 11	1, 2, 5, 11	1, 2, 5, 11
Interpret and Integrate Ideas and Information	9, 15	9, 15	9, 15
Evaluate and Critique Content and Textual Elements	N/A	N/A	N/A

The majority of items (9) focused on learners' ability to *Focus on and Retrieve Explicitly Stated Information*. Of the remaining six items, four items were pitched at *Make Straightforward Inferences* and two items tested how well learners were able to *Interpret and Integrate Ideas and Information*. The text did not include any items at the higher order process of *Evaluate and Critique Content and Textual Elements.* As it is a PIRLS Literacy text, a maximum of 25% of the items have to be at the two higher order processes, whereas 50% of the items may be at the *Focus on and Retrieve Explicitly Stated Information* process level (Mullis & Martin, 2015). Based on the information in Table 7.6, it would appear that across the three languages, all the items remained at the same comprehension level. This finding indicates that even though the test items were translated, none of the items' difficulty level had changed.

7.3.3 Professional Opinion on Translated Texts

For both Afrikaans and isiZulu, the incidence of low frequency words is low. However, isiZulu expert 2 (Z2) did make special mention of the following words that are not commonly used: '*ipharele*' (pearl); '*zamngunga*' (encircled/surround); '*zazimagange*' (were eager); '*emagobolondiweni*' (in shells) and '*ezimbaza*' (oyster). Words such a '*pearl(s)*', '*oyster(s)*' and possibly *'shellfish*' may also be considered low frequency words for the other two languages. The K-levels[45] indicate that in English *pearl(s)* is in the word frequency range of 4 000 (K4) whilst *oyster(s)* is in K6. South African learners, or indeed learners from across the word, may not be familiar with these words. Afrikaans expert 2 (A2) explained that the word '*drom'* (*saam*), '*oester(s)'* and '*pêrel*' are low frequency words that are not commonly used in everyday Afrikaans – this matches the K-levels of the English corpus. A2 further explained that "dit is nie dat dit 'kompleks' geskryf is nie, dit is eerder ongewoon geformuleer [sic]" (the writing is not complicated, rather, it is unusually formulated). The word '*drom*' is a homonym, could mean either '*drum*' or '*gathering round*', in the text it refers to the latter which could be unfamiliar to learners. In terms of the complexity of the language used across the text, the Afrikaans was written in a simple and straightforward style and easy to understand. isiZulu experts Z1 and Z2 differ in their opinion concerning the complexity of the isiZulu version of *The Pearl*. Z2 explained that because of the "words that are not commonly used in the language… especially in children stories can be confusing to young learners. Simple words used in everyday Zulu can be used to instead [sic]". Even though A2 and Z2 identified and explained why certain low frequency words may be difficult to understand for Grade 4 learners, it is part of the reading process that learners be exposed to a wider range of vocabulary, especially in a narrative text. In so doing, learners are exposed to new words and integrate these words in their vocabulary. *The Pearl* has a low incidence of low frequency words which means that learners should be able to understand the text.

In terms of cultural equivalence, the translated versions of the ST should be equally interpreted and understood by the English, Afrikaans and isiZulu learners. Z2 commented that "… it discusses something that is real and can be associated with their everyday life". However, A2 mentioned that within the South African context, *The Pearl* item 2 could be differently

[45] Based on the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) lists.

interpreted as the learners' answers may be influenced by their frame of reference, experience or culture given that South Africa is a multilingual and multicultural country.

Regarding the construct relevance of *The Pearl*, the Afrikaans and isiZulu versions are viewed as being relevant as South African Grade 4 learners should be able to cope with the content and theme of the story. The learners should also be familiar with the type of text (genre). Z1 and A1 agree that the Grade 4 learners should be able to relate to the setting and events that occurred during the story. However, Z2 explained that because of the low frequency words mentioned earlier, such as '*ipharele'*, the English version of the word could have been included in brackets so that the learners understand that they are going to read about pearls in this text. Yet, the word '*ipharele'* has been added to an isiZulu dictionary (Dent & Nyembezi, 2003) which means that it is a word that is used by isiZulu-speakers. Although '*ipharele*' has been added to the dictionary, most isiZulu learners might not be familiar with the word. Both A1 and A2 noted that some of the pictures in the text could have been better illustrated to clearly indicate what is happening in the story. A2 felt that "die illustrasie is van goeie gehalte. Die eerste illustrasie is verwarrend, aangesien die teks dui op 'dieper water duik' en die prentjie beeld 'n seun uit in vlak water sonder enige duiktoerusting [sic]" (the illustrations are of good quality. The first illustration is confusing as the text refers to 'diving into deeper water' but the picture presents a boy in shallow water without any diving equipment).

Most of the experts agreed that *The Pearl* looks like a typical literary story that Grade 4 learners read, with Z1 adding that the story's "background information also related to South Africa". Z2 on the other hand explained that she would rather use this text for Grade 5 learners as "Grade 4 learners are in the first year of Intermediate phase and are used to reading simpler and shorter stories". The national curriculum in South Africa requires Grade 4 learners to be able to cope with 150 to 200 words to test reading comprehension (DBE, 2011).

The PIRLS Literacy texts are divided into sections with answers. The layout of *The Pearl* is similar across English, Afrikaans and isiZulu – the story consists of six pages of text and six pages with questions. The text and question items make use of a large, easy to read font as well as large spacing to ensure that the layout of the story is not cramped. Each page of the story contains specific pieces of text that are found at the same places across the ST and translated texts (TT). Furthermore, the visuals included for the story accompany a specific piece of the text. This finding adds to the face validity of *The Pearl*. As the layout of the texts across the

languages is identical, it can be argued that it reduces extraneous[46] features of assessment across languages – meaning that the relevant information for answering a question is found on the same page irrespective of the text language.

According to the experts, the *Processes of Comprehension* for each question for the Afrikaans and isiZulu versions of the text did not change from the ST to the translated version. Z2 did mention that item 6 asks "Where does Reuben go when he leaves", the isiZulu version asks "Uyaphi u-Reuben lapho eshiya indawo yakubo?" (Where Reuben goes when he leaves his place?). The Afrikaans version asks "Waar gaan Ruben heen wanneer hy weggaan?" (Where does Ruben go when he leaves?). This item shows that the isiZulu version of the item included additional information whereas neither the ST nor the Afrikaans version did; for example, the noun '*indawo*' (place/locality/hometown) and the possessive pronoun '*yakubo*' (from his) were included. Notably, this item did not display non-equivalence even though the experts flagged it. A2 mentioned "die 5 boublokke vir lees sluit in klankleer (foneties), fonemiese bewustheid, woordeskat, leesbegrip, en vlotheid" (the 5 building blocks for reading includes sounds, phoneme awareness, vocabulary, comprehension and fluency) and that *The Pearl* does not test all five these building blocks as it is a reading comprehension assessment. A2 suggested that *The Pearl* only partially tested the four processes as the Afrikaans version contained words not often used in everyday discussions. However, the same argument can be made for English and isiZulu. Overall, the majority of the experts agreed that the question items tested what they intended to measure, and as Z2 explains, "questions intend to test comprehension" and "comprehension of the text was assessed".

7.3.4 Per-item Analysis for The Pearl

This sub-section provides a summary of the items including the percentage of incorrect responses per language, zero percentage per item, items that showed differential functioning across Afrikaans, English and isiZulu, as well as a sample of learner responses for each item displaying differential functioning. A summary of the DIF items is provided in Table 7.8, which presents the total number of learners who had incorrectly answered each of *The Pearl's* items.

[46] Extemporaneous features such as self-corrections, pauses and false starts.

Special attention is given to those items that displayed non-equivalence amongst the three languages. These include items 2, 5 and 14, all of which are part of the lower order comprehension processes. Item 2 is a multiple choice (MC) item, worth one mark, that tested learners' abilities to *Make Straightforward Inferences* – the highest percent of incorrect responses is isiZulu (69%), followed by English (40%) and Afrikaans (39%). Furthermore, although the Afrikaans group (6%) had the highest percent missing data, most learners attempted to answer this item.

Item 5 required learners to write an answer for two marks and tested whether learners could *Make Straightforward Inferences*. The isiZulu learners performed the poorest on item 5 (79%) followed by English (62%) and Afrikaans (60%). However, as this is a partial credit item, it is important to look at those with a partial credit, meaning the percentage of learners who obtained one mark out of two.

Approximately one-third (31%) of the Afrikaans group obtained one mark for item 5, with 23% and 20% of English and isiZulu learners, respectively, also obtaining one mark. The missing data is relatively low as both Afrikaans and isiZulu had three percent missing values, followed by English (2%). This item was particularly more difficult for the isiZulu learners compared to the other two languages. Note that each item is discussed according to the order used in Table 7.8 (following page).

Table 7.8: The Pearl: Percentage of Grade 4 Learners who Incorrectly Answered Items per Language

Item No	Processes of Comprehension	English			Afrikaans			isiZulu		
		N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing
3	Focus and Retrieve Explicitly Stated Information	347	43	0,3	198	42	2	291	55	2
4	Focus and Retrieve Explicitly Stated Information	347	44	1	198	45	1	291	53	2
6	Focus and Retrieve Explicitly Stated Information	347	39	2	198	47	3	291	50	2
7	Focus and Retrieve Explicitly Stated Information	347	36	1	198	42	5	291	53	4
8	Focus and Retrieve Explicitly Stated Information	347	48	1	198	45	5	291	52	5
10	Focus and Retrieve Explicitly Stated Information	347	41	3	198	40	9	291	44	7
12	Focus and Retrieve Explicitly Stated Information	347	36	3	198	45	9	291	48	8
13	Focus and Retrieve Explicitly Stated Information	347	37	7	198	45	9	291	56	9
14*	***Focus and Retrieve Explicitly Stated Information***	***347***	***25***	***2***	***198***	***32***	***9***	***291***	***49***	***11***
1	Make Straightforward Inferences	347	32	1	198	30	5	291	52	1
2*	***Make Straightforward Inferences***	***347***	***40***	***1***	***198***	***39***	***6***	***291***	***69***	***3***
5*	***Make Straightforward Inferences***	***347***	***62***	***2***	***198***	***60***	***3***	***291***	***79***	***3***
11	Make Straightforward Inferences	347	75	6	198	69	11	291	80	7
9	Interpret and Integrate Ideas and Information	347	67	4	198	63	4	291	78	5
15	Interpret and Integrate Ideas and Information	347	78	3	198	79	7	291	87	9

*Indicates items that displayed DIF.

Item 14 tested whether learners could *Focus on and Retrieve Explicitly Stated Information*. It was a MC type item with a mark allocation of one. The highest percent of incorrect answers were from the isiZulu (49%) group whereas the Afrikaans group had 32% incorrect responses with English (25%) having the lowest percent of incorrect responses. This item was easier for the learners who completed the test in English. There is a notable difference in missing data amongst the three languages – isiZulu (11%) had the highest missing data, shortly followed by Afrikaans (9%). In contrast, the English group only had a missing value of 2% indicating that these learners more often than not attempted to complete item 14. Even though this is a literal item, Afrikaans and isiZulu learners may by now have experienced test fatigue, or found the item too demanding to complete.

These items, to the best of the experts' knowledge, were accurately translated without losing any crucial information even though some of the panellists felt that some items could have been phrased differently to enhance the learners' understanding of the item. These differing viewpoints may reflect different stylistic preferences rather than indicating gross translation errors in the texts.

In order to further unpack these items, a selection[47] of learner responses[48] was captured to further explore possible reasons for the non-equivalence. However, if the item was a MC type item, the entire sample's responses were used. The Rasch analysis of *The Pearl* is presented in Chapter Five of this thesis. To recapitulate, *The Pearl* item 2 asks the following question:

Why are the children all eager to touch the pearl?

a. They want to take it away.
b. They think it is special.* (correct answer)
c. They think the boy will drop it.
d. They do not believe it is real.

Table 7.9 presents the number of persons who selected each of the distractors.

[47] The Centre for Evaluation and Assessment's (CEA) data manager randomly selected the booklets.
[48] Constructed Response (CR) items only.

Table 7.9: The Pearl Item 2: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	31	9	24	12	71	24
B*	205	59	109	55	79	27
C	53	15	32	16	42	14
D	54	16	22	11	89	31
9 (not attempted)	4	1	11	6	10	3
Total	347	100	198	100	291	100

*Correct response (distractor B)

Most of the learners (59%) who completed the test in English selected the correct option (distractor B), closely followed by the Afrikaans learners (55%). However, only slightly more than one-quarter (27%) of the isiZulu learners selected B. Even though fewer isiZulu learners were able to select the correct distractor, the information in the text prompts an answer for question item 2, as it explicitly states "*ngoba zazibona ukuthi yinhle futhi iyabenyezela*" ('because they saw that it was beautiful and shiny'). Therefore, there is sufficient text evidence to infer that the pearl is something special and the picture supports this interpretation. Learners had the opportunity to review the provided answers and refer back to the first page of the text to infer that the pearl is something special. Moreover, this finding suggests that distractors A (24%) and D (31%) were too tempting for the isiZulu group. As isiZulu expert 2 (Z2) indicated, the word '*ipharele*' may have been unknown to learners as they probably had not seen a pearl or a picture of one – in other words, these learners may not have had sufficient prior knowledge about pearls or how special they are. Yet, learners should be able to cope with the text as it provides them with information about the pearl (that is, what it looks like, where it can be found, trade value) to assist them in building a rudimentary schema on the topic that they may initially have known little about. Z3 noted "an illustration of a pearl being taken out of the shell would have helped a lot to get a clear picture of what it is". A more skilled reader may have expectations that they can learn something new from the text, and would use the text to find the answers, whereas unskilled readers may make use of their own background knowledge to answer questions and possibly overlook the evidence contained in the text.

The next item that showed differential functioning between the languages, is item 5. To reiterate, *The Pearl* item 5 asks the following question:

What does Reuben do differently after he gets the pearl?

Item 5 of *The Pearl* required the learners to *Make Straightforward Inferences* based on the information in the text. As this item is a constructed response type item, learners had to list two things that Reuben did differently after he received the pearl in order to be awarded the maximum point allocation of two. Possible correct responses include that Reuben no longer plays with his friends, he reads about pearls, for his birthday he asks for pearls or states that he wants to become a pearl merchant[49]. There are at least five things that he did differently (*cf.* Appendix C). These things are present in the text but are spread across two paragraphs and the learner had to identify them as 'being different'. If the learner only provided one correct response, then that learner only received a partially correct score (one point).

Table 7.10 provides a selection of ten Grade 4 learners' responses to *The Pearl* item 5. During the PIRLS scoring training, scorers were tasked to not focus on spelling or grammar. Scorers had to read the learners' answers, and based on what the answer says, award a point. See for example English learner 10's answer. Despite spelling mistakes, the learner wrote two reasons why Reuben acted differently after he received his first pearl and obtained two points. The colours of learner responses indicate the correctness of each answer. Learner responses in green indicate a correct answer, dark orange are partially correct answers, red are incorrect answers and purple indicates nonsensical answers.

[49] *cf.* Appendix C for scoring guide.

Table 7.10: The Pearl Item 5: Grade 4 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	outdoors, whildren he stayed inside the played outdoors, he stayed inside	Not attempted	Ingoba wali thala embi shi Llihle futhi bobefuna uku lithsa
2	Not attempted	Hy leer hoe perels na Hy gaan endag 'n perelhandelaar	Wahamba wayohlala indiniwa funda amyalelo wampha abangibibakhe babona enga saphamelinga phande
3	Not attempted	Ruben het van die gehow En dit was mooi.	ukufunda incwadi nokukhuluma nabahgahi bakhe
4	Can I have it, Reaben. It really belongs to Josh.	Hy was gullukig. Hy het gese Ruben kan ma die ding kry	wayehlala endini akunde ukukhula ngayo ngaphankthi
5	Shellfish that lives in the sea. Present.	Leer oor perels Hy how baie we van a perels	lapho zona zidlala ngapnandle afunda indlela amapherele
6	He studied and reading about pearls. He wants to be a pearl trader	Hy het in die huis gebly as die ander kinders biute is. En as sy ours hom vra wat wil hy he vra hy altyd vir 'n pêrel	Wafunda ngalo izinto eziningi Umdeni wakubo umbuza angathanda umuphe sipho sini njalo nje wayecela ipharele
7	While hes friends went to play he stayed inside. Now when hes birthday comes he always askeds for pearls.	Van daardie dag af sien die ander kinders minder van Ruben. Vra hy altyd vir 'n pêrel	kwembaza oyster lwembaza shellfish

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
8	he learned how pearls grow inside oysters. A kind of shellfish that live in the sea	Hyadetikgetgetget Kgetgetegeye Hkaetge	Oyster Shellfish
9	Whe same si thesa llily? Ruamgil to wamesrimd?	Dit watter soont geskin n perel Ek gaan eenda n	ukutshengisa abantu ukuthi athi lehlukile kakhulu
10	While they played outdoors t he stayed inside reading about pearls. A kind of shellfish that lives in the sea. H learned how pearls grow inside oysatens.	Ruben bly binne en lees oor pêrels na. Hy leer hoe perels groei in oesters n soort skulpvis wat in die see leef.	Oyster Shellfish

Based on the learner responses in Table 7.10, several learners gave incomplete answers, for example, Afrikaans learner 1 who left the answer line empty. Learners also wrote incomplete sentences, such as English learner 5 who only mentioned *present* as an answer. Some learners, such as Afrikaans learner 8 and English learner 9, wrote nonsensical responses where they only copied letters randomly.

There may be several reasons why learners write such answers, as captured in Table 7.10: learners did not have basic reading and writing skills, learners were not interested in taking the test and so did not bother to answer appropriately, learners may not have understood the question and learners may have experienced mind-wandering and only scribbled letters. Learners may have scribbled letters to appear busy and possibly to hide their lack of reading literacy skill from their peers or invigilator during the test time.

Two isiZulu learners wrote *Oyster* and *Shellfish*. This finding possibly means that these learners did not fully understand what the item requested of them as they took the English words from the text as their answers. The remainder of the responses that obtained a zero point, did not provide sufficient reasons as to what Reuben did differently after he received the pearl. The next table provides the breakdown of the partial credit item per language.

Table 7.11: The Pearl Item 5: Partial Credit Breakdown

Points	English Persons	% Obtained	Afrikaans Persons	% Obtained	isiZulu Persons	% Obtained
0	135	39	58	29	169	58
1	81	23	61	31	59	20
2	124	36	74	37	54	19
9 (not attempted)	7	2	5	3	8	3
Total	347	100	198	100	290	100

Approximately one third (29%) of the learners who completed the test in Afrikaans received a zero, whereas 39% of English learners and more than half (58%) of the isiZulu learners received a zero. Very few learners across the three languages did not complete the item. This means that almost all the learners tried to answer the question. Just over one third of Afrikaans (37%) and English (36%) obtained two points, while only 19% of isiZulu learners received two points. Based on this finding, and the selection of learner responses captured in Table 7.10, it

would appear that besides poor comprehension, there may be several other reasons for poor performance; for example, learners do not read questions with sufficient precision, they fail to adjust answers to mark allocation or to the space provided for answers, they are not used to providing longer answers, they are not used to reading texts of this length and so do not have sufficient test-taking stamina and then experience test fatigue.

The last item of *The Pearl* that displayed DIF was item 14 and required learners to *Focus on and Retrieve Explicitly Stated Information*. It asks the following question:

What does Josh say they should do with Reuben's money?

a. get a new house
b. buy lots of pearls
c. share it with everyone* (correct answer)
d. take it back to the city

Table 7.12 presents the number of persons who selected each of the distractors.

Table 7.12: The Pearl Item 14: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	46	13	39	20	76	26
B	23	7	15	8	45	15
C*	251	72	118	60	114	39
D	19	5	9	5	23	8
9 (not attempted)	8	2	17	9	33	11
Total	347	100	198	100	291	100

*Correct response (distractor C)

Almost three-quarters (72%) of the learners who completed the test in English selected correct distractor C whereas 60% and 39% of the Afrikaans and isiZulu learners, respectively, selected the correct option. This is a literal type item as it required learners to retrieve information which was explicitly stated in the text.

Based on Table 7.12, it would appear that both the Afrikaans and isiZulu learners found distractor A tempting – in the text Reuben asks Josh what he would like in return for his generosity, and specifically asked whether Josh would like a new house. The question involves perspective taking, *what does X think*; as well as recursion[50], *where X said that… Y*. It could be that learners do not understand who is saying what during the text. The answer was stated in a previous section of the text and not on the same page where the correct answer could be found. This finding may also suggest that learners accessed previously given text information in their memory (someone said something about a house) rather than finding the correct answer.

7.4 ANALYSIS OF *FLOWERS ON THE ROOF*

The source text (ST) as well as the Afrikaans and isiZulu versions of this story can be found in Appendix E. *Flowers on the Roof* is the only released text that was shared between PIRLS Literacy 2016 (n=761[51]) and PIRLS 2016 (n=830[52]).

7.4.1 Overview of Flowers on the Roof

Flowers on the Roof is a realistic fiction literary text that was first introduced during PIRLS 2001 and is shared with PIRLS Literacy 2016. The story revolves around two characters, Granny Gunn and a boy, who is also the narrator of the story. At the beginning of the story, the boy introduces his new neighbour, Granny Gunn, where she lived and how many animals she had. However, she fell ill and had to move to the city, next door to the narrator. The boy was excited to have a new neighbour as he hoped it would be another little boy that he could play with. He was disappointed when it turned out to be Granny Gunn. Granny Gunn was feeling homesick and did not like her new flat. He suggested that she should fetch her animals to make her feel more at home.

Flowers on the Roof is set mainly in Granny Gunn's new flat in town. The explicit theme for this story involves Granny Gunn being adaptable as she learns that wherever you move to, you can make it a home. For the boy, he learned that it does not matter who moves in next to you as you can make a friend with unexpected people such as a granny. The abstract theme focuses on intergenerational friendship and comfort, for example, making new friends and adding

[50] Recursion has embedded clauses.
[51] Fifty-nine Grade 4 learners were removed from the analysis due to extreme scores.
[52] Thirty-three Grade 5 learners were removed from the analysis due to extreme scores.

things around you so that you feel at home. Table 7.13 portrays the characters' traits, attributes and functions in *Flowers on the Roof*.

Table 7.13: Character Description of Flowers on the Roof

Character Name	Traits or Attributes	Functions
Granny Gunn	She is an independent granny who loves animals. She is also eccentric and resourceful.	She shows the importance of being independent and having ownership of things. Granny Gunn also shows how difficult it can be to move to a new place, leaving the home you loved for so long. But if you make an effort, you can make your new house feel like home.
The Boy	He is a kind, open-minded and welcoming young boy.	He shows the invaluable characteristic of being kind and welcoming to other people and in the process, making new friends.

Literary or rhetorical devices such as simile, imagery and dialogue were used in *Flowers on the Roof*, as identified by the IEA.

- *Simile* is a figure of speech that is used with words such as 'like' or 'as'. A simile makes a comparison between the things and points out the similarities between these two things (Literary Devices Editors, 2019). By using similes in a story or text, it assists the reader in gaining a better understanding of what the author intended to say.
- *Imagery*: as previously discussed, see page 238.
- *Dialogue*: as previously discussed, see page 238.

Table 7.14 presents examples of the abovementioned literary devices across the three languages for *Flowers on the Roof*.

Table 7.14: Literary Devices for Flowers on the Roof across Languages

Literary Devices	English	Afrikaans	isiZulu
Simile	"Her farmhouse was just like a doll's house."	"Haar plaashuis was net soos 'n pophuis."	"Umuzi wakhe wasemakhaya wawufana nse nendlu kanodoli."

Literary Devices	English	Afrikaans	isiZulu
Simile	"I feel as if I'm back home"	"Ek voel asof ek terug is by die huis"	"Ngizizwa sengathi sengibuyele ekhaya"
Imagery	"The hens are cackling all around me, and if I squint, I can easily imagine that the mountains I see are those near my farm. All that's missing is the smell of earth and grass."	"Die henne kekkel al hier rondom my, en as ek my oë op skrefies trek, kan ek my maklik verbeel dat die berge wat ek sien, die berge naby my plaas is. Al wat kortkom, is die geur van grond en gras."	"Izikhukhukazi ziyakekela lapha eduze kwami, kanti mangicijisa amehlo, lezi zintaba engizibona la, zingenza ngicabange ukuthi ezangakithi le duze nepulazi lami. Okungekho nje iphunga lenhlabathi nelotshani."
Dialogue	Dialogue between Granny Gunn and the doctor. Dialogue between Granny Gunn and the boy.	Dialoog tussen Ouma Gunn en die dokter. Dialoog tussen Ouma Gunn en die seun.	Ngaphakathi kuka Gogo Gunn kanye nodokotela, Kanye nangaphakathi kuka Gogo Gunn kanye nomfana.

Table 7.14 shows that the literary devices used in the text were the same across the ST and the Afrikaans and isiZulu versions. As with the previous text, this finding indicates that although the text was contextualised and translated into Afrikaans and isiZulu, the translated versions were able to maintain the literary devices. The translated versions also did not lose any of the meaning of the three literary devices, which means that the ST and translated texts (TT) should read similarly.

The plot of *Flowers on the Roof* includes a problem which is then resolved. The problem occurs when Granny Gunn decided to move to town as her doctor requested, due to her ill health. However, she did not like her new urban dwelling as it did not have grass and flowers on the roof and none of her animals were with her. The little boy, because of his kind and welcoming demeanour, helped Granny Gunn to make her new flat feel like home. Below is a list of the major events that occur during the text:

- Granny Gunn lives alone in a farmhouse with a cow, seven hens, two sheep and a cat.
- Granny Gunn becomes ill.

- Her doctor recommends that she moves into town.
- She sells her farm and the neighbours said that they will take care for her farm animals (excluding the cat).
- Granny Gunn packs her things and moves into her new flat.
- The boy is excited and hopes that another little boy moves in.
- Granny Gunn is not impressed with her new flat and says that she is going to move back to her farm.
- Granny Gunn discovers the balcony on the outside of her window.
- From the balcony she can see the mountains and the sea in the distance.
- Granny Gunn decides to stay after she sees the beautiful scenery.
- She still looked unhappy the following day.
- The boy asks whether she misses her farm animals. After she admits it, he suggests that she fetches them.
- Granny Gunn goes to back to the country and brings her chickens to her new flat.
- Granny Gunn states that her new flat almost feels like her home in the countryside.
- She decides to attach grass to the roof and starts a rooftop garden.
- Granny Gunn is happier now because her animals are in the flat and there is grass on the roof.
- However, the little boy wonders how she will get her cow into her flat (IEA, n.d.-b).

The list above highlights the main events that occurred in the story and the same idea units occur in each language text. It focused on Granny Gunn and a little boy who had helped her make her new flat a home. Some key vocabulary was identified by the IEA, and is presented in Table 7.15.

Table 7.15: Key Vocabulary and Expressions for Flowers on the Roof across Languages

IEA Vocabulary	Frequency English	Frequency Afrikaans	Frequency isiZulu
balcony	3	3	3
fetch	1	1	1
dreadful	1	1	1
clucking [a]	2	1	1
lift	2	2	1

IEA Vocabulary	Frequency English	Frequency Afrikaans	Frequency isiZulu
squint	1	1	1
turf	1 (grass, grass, grass, turf)	1 (gras, gras, gras, grassooie)	1 (ngotshani, nelotshani, notshani, zotshani)

[a] The word 'cackling' was changed to match the South African context to 'clucking'.

Table 7.15 depicts the key vocabulary as well as the frequency with which these words occurred in English, Afrikaans and isiZulu. These key vocabularies were selected by the IEA based on the function of the word, for instance depicting emotions, actions and descriptions. The words '*flowers*' and '*roof*' were not selected as key words. The English word '*balcony*' was translated into '*balkon*' (balcony) in Afrikaans, however, the isiZulu word '*kuvulanda*'/'*uvulanda*' (veranda/balcony) was used. The compound word '*kuvulanda*' means at/on/from/in the veranda whereas '*uvulanda*' is the word for veranda. The word '*cackling*' was contextualised to '*clucking*' in the South African ST version of the text. Thereafter, '*clucking*' was translated '*kekkelgeluid*' in Afrikaans and '*ziyakekela*' in isiZulu, the *zi-* is a subject concord and *ya-* is a prefix used that indicates the tense (Keet & Khumalo, 2017b). In English, the word '*turf*' was used and in Afrikaans, it was translated to '*grassooie*' (grass turf). The isiZulu word for grass is '*utshani*' and where turf was first mentioned in the text, the word '*zotshani*' was used. The *za-* is a possessive concord added to the root word '*utshani*'. As previously explained, the structure of isiZulu is complex as it adds prefixes, which may include subject concords and verbal prefixes which indicate tense and object pronouns, to root words. As such, the root word could be used several times but the prefixes could change the meaning of the word stem (Land, 2015). Table 7.16 shows the word profile of *Flowers on the Roof* for English.

Table 7.16: Word Profile of Flowers on the Roof

Word frequency level (English)	Occurrence	% of Tokens	Cumulative Tokens (%)	Examples
K1	779	93.7	93.6	granny, grass, flowers
K2	27	3.2	96.8	block, cow, ill
K3	1	0.1	96.9	sighed

Word frequency level (English)	Occurrence	% of Tokens	Cumulative Tokens (%)	Examples
K4	1	0.1	97.0	fond
K5	13	1.6	98.6	balcony, winked, hens
K6	1	0.1	98.7	squint
K11	2	0.2	98.9	clucking
Off list	0	0	≈100	

The word profile of *Flowers on the Roof* conforms to the text-lexis that is required for reading comprehension (for example, Laufer, 1989; Schmitt et al., 2017) as 95% of the words used during the text comprised K1 and K2 levels – most of the words used were between the 1 000- and 2 000-word frequency. This finding indicates that more than 95% of the words are considered as high frequency words, so it should be relatively easy to read as both Grade 4 and 5 learners should have these words in their vocabulary. The story also made use of words such as '*cluck*' and '*squint*' which are beyond the high frequency range of K-level words. The words '*squint*', '*balcony*', '*winked*' and '*hens*' are between the 3 000 and 8 000 rage, meaning that they fall in the mid-frequency vocabulary range (Schmitt & Schmitt, 2014); whilst the word '*cluck*' is beyond the 9 000 word threshold, indicating that it is a low frequency word (Schmitt & Schmitt, 2014). *Flowers on the Roof* is a narrative text and balances the range of frequency words, as most are contained within the K1 and K2 frequency ranges. Table 7.17 indicates the readability analysis of *Flowers on the Roof* for each of the selected languages.

Table 7.17: Readability of Flowers on the Roof per Language

Readability Features	English	Afrikaans	isiZulu
No of words	804	859	577
Pages/paragraphs	4	4	4
Sentences per paragraph	20.2	19.5	19.2
Words per sentence	9.9	11.0	7.4
Mean word length	3.9	3.8	7.1
Passives [a]	1%		
Flesch reading ease [b]	92.5		
Grade level [c]	2.8		

[a, b, c] These features can only be calculated for English.

Table 7.17 shows the text length across English, Afrikaans and isiZulu. As anticipated, the word count of the English and Afrikaans texts was closer compared to isiZulu[53]. The length of *Flowers on the Roof* might seem long for South African Grade 4 and 5 learners, yet according to Hasbrouck and Tindal's (2017) reading norms, Grade 4 English learners at the 50th percentile should be able to read 112 words correctly per minute (WCPM) by the middle of the year and 123 WCPM by the end of the year. As such, it should take the average learner about 7 minutes to read *Flowers on the Roof* (*cf.* Appendix E). A discussion of the items that displayed non-equivalence between the languages is presented in section 7.3.4.

Even though the word count differs between the three languages, the number of sentences per paragraph was similar across the languages. In terms of the number of words per sentence, the Afrikaans and English versions had 11.0 and 9.9 words per sentence, respectively – whereas the isiZulu (7.4) has somewhat lower words per sentence count. Moreover, the mean word length also differs between the three languages. The mean word length of the English (3.9) and Afrikaans (3.8) texts were similar whereas the mean word length of the isiZulu version was estimated at 7.1, which reveals the conjunctive orthography of isiZulu.

7.4.2 Processes of Comprehension of Flowers on the Roof Items

This section focuses on the four *Processes of Comprehension*, which indicate item difficulty, for *Flowers on the Roof* across the selected languages. This section also provides partial evidence whether non-equivalence is present for this text and accompanying items. *Flowers on the Roof* consists of 13 question items and has a total mark allocation of 16.

Of the 13 items, three items counted two marks each. Seven of the items took the form of MC type items and the remaining six were CR items. The CR items account for nine points whilst the MC items counted seven points. This text was developed for the PIRLS assessment in 2001 and tests a wide range of the *Processes of Comprehension*. Table 7.18 presents the items for *Flowers on the Roof* in relation to the different processes per language.

[53] *cf.* Page 235 for explanation about orthographies about the different languages.

Table 7.18: Item Difficulty of Flowers on the Roof across Languages

Processes of Comprehension	English Items	Afrikaans Items	isiZulu Items
Focus on and Retrieve Explicitly Stated Information	4, 9	4, 9	4, 9
Make Straightforward Inferences	2, 3, 5, 6, 8, 10	2, 3, 5, 6, 8, 10	2, 3, 5, 6, 8, 10
Interpret and Integrate Ideas and Information	1, 7, 12	1, 7, 12	1, 7, 12
Evaluate and Critique Content and Textual Elements	11, 13	11, 13	11, 13

As this text was developed for PIRLS and not PIRLS Literacy, the text was accompanied by fewer lower order items that test the *Focus on and Retrieve Explicitly Stated Information* process. Six of the items focused on *Make Straightforward Inferences*. Of the higher order items, three items tested learners' ability to *Interpret and Integrate Ideas and Information* and two items tested whether learners could *Evaluate and Critique Content and Textual Elements*. In addition, the text adhered to the percentage split between the four *Processes of Comprehension*, required by PIRLS, namely that 20% of the items may test *Focus on and Retrieve Explicitly Stated Information*, and 30% of the items may test *Make Straightforward Inferences*. The remaining 50% of the items must be split 30% and 20% for the two higher order skills, namely *Interpret and Integrate Ideas and Information and Evaluate* and *Critique Content and Textual Elements* (Mullis & Martin, 2015). From the information captured in Table 7.18, the items remained on the same comprehension level and difficulty across the three languages.

7.4.3 Professional Opinion on Translated Texts

Both translated versions of *Flowers on the Roof* had low incidence of low frequency words, similar to the ST. isiZulu expert 1 (Z1) stated that "most of concepts known to the learner, concrete things [sic]". Z2 did mention that "*uyajokola*" could have been replaced with a more commonly used word, such as "*uyahamba*" which means that she or he leaves. In terms of the Afrikaans version, A1 mentioned that the word "*koddig*" (funny) is not often used in everyday Afrikaans. The English version of the sentence in which this word was used, is as follows: "Granny Gunn winked at me and gave me a funny grin" and the Afrikaans version reads:

"Ouma Gunn het vir my geknipoog en koddig vir my geglimlag" (Granny Gunn winked at me and gave me a funny smile). Even though A1 flagged the word "*koddig*" as an unfamiliar word, the point of literary texts is to enrich learners' vocabulary by including fewer familiar words and not merely recycling a narrow range of everyday high frequency words. Moreover, as a picture of Granny Gunn is included which shows that she is grinning and winking, learners could have made the connection between "*koddig*" and how she smiled. The Afrikaans and isiZulu experts agree that the language used in the text is plain and easy to understand; as Z2 explains, the "words used in the text were simple and appropriate for intermediate phase learners." However, Z3 indicated that the overall translation of *Flowers on the Roof* was very poor as it included some isiXhosa words, used 'free translation' (does not strictly follow the ST), included words not in either the English or Afrikaans version, in some cases invented a new word, and made use of a slang word.

In terms of the cultural equivalence of the translated versions of the ST, the experts agreed that *Flowers on the Roof* should be equally interpreted and understood by the Grade 4 and 5 English, Afrikaans and isiZulu learners. Z1 continued by saying "…the [sic] are no cultural impediment. The only issue is that of Granny who has to go on the roof which is rare in isiZulu way of living". It should be noted that in most cultures, it would not be commonplace for a granny or any person except construction workers, to go on the roof, and planting greenery on the roof is perhaps a more Scandinavian tradition. A1 did mention that in South Africa, farmhouses would likely be made of brick and metal (zinc) roofs due to veld fires. The experts agree that the text should not be interpreted differently among the different language and cultural groups. Not all learners in South Africa are familiar with farms. However, literary texts transport the reader from out their known worlds and introduce them to other ways of being.

Besides growing things on a roof, the experts agreed that the text, *Flowers on the Roof*, is deemed to be construct relevant as it does not feature any unique features that may pose a hindrance for some learners. Z1 remarked that the text "talks about things the learner is used to and can associate with them". Both Grade 4 and 5 South African learners should be able to cope with the content, topic, genre, events as well as language used in this story. Z2 stated that the pictures "…are relevant to the story and are appealing to your learners. Learners will be able to predict the story by looking at the illustrations". Z3 agrees that the learners should be able to read the story even though it is abstract and that the illustrations assist with the understanding of the text.

The experts agreed that *Flowers on the Roof* looks like a typical realistic fiction story that South African Grade 4 and 5 learners should be able to cope with. The text is somewhat longer than the required number of words stipulated in the national curriculum (*cf.* DBE, 2011). As the text makes use of humorous and quirky pictures to illustrate incidents as the story progresses, it should draw the attention of the reader. All the experts liked the illustrations, with A1 stating that the illustrations used were excellent. Z2 commented that "lots of illustrations were used, learners like illustrations" and Z1 noted that "the setting in the story is local and is common". This text was developed for the PIRLS 2001 cycle but was used as a trend text during the PIRLS 2016 cycle. It was also used for PIRLS Literacy 2016 in order to create an anchor between those countries who tested at PIRLS level and those at PIRLS Literacy level. The national curriculum stipulates that Grade 4 learners should be able to engage with 150 to 200 words to test their reading comprehension whereas Grade 5 learners should be able to cope with 200 to 250 words (DBE, 2011). Both of these word counts are lower than the number of words used in *Flowers on the Roof*, yet the length should not serve as an obstacle as Z2 explains "there is no area in the passage that I thought would be difficult to learners [sic]". Furthermore, the layout of *Flowers on the Roof* is the same across the English, Afrikaans and isiZulu versions. It consists of four pages of text followed by five pages with question items. The text and question items have a slightly smaller font than the other PIRLS Literacy texts. The spacing is also narrower than PIRLS Literacy texts, but are similar to other PIRLS texts. Even though the spacing is narrower, it is still fairly easy to read. The layout is specifically set so that each page of the story contains the same pieces of text and images, at the same places, across the ST and TT, which adds to the face validity of *Flowers on the Roof*. This means that as the layout is the same across the ST and TTs, sections of information relevant to answering questions are located in the same paragraphs across the different languages.

With regard to the *Processes of Comprehension*, the Afrikaans and isiZulu versions of *Flowers on the Roof*, remained the same as the ST, meaning that none of the items changed its difficulty. This finding means that none of the items, based on the translations, displayed bias in this regard. It is worth noting that Z2 specifically mentioned that in the isiZulu version, the splitting of words was incorrectly done, for instance "*ey-iveni*" (van) which splits the prefix "*ey-*" and the root word "*iveni*". Z2 also indicated that where possible, the words should be used in full, for example "*kozilanda*" (fetch them) should be "*ukuyozilanda*" (to fetch them). The Afrikaans experts mentioned that some of the phrasing of the items may have been a direct translation, for example "*Watter van hierdie*" (Which of the) could rather have been "*Watter een van die*"

(Which one of the). The experts agree that the question items in *Flowers on the Roof* did measure what they intended to measure as they tap into the different processes of comprehension.

7.4.4 Per-item Analysis for Flowers on the Roof (Grade 4)

This sub-section provides a summary of the *Flowers on the Roof* items. It indicates the percentage of incorrect responses for each language, zero percent per item, items that displayed differential functioning as well as a selection of learner responses for each item that displayed DIF. Table 7.19 presents the number of Grade 4 learners who had incorrectly answered each of *Flowers on the Roof* items.

Five items of the *Flowers on the Roof* text indicated differential functioning among the Grade 4 learners. Items 1 and 5 are both MC type items while items 6, 9 and 12 are items which require learners to write an answer. Item 9 is a literal type (access and retrieve) item, which is considered to be easy, yet the Grade 4 learners found the item very challenging with most learners providing an incorrect response – isiZulu (89,9%), Afrikaans (86,2%) and English (72,1%). The missing values for item 9 were less than 10% for each language.

Items 5 and 6 require learners to make inferences based on the information contained in this text and are fairly easy and part of the lower order comprehension skills tested during PIRLS. Both these items proved to be difficult for the Grade 4 learners as the majority of the learners obtained a zero score. In terms of item 5, the English group (75%) experienced the highest percent incorrect whereas isiZulu and Afrikaans had 61% and 56% respectively. Seemingly most learners attempted to answer item 5 as the missing data is low and ranges between 1% and 3%. For item 6, it would appear that the item was more difficult for the isiZulu learners as the majority (87%) obtained a zero mark, with three-quarters of English (77%) and Afrikaans (76%) learners incorrectly answering the item. The missing values for item 6 was below 10%, with the isiZulu group's missing data being the highest (9%).

The remaining two items, 1 and 12, form part of the higher order comprehension skills as they tested learners' ability to interpret and integrate information from the text. It is expected that learners may have some difficulty in correctly responding to this type of item as it requires higher thought processes.

The majority of the isiZulu (76%) learners struggled with item 1. It could be considered more difficult for the isiZulu learners as fewer Afrikaans (67%) and English (54%) learners provided incorrect responses. The missing values for item 1 were quite low for each of the languages.

Item 12 is considered the most difficult item of the *Flowers on the Roof* items as very few learners were able to provide even a partially correct response. None of the learners were able to obtain full marks for this item. Moreover, this item showed moderately high missing values; 14% of the isiZulu learners did not attempt item 12. This missing percentage is closely followed by English (12%) and Afrikaans (11%). Special attention is paid to the aforementioned items, presented in Table 7.19, and these are discussed after the table.

Table 7.19: Flowers on the Roof: Percentage of Grade 4 Learners who Incorrectly Answered Items per Language

Item No	Processes of Comprehension	English			Afrikaans			isiZulu		
		N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing
4	Focus on and Retrieve Explicitly Stated Information	341	46	2	203	57	2	276	60	1
9*	***Focus on and Retrieve Explicitly Stated Information***	341	72	7	203	86	4	276	90	8
2	Make Straightforward Inferences	341	36	5	203	40	6	276	53	9
3	Make Straightforward Inferences	341	48	5	203	55	3	276	61	6
5*	***Make Straightforward Inferences***	341	75	3	203	56	2	276	61	1
6*	***Make Straightforward Inferences***	341	77	4	203	76	3	276	87	9
8	Make Straightforward Inferences	341	79	6	203	80	3	276	92	4
10	Make Straightforward Inferences	341	63	8	203	56	4	276	63	12
1*	***Interpret and Integrate Ideas and Information***	341	54	2	203	67	3	276	76	2
7	Interpret and Integrate Ideas and Information	341	89	4	203	88	3	276	95	4
12*#	***Interpret and Integrate Ideas and Information***	341	88	12	203	89	11	276	86	14
11	Evaluate and Critique Content and Textual Elements	341	62	6	203	57	6	276	71	10
13	Evaluate and Critique Content and Textual Elements	341	49	11	203	58	10	276	55	15

*Indicates items that displayed DIF.

#Item counts three marks but none of the learners obtained full marks for this item. The item shows the percentage of learners who were not able to provide a partially correct response.

MC Item 1, worth one point, tested a higher order literacy skill, namely *Interpret and Integrate Ideas and Information* and required learners to recognise perspective. To recapitulate, *Flowers on the Roof* item 1 asked the following question:

Who is telling the story?

a. a granny
b. a child* (correct answer)
c. a doctor
d. a farmer

Table 7.20 presents the number of persons who selected each of the distractors.

Table 7.20: Flowers on the Roof Item 1: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	141	41	115	57	182	66
B*	151	44	59	29	61	2
C	23	7	18	9	20	7
D	19	6	4	2	7	3
9 (not attempted)	7	2	7	3	6	2
Total	341	100	203	100	276	100

*Correct response (distractor B)

Forty-four percent of the learners who completed the test in English selected distractor B, while only 29% and 2% of the Afrikaans and isiZulu learners, respectively, selected it. It would appear that a large majority of learners were tempted by distractor A indicating that the learners thought the granny told the story. This finding could be due to several possible reasons in that learners may have experienced confusion between the main character and the narrator and/or lack of familiarity with literary conventions such as narration.

Item 5 displayed non-equivalence among the languages. To reiterate, *Flowers on the Roof* item 5 required learners to take perspective of the situation and asks the following question:

Granny Gunn did not like the walls and windows in her new flat. Why else was she unhappy?

a. She was ill.
b. She missed her cat.
c. She did not like the balcony.
d. She felt homesick.* (correct answer)

Table 7.21 shows the percent of learners who selected each of the distractors.

Table 7.21: Flowers on the Roof Item 5: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	75	22	50	25	79	29
B	101	30	28	14	63	23
C	80	23	35	17	27	10
D*	76	22	85	42	103	37
9 (not attempted)	9	3	5	2	4	1
Total	341	100	203	100	276	100

*Correct response (distractor D)

It would appear that the Afrikaans learners found this item somewhat easier than the other two languages, as 42% of the Afrikaans learners selected distractor D while only 37% and 22% of the isiZulu and English learners, respectively, selected it. Across all three languages, distractor A seemed tempting as learners accessed memory and recalled that Granny Gunn was sick and that is why she had moved to the city. The question required the learners to take perspective of the situation as it asks 'Why else was she unhappy?'. The question item already told the learners that she was unhappy about the walls and windows and required them to infer, based on large portions of the story, why else she was unhappy. In the text, Granny Gunn was still unhappy even though she had her cat, Robert, with her as well as her view from the balcony. The little boy asked whether she missed her animals after seeing that she was still unhappy – from this can be inferred that the granny was homesick.

The next item that displayed non-equivalence among the languages is item 6:

Why did Granny Gunn scream when the cat jumped out of the window?

This item required learners to *Make Straightforward Inferences* from the text. This is a constructed response type item worth one mark which required taking perspective of the situation. Learners had only to list one appropriate inference to explain why Granny Gunn had screamed[54]. Table 7.22 provides a selection of ten Grade 4 learners' responses to *Flowers on the Roof* item 6.

Table 7.22 (following page) shows that learners gave incomplete answers, for instance isiZulu learner 2 and 3. Learners also copied sentences from the text, see English learner 1, Afrikaans learner 2 and isiZulu learner 1. Few learners, such as Afrikaans learner 5, provided nonsensical responses where they copied letters randomly. Based on the text, the learners could have made several inferences of why the granny had screamed; for instance, Granny Gunn did not know that there was a balcony or she thought the cat could get hurt (*cf.* Appendix C). As this is an inference-type item, the answer is not explicitly stated in the text and required learners to think of a reason for the granny screaming when the cat jumped out of the window. Learner responses in green indicate a correct answer, red responses are an incorrect answer and purple indicates nonsense answers.

[54] *cf.* Appendix C for scoring guide.

Table 7.22: Flowers on the Roof Item 6: Grade 4 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	Then she suddenly gane little scream. Robert the cat jumped out of th window	d die mense op die buurplaas die dokter	Ionga limbe uGogo uGunn wahl-aselwa ukugula
2	Still atleast she had a cat	Robert die kat het by die venster uitgespring!	Not attempted
3	Granny Gunn is much happier now. Shes made abit of countryside here in the tuaop.	Sy was so hartseer dat sy uitendelik besluit het om haar kat, Robert, saam.	Not attempted
4	She thought Robert fell out of the flat.	sy het nie van die balkon gehou nie sy het terugverlang plaastoe	Ingoba uGogo waye ngalivalile kahle ngu uGogo laphuna ikati
5	It was going to fall of the flat.	bmlumeume abinbemo slomlumbe	ingoba lizofa futhi yena uGogo u-Gunn kakhulu akanawa
6	She was afraid.	Hy het nie tug ge kom nie	ugogo ushaya izihnuknu phekle ndu
7	Because the cat will go to in front of walking cars.	Sy was bang hy val	Wayesaba ukulibamba uGogo u-Gunn wayesaba no kuthi lizomukwebha
8	Granny Gunn is much happier now	Die ouma het baj gu sukl Hulle het gu sukl om hius reg ti maak	Walethanda iz izilwane zake ngoba zazimutabuli sa waye jabulile nezilwane zijabulile
9	Then she sudde gave her a little scream	Ouma Gunn het vir my geknipoog en koddig vir my gegil	yingoba wayelithanda kakhulu
10	Cranny is telgking the cat	Ek dink ons moet môre drop toe gaan.	Ingoba lalifuna ukubuyela laphi eliphuma kona

Flowers on the Roof item 9 also showed differential functioning between the three languages. It asks the following question:

Write two ways in which Granny Gunn made her new flat feel like home.

This item required learners to find two actions granny had taken to make her new flat feel like home. As such, it required the reader to see the world from X's perspective – in this case, the granny's perspective. As this is a literal question, the information was explicitly stated in the text – as such, learners needed only to locate the information in the text. Learners would have obtained two marks, one for each action taken. Some of the possible answers[55] include that the granny had put grass on the roof and she brought her chickens to her new flat; or the granny brought her cat with her and went to fetch her farm animals. If, however, the learner only provided one correct response, then that learner would have obtained a partial correct. Table 7.23 (following page) shows a selection of ten Grade 4 learners' responses to *Flowers on the Roof* item 9.

[55] *cf.* Appendix C for scoring guide.

Table 7.23: Flowers on the Roof Item 9: Grade 4 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	Soon she had sold her farm and bought a flat in our block of flats	Sy was siek die mense op die burmrplaas Skryf twee maniere neer waarop ou verander het sodat die weer soos haar	kwabekun- guGogo u-gunn phezulu ophahleni
2	Cow How is going to lift the cow cat The cat is too small	Ouma Gunn is nou baie gelukkiger Sy kan enigiets doen! Daar is net een ding wat haar nou bekommer.	Nami ngabe ngesese ngempela Kwaba into ele elukhuni kani kanye
3	She can don anything! Theres only one thing that bothers her now. How is she going to get the cow in to the lift	Ouma Gunn het want mure Ouma Gunn het bruin kat	wase ethenga indlu laphakithina emakutethini Wabe enesasasa kakhulu futhi ekujihile
4	She put grass on top of the roof. She brought her animals to live with her in town.	Ek soe of as ek kan wanag dy die huis het 'n stutkkie hy sbak kar En daar grooi moos dlommo gaeie	wayekha ndu yakheyesibili ukuzae anbe nezindlu ezilanbili ugogo sengi zoba nerdly yaimi lapha e dolobheni nami
5	She brought her animals. She put grass on the roof.	bunlfrenemlf vefumbofu oobumbfurlgl rbusmlglof	wawakhela ikatilakhe indlu kanye nezinkukhu zakhe kanye nokunye futhi wahamba uGogo u-Gunn wayegula kakhulu wahamba

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
6	She wanted her cat. She took the thinks of her house and put it in the flat	Sy mis die plaas dis nie vir haar nie mooi nie	izwakore njengek noyolankhe edidala ia izindlela ezimblili ugogo enzeindlu
7	he clamb up the house th he hit the house	Oor dit lelik was Dad dit soos die plaas voel.	udokotela uRobert uGogo Gunn
8	Granny Gunn mad her new. Write ways in which her.	Ouma sy kynt se maan was Gum hy het hom grus guho Ouma Gunn het verby my na die balkon toe gestrm.	Ngiziza umangisese khalga lami elidala Ikhaya lami elisha ngiyalithanda nmakhu yanmi womabili
9	Granny Gunn	Sy Sy het nie van die groot vensters gehou nie En sy het nie van d die balkon gehou nie	yebo wayengalifuni wayethanda isitezi
10	Not attempted	En daar groei weer eens blomme op die dat doen!!! Hoe gaan sy die koei in die hysbak kry?!	kwathi umi sebebuyele edolodeni ugogo wakweza izigephu phezu kophlala lwendlu. uGogo u-Gunn aheve ethokozile mapje. Phela waba esizakhele uphandla lendle edolobheni.

From the learner responses captured in Table 7.23, some learners provided incomplete answers such as English learner 10. Learners also provided incomplete sentences, such as isiZulu learner 7 who mentioned *udokotela* (doctor); *uRobert* (Robert); *uGogo Gunn* (Granny Gunn). Afrikaans learner 1 wrote the second part of the question as the answer. Some learners also provided nonsensical responses where letters were copied randomly; see for example Afrikaans learner 5. It would appear that this item was somewhat difficult for the learners as very few learners were able to obtain full marks. Learners were required to give two actions that the granny had done to make her flat feel like home, and these are mentioned from the middle towards the end of the story. It would appear that some learners did not have the necessary reading and writing skills to answer this question. Other reasons may include test fatigue, not being interested in the test or not understanding the question posed. Table 7.24 depicts the breakdown of the partial credit item per language.

Table 7.24: Flowers on the Roof Item 9: Partial Credit Breakdown

Points	English Persons	% Obtained	Afrikaans Persons	% Obtained	isiZulu Persons	% Obtained
0	174	51	137	67	224	81
1	72	21	38	19	24	9
2	72	21	20	10	7	3
9 (not attempted)	23	7	8	4	21	8
Total	341	100	203	100	276	100

Even though the answers to this question are explicitly stated in the text, very few of the learners were able to obtain full marks for this item, with only 3% of the learners who completed the test in isiZulu having correct answers. Half of the English learners (51%) and about two-thirds (67%) of Afrikaans learners incorrectly answered this question. The isiZulu (81%) learners had the highest percentage of incorrect answers. This item also shows higher missing percentages than other items highlighted in this sub-section. It would seem that learners struggle to provide long answers, struggle to retrieve explicitly stated information from the text, are not skilled in reading and writing and may experience test fatigue.

The next item that indicated non-equivalence was *Flowers on the Roof* item 12. This item required the reader to take second order perspective of the situation, in other words, X's perspective changes because X sees Y's perspective. The item asked the following:

What were the little boy's feelings about Granny Gunn when she first moved in and at the end of the story? Use what you have read to describe each feeling and explain why his feelings changed.

This story, as with other literary texts, enables the reader to get into the minds of the characters and become privy to their feelings. When learners are questioned about the characters' feelings it requires higher order understanding. For this item, learners were required to recognise that the little boy's feelings had changed as the story progressed; then the learners had to describe these feelings and explain why they changed. This item counts three points[56].

Full marks were awarded if learners provided how the little boy felt when the granny moved in, at the end of the story and how his feelings had changed. However, none of the learners across the languages received full marks for this item. Learners received either one or two marks depending on the extensiveness of their answer. Table 7.25 presents a selection of ten Grade 4 learners' responses to *Flowers on the Roof* item 12.

Similar patterns are observed for *Flowers on the Roof* items 12 and 9. Based on the learner responses captured in Table 7.25, it would appear that learners copied answers from the text even though this item required learners to integrate ideas from across the text. An example thereof is Afrikaans learner 10 and isiZulu learner 1. Some learners provided incomplete answers such as English learner 9, while others provided incomplete sentences such as Afrikaans learner 9. Learners also seem to include parts of the question as their answer; see for instance English learner 8. Afrikaans learner 5 wrote nonsensical responses, where letters were copied randomly and grouped into 'words'. In addition to these types of responses, English learner 5 used ellipsis in his/her response. Ellipsis is a commonly used literary device where a person can easily infer the omitted information from the context (Lexico, 2020a). However, in formal schooling, learners are usually taught to provide full sentences when asked a question.

[56] *cf.* Appendix C for scoring guide.

Table 7.25: Flowers on the Roof Item 12: Grade 4 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	There was no one at home when I came to visit her the next day Granny Gunn had taken the bus out into the country That night I woke up to hear a strange clucking sound comeing up the stairs	om na grappie by die verhaal te ver om die les van die verhaal, te verdee om die storie geloofwaardik te maak om die leser te help verstaan wat was die klein seuntjie se gevoelens sy aanvanklik inbetertrek het en aan	uGogo u-Gunn wathi ukungicwayic-wayizela engi buka ngokumoyizela okun gajwayelekile
2	Not attempted	Sy is nou net so lief vir haar daktuim as wat sy vir haar ou plaas was.	ukusiza ofundayo ukuba oqonde nga kwanzeka
3	Granny Gunn packed all her things into a van and was soon on her way to her new home. She was very excited and really lookinf forward to seeing the town. I was very excited too.	Gelukkig was die mense by die buurplaas so gaaf om te se dat hulle die siere sou versorg. Dit was nog steeds vir Ouma Gunn baie moeilik om vir haar dierevriede totsiens te se. Sy was so hartseer dat sy uiteindelik besluit het om haar kat, Robert, saam.	Not attempted
4	He was excited because he could not wait to see who was moving in. He was sad because the person that was moving in was a granny. He was happy because he the granny was fun and exciting.	Om 'n grappie verhaal met hiel om die les van die verhaar te verdiodlik volg om die soerie ge loofwaan dig te maam die help verstaan wat gebeur	wamtha nda ugogo ugunn wamsebenzela besizane amzize ugogo uGunn-ugogo uGunn wamupha Imali.

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
5	He was very exicyted because he saw his granny.	bmlqlu blneme mluqlunbf bume mluqolnemlu gluref lunegl wealglume Sbenim lunigl mu welo	umfanyana wathatha isinqumosoku gada izinkukhu zika Gogo u-Gunn kanye nekati laGogo u-Gunn nokunye kwezi mfuyo zaGogo u-Gunn waye umfana owaphatha ikhayalika Gogo u-Gunn waye
6	He could be unhappy because his Granny is no right at all and he must be very very … sad for that.	Hy was bly, opgewonde en om ieman te onmoet	Not attempted
7	The boy feel unhappy to his Granny. The Granny feel so sad to live his house. The boy was happy to go to the other house.	Hy het sleg gevoel. Hy het meer van haar begin hou. Hy het later baie van haar gehou.	Not attempted
8	little boys feelings about	Ouma Gunn het al haar goed in 'n vragnotor goepk en was sy op pad na haar nuwe huis was uitgesine om die dap darn 'n kat gehad.	Not attempted
9	Not attempted	Hy was ongelukkig En sy was bang En sy was gelukkig	wayejabula kakhulu ngoba wayejabulle
10	What is the boy is feeling about Granny is likes the animals Crann What she did Granny she likes her home	Ouma Gunn wat nie te gelukkig toe sy haar nuwe woonstel nie. Dit is sowaar verskriklik!	Umfanyana wayenalamazi" u-Gogo Gunn wayehlale yedwa izimbali ophahleni lwendlu kuGogo u-Gunn wagcina esihlala edolobheni

This item was difficult for learners as most of them were not able to correctly answer the question. This item tested a higher order skill, which is more difficult as learners were required to integrate information from across the text in order to describe the little boy's feeling along with an explanation. As this item was at the end of the test, learners may have experienced test fatigue as they may not have had the required test-stamina to complete it within the allotted time frame. Table 7.26 shows the breakdown of the partial credit item per language.

Table 7.26: Flowers on the Roof Item 12: Partial Credit Breakdown

Points	English Persons	% Obtained	Afrikaans Persons	% Obtained	isiZulu Persons	% Obtained
0	211	62	152	75	221	80
1	61	18	21	10	13	5
2	28	8	8	4	3	1
3	0	0	0	0	0	0
9 (not attempted)	41	12	22	11	39	14
Total	341	100	203	100	276	100

None of the learners were able to obtain full marks for *Flowers on the Roof* item 12. The majority of each language received a zero. Moreover, less than 10% of each language's learners were able to obtain two marks. All three languages had more than 10% missing scores. Based on the poor results of this item, Z2 suggested "the question could have been divided into two different parts" to assist the learners completing it. The next section provides the Grade 5 learner item-analysis of *Flowers on the Roof*; however, only items that were not presented in this section were explored as similar patterns were observed.

7.4.5 Per-item Analysis for Flowers on the Roof (Grade 5)

As *Flowers on the Roof* was used in both PIRLS Literacy 2016 and PIRLS 2016, one would expect the Grade 5 learners to show increased reading stamina and a deeper understanding of the narrative text in terms of literal and higher order comprehension skills. This sub-section only focuses on items that were not highlighted by the Grade 4 results, and items include item 7, 8, 10 and 13. These four items displayed differential functioning and a selection of Grade 5 learner responses is provided for each of these items. Table 7.27 shows the number of Grade 5 learners who incorrectly answered the *Flowers on the Roof* items.

Table 7.27: Flowers on the Roof: Percentage of Grade 5 Learners who Incorrectly Answered Items per Language

Item No	Processes of Comprehension	English			Afrikaans			isiZulu		
		N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing
4	Focus on and Retrieve Explicitly Stated Information	350	38	1	167	36	1	346	51	1
9*	***Focus on and Retrieve Explicitly Stated Information***	350	60	6	167	78	1	346	87	9
2	Make Straightforward Inferences	350	26	3	167	25	1	346	46	6
3	Make Straightforward Inferences	350	36	2	167	38	1	346	58	2
5*	***Make Straightforward Inferences***	350	74	1	167	40	1	346	60	1
6	Make Straightforward Inferences	350	66	2	167	66	2	346	88	4
8*	***Make Straightforward Inferences***	350	61	4	167	53	2	346	89	6
10*	***Make Straightforward Inferences***	350	24	6	167	29	1	346	78	8
1	Interpret and Integrate Ideas and Information	350	43	1	167	46	3	346	70	2
7*	***Interpret and Integrate Ideas and Information***	350	86	4	167	77	2	346	92	5
12*#	***Interpret and Integrate Ideas and Information***	350	90	10	167	96	4	346	87	13
11	Evaluate and Critique Content and Textual Elements	350	49	7	167	55	2	346	69	5
13*	***Evaluate and Critique Content and Textual Elements***	350	45	10	167	54	3	346	60	9

*Indicates items that displayed DIF.

#Item counts three marks but none of the learners could obtain full marks for this item. The item shows the percentage of learners who were not able to provide a partially correct response.

The Grade 5 learners also found items 5, 9 and 12 difficult, which were all questions requiring taking perspective of the situation. Figure 7.1 shows the percentage of incorrect responses of Grade 4 and 5 learners for items 5, 9 and 12. One would expect the Grade 5 learners to perform better in lower order comprehension items, such as item 9, compared to the Grade 4 learners. Even though the Grade 5 learners did improve overall, the improvement for isiZulu learners was small in comparison to the other two languages. It would appear that the Grade 5 English and Afrikaans learners experienced the text processing easier than the isiZulu group.

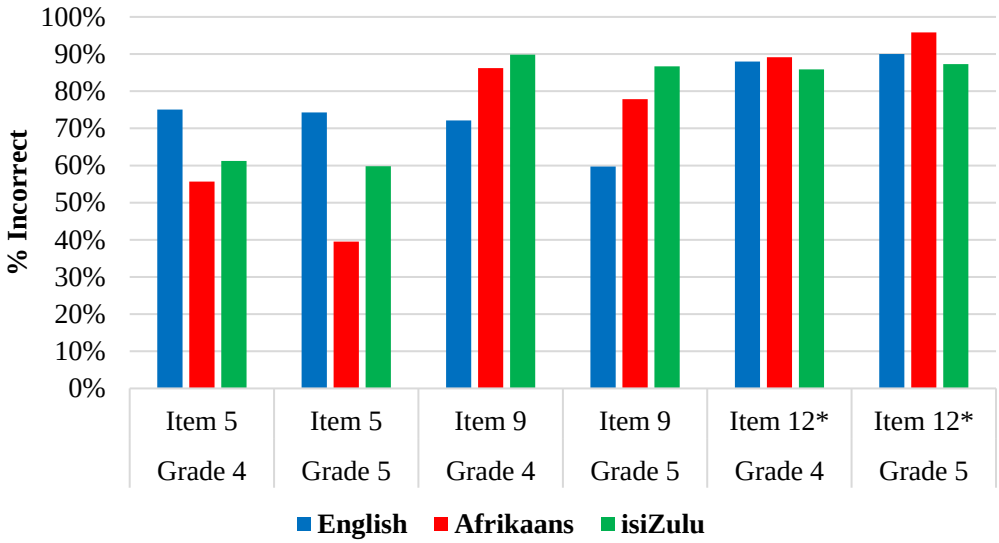

Figure 7.1: Percentage of Incorrect responses of Grade 4 vs. Grade 5 Learners on Selected Flowers on the Roof Items

There were slightly fewer incorrect responses for English and isiZulu learners from Grade 4 to Grade 5 on item 5 – this finding suggests that both grades experienced the item as difficult. Furthermore, the Afrikaans Grade 5 learners showed the most progress as 40% of Grade 5 learners provided incorrect responses – in other words, most Afrikaans Grade 5 learners experienced the text processing easier and were able to obtain the mark. Item 9 saw some improvement from Grade 4 to Grade 5 as the English, Afrikaans and isiZulu Grade 5 learners provided fewer incorrect responses. These learners were able to find the correct answer in the text. In terms of item 12 (higher order skill), both Grade 4 and 5 learners found this item difficult as none of the learners were able to obtain three marks; however, some learners were able to obtain a partially correct response (either a one or two). Item 12 also had the highest percentage missing data of all the *Flowers on the Roof* items. It was suggested by one of the experts that the item may have performed better if it were split into two items.

Beyond the abovementioned items, an additional four items showed differential functioning among the Grade 5 learners, namely items 7, 8, 10 and 13. Item 7 is a higher order (interpret and integrate) item and asks learners to provide an answer worth two marks. Items 8 and 10 both test lower order comprehension skills (straightforward inferences) where learners had to provide a response for one mark each. The last item, item 13, tested learners' ability to evaluate and critique contents of the text. This item is a higher order MC item requiring learners to make a decision between four possible answers.

Item 7 tests the higher order literacy skill *Interpret and Integrate Ideas and Information* and required learners to provide an answer[57] that links Granny Gunn's feelings about her view from the balcony to her home in the country, in order to gain two points. In other words, the learners had to take perspective of the situation and interpret Granny Gunn's feelings about the mountains and the sky; for example, the mountains reminded her of the country, or she was thinking about her farm and missing it. Learners received one mark if they provided a correct response that describes Granny Gunn's feelings about the mountains or why she crouched down – the learner response, however, does not provide a connection between the granny's feelings and the countryside. All three language groups found this item difficult as the English learners had 86% incorrect responses, while the Afrikaans and isiZulu had 77% and 92%, respectively (*cf*. Table 7.27). Item 7 asks the following question:

When Granny Gunn was on the balcony, she crouched down so that she could not see any of the rooftops – only mountains and the sky. Why did she do this?

Based on the Grade 5 learner responses captured in Table 7.28, learners provided incomplete answers – such as isiZulu learner 5 and English learner 10. Some learners provided incomplete sentences, for instance English learner 8. Other learners copied sentences, or parts of sentences, from the text and provided these as their answers, see for example isiZulu learner 1 and 2. The response by Afrikaans learner 10 does not make sense, as words are strung together to form a sentence that in itself does not make sense.

[57] *cf.* Appendix C for scoring guide.

Table 7.28: Flowers on the Roof Item 7: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	Because her farmhouse had grass and flowers on the rooftop but the flat has no grass and flowers	Ouma Gunn op balkon was nie, dakke kon sie nie berge en die lug. Ouma Gunn in die dorp gewoon.	Awugula kakhulu kodwa sen gathi umgathuthela edolabheni kwasha udokotela.
2	She remembered her farmhouse	Gnun balkon was sy sodat nie lug van dakke sien nie berge en lug.	uGogou-Gunn akathokozaga lu lapho eqalaza endlini yakhe ent-she eyifulethi
3	So that she does not remember her animals	Sy het verlang na haar plaashuis. Sy het haar diere op die plaas gemis.	uphuphume impilo lgama langemp ngu Gunn jona kodwa mina ngimbiza ngoGogo Gunn
4	It might have caught her attention or it reminded her of the country.	Sy het besluit om te bly	Wayefuna ukubona ekhayalakhe.
5	So she can p be a little bit happy.	Want sy het verlang na haar plaas. Sy het verlang na haar diere.	Not attempted
6	She did that because she was scared of falling and because she doesn't like the sea.	Sy het dit geniet to sy uitkyk	Ingoba wayefuna kuhlale izilwane zake
7	She wanted to see a great view.	Sy het gesien toe sy op die balkon.	Yingoba ukhumbula ekhaya epulazini.
8	to feel well	Daar was net gras op die dak.	Wayefuna ukubona izindlu zabaye abantu ukuze abone.
9	Because she could see the mountains far away and even a bit see. Granny Gunn crouched down so that she couldn’t see any of the rooftops - only the mountains and the sky.	Sy het dit gedoen om die goeie dae van haar plaas te onthou, sonder al die geboue	Not attempted
10	Not attempted	Dit is sowaar verkule het so gese Die mure is heelglo en die res is to wat en wet	ubuka nje lawa mafasitela makhulu ayizibhandlazi nje

This item appears to be difficult for the learners as few of them were able provide a response that includes the connection between the granny's feelings and the countryside. Table 7.29 presents the breakdown of the partial credit item per language.

Table 7.29: Flowers on the Roof Item 7: Partial Credit Breakdown

Points	English Persons	% Obtained	Afrikaans Persons	% Obtained	isiZulu Persons	% Obtained
0	251	72	101	60	304	88
1	50	14	27	16	15	4
2	36	10	35	21	11	3
9 (not attempted)	13	4	4	2	16	5
Total	350	100	167	100	346	100

Of the learners who completed the test in isiZulu, almost 9 out of 10 learners received a zero whereas 60% of Afrikaans learners received zero. While this item appears to be difficult for learners across all languages, it was more so for the isiZulu learners. Twenty-one percent of Afrikaans learners obtained two points compared to 10% and 3% of the English and isiZulu learners, respectively. Based on the learner responses provided in Table 7.26, some learners had difficulty in providing long answers, with most giving only short answers. As this is a higher order literacy skill item, it may also indicate that learners are not used to interpreting and integrating information from across the entire text to answer a question. The majority of the learners were not able to make the connection between Granny Gunn's feelings about her farm in the country.

Flowers on the Roof item 8 is the next item that showed non-equivalence. It required learners to *Make Straightforward Inferences* about why Granny Gunn had winked and grinned at the little boy. The item required learners to provide an inference[58] worth one mark that showed that Granny Gunn had remembered that she could bring more of her animals to her new flat. Learners received one point if they could infer from the text why Granny Gunn had winked and grinned at the boy – indicating that she had a plan or that the boy had given her an idea.

[58] *cf.* Appendix C for scoring guide.

The majority (89%) of the isiZulu learners obtained a zero for this item while 61% and 53% of the English and Afrikaans learners, respectively, obtained a zero score (*cf.* Table 7.27). It asked the following question:

Find the part of the story by this picture of Granny Gunn:  . *Why did Granny Gunn wink and grin at the little boy?*

As with the previous item, similar patterns are observed for *Flowers on the Roof* item 8. Of the learner responses captured in Table 7.30, some did not provide an answer; for example, English learner 9 and isiZulu learner 5. In some instances, learners copied words either from the text or from the question item; for example, Afrikaans learner 1 and isiZulu learner 1 copied parts of the question as the answer while isiZulu learner 2 copied words from the text. Learners also provided incomplete sentences such as Afrikaans learner 2 who stopped in the middle of the sentence. It would appear that Afrikaans learner 10 copied letters and words which do not make sense. A reason for the poor results could be that learners were not sure what the facial cues 'wink' and 'grin' means and could not make the required inference from the text, even though a picture was provided to show the granny winking and grinning.

Table 7.30: Flowers on the Roof Item 8: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	The little boy told Granny Gunn to go fetch the animals.	Ouma Gunn klein seuntjie knipoop geglimlag	Thola ingxohye yalolu ngalo mfankiso uGaun kunyanl uSunn acwaylew
2	The little boy gave her an idea of going and taking her animals.	Sy het 'n stukkie platteland in	uGogou-Gunn wathi ukungic wayic-wayizela engibuka okun
3	he gave her an idea to go and fatch the animals.	Omdat sy was bly om hulle (haar diere) te gaan haal.	gingakuxoxela ngogogo othize engimaziyo akeve evisimanga
4	Because she had an idea.	Hy het vir haar se sy kan die diere gaan haal	Ngoba wayenga namanga
5	She winked and grin because the little boy gave him an idea.	Ouma Gunn het met die bus platteland toe gery.	Not attempted
6	She winked and grined at the boy because she had a plan.	Omdat sy 'n plan gehad het.	Not attempted
7	Because there was know at home.	want sy was gelukkig	kungoba ulojagula
8	to play with him	Sy was fly.	ingoba ugogo uyamathanda ufuna usakhe
9	Not attempted	Sy het gedink aan wanner sy al haar diere woonstel to vat.	yingoba wayengathandi ukunlala edolobheni wayefuna uku phindela emakhaya
10	Not attempted	Se het vir di seki oor egl.	Kwaba kungekho muntu ngakusasa laphongi thi ngizomuvakashela.

The next item that displayed differential functioning was item *Flowers on the Roof* item 10. This item required learners to *Make Straightforward Inferences* from the text by specifically focusing on Granny Gunn's feelings at the end of the story, as such, the item required the reader to take perspective of the situation. Most (78%) of the learners who completed the test in isiZulu incorrectly answered the question while 29% and 24% of the Afrikaans and English learners, respectively, also provided incorrect answers (*cf.* Table 7.27). This item posed the following question:

At the end of the story, how did Granny Gunn feel about her new home?

As this is a dichotomous item, learners would have received one point[59] if they could provide a reasonable inference as to how Granny Gunn, at the end of the story, felt about her new home. Their answer should have included a positive feeling, for instance, she decided that she liked it or she was happy, as it felt like home. Table 7.31 shows a selection of ten Grade 5 learners' responses to *Flowers on the Roof* item 10.

The majority of English and Afrikaans learners correctly answered the question by indicating that she was happy with her new flat by the end of the story. It would appear that the isiZulu learners found this item more difficult than the English or Afrikaans learners. Some of the learner responses captured in Table 7.31 indicate the positive feelings the granny had; for example, English learner 1 and Afrikaans learner 5. Other learners simply did not provide an answer; for instance, English learner 9 and isiZulu learner 5. A few learners copied words from the text such as isiZulu learner 2, who copied a segment from a sentence in the text. Moreover, a few learners provided incomplete sentences, such as Afrikaans learner 3 who only wrote the word '*bekommerd'* (worried) as an answer – this answer was added in the scoring guide as an incorrect response as it does not demonstrate the granny's positive feeling at the end of the story.

[59] *cf.* Appendix C for scoring guide.

Table 7.31: Flowers on the Roof Item 10: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	She Granny Gunn is now happier	Ouma Gunn het haar nuwe woonplek gevoel?	ukufakza wagcina eseziziwa
2	She felt take she was bock home	Guna gevoel van die oor nuae.	Awuguli kakhulu kodwa
3	She felt very happy	Bekommerd	uGogo uGunn wabe ezibwane
4	She loved it	Sy was by	Wagcina eselithanda
5	She felt happy because it looked like the farm	Sy was gelukig.	Not attempted
6	She felt a little satisfied about it.	Sy het (op) baie gelukkige gevoel	Not attempted
7	She felt happy.	Sy was gelukig ge wees oor haar huis	Wayese jabule kakhulu
8	Shee feeled happier abot her home	Sy was fly.	kwabu nenzinto eziningi
9	Not attempted	Bekommerd, want hoe gaan sy die koei in die hysbak kry.	wayeseweziwa kahle ngekha lakhe elisha
10	Not attempted	Sy het hulle vengus uibere	wayesezizwa e jabele kakhulu

Flowers on the Roof item 13 is a MC type item that displayed differential functioning. The question required learners to *Evaluate and Critique Content and Textual Elements* in order to answer the question. Learners had to, based on the details, construct a theme at a higher level of abstraction – constructing a situation model, thus, concluding that when you move to a new house or flat, you are able to make it feel like home by bringing familiar things with you. To recapitulate, it asked the following question:

Which of the following might you learn from this story?

a. Old people will never be happy if they change where they live.
b. You can make a new place feel like home if you bring familiar things with you.* (correct answer)
c. You can get used to living with animals, even though they are noisy.
d. Children and old people do not make good friends.

The following table depicts the percent of learners who selected each of the distractors.

Table 7.32: Flowers on the Roof Item 13: Grade 5 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	94	27	50	30	103	30
B*	158	45	72	43	108	31
C	52	15	32	19	73	21
D	12	3	8	5	32	9
9 (not attempted)	34	10	5	3	30	9
Total	350	100	167	100	346	100

*Correct response (distractor B)

Almost half (45%) of the English learners selected the correct distractor, closely followed by the Afrikaans (43%) learners. Thirty-one percent of the learners who completed the test in isiZulu selected distractor B. In the story, the granny first brought her cat with her, and then later she went to fetch her chickens and affixed some grass to the roof. Across the languages, it would appear that distractor A was somewhat tempting for the learners as almost one-third of learners selected it. Learners may have fixated on the negative feelings the granny had

experienced before she moved and when she first moved into her new flat. One of the pictures also showed how unhappy she was. As such, learners possibly focused on these negative feelings rather than carefully reading the question item and focusing on how she felt about her new home at the end of the story. The missing data for the item is low which indicates that most learners had answered the item even though it was the last item of the story.

As both Grade 4 and 5 learners completed the *Flowers on the Roof* question items, it is instructive to see whether there are some signs of development from Grade 4 to Grade 5. Figure 7.2 shows the raw mean score of both the Grade 4 and Grade 5 learners for this text (*cf.* Chapter Five and Six).

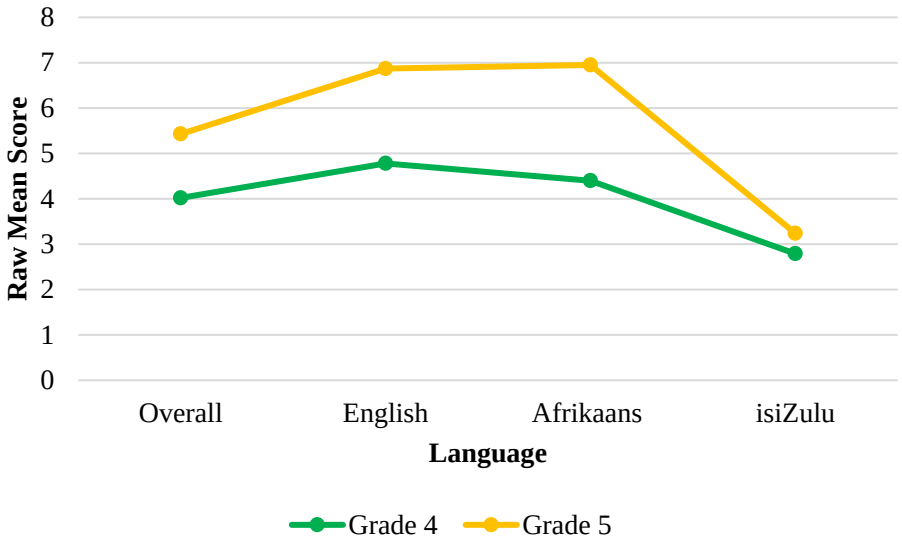

Figure 7.2: Flowers on the Roof: Grade 4 and 5 Raw Mean Score per Text by Language

Flowers on the Roof consisted of 13 items with a total mark allocation of 16. Based on Figure 7.2, the Grade 5 learners (5,43), performed somewhat better in comparison to the Grade 4 learners (4,02). This finding indicates that there appears to be some growth from Grade 4 to Grade 5. As it is difficult to ascertain from the above figure exactly where Grade 5 learners had performed better than the Grade 4 learners, Table 7.33 was devised. It indicates the percentage correct for each item[60] per language and grade.

[60] Excludes missing data.

Table 7.33: Flowers on the Roof: Item Percentage Correct per Language and Grade

Item Number	Processes of Comprehension	Afrikaans		English		isiZulu	
		Grade 4	Grade 5	Grade 4	Grade 5	Grade 4	Grade 5
1	Interpret and Integrate Ideas and Information	30	53	45	57	23	28
2	Make Straightforward Inferences	57	75	62	73	42	52
3	Make Straightforward Inferences	44	61	50	63	35	41
4	Focus on and Retrieve Explicitly Stated Information	41	64	53	61	39	49
5	Make Straightforward Inferences	43	60	23	25	38	39
6	Make Straightforward Inferences	21	33	20	32	4	9
7	Interpret and Integrate Ideas and Information	10	21	7	11	2	3
8	Make Straightforward Inferences	17	46	15	36	4	5
9	Focus on and Retrieve Explicitly Stated Information	10	21	23	37	3	5
10	Make Straightforward Inferences	42	71	32	74	29	16
11	Evaluate and Critique Content and Textual Elements	39	44	34	48	21	27
12	Interpret and Integrate Ideas and Information	0	0	0	0	0	0
13	Evaluate and Critique Content and Textual Elements	35	44	45	50	35	34

Overall, it would appear that the Grade 5 learners performed higher on most items. All languages showed improvement on the literal type items that only required learners to find the answer in the text (for example, items 4 and 9). The next lower order comprehension skill

tested was learners' ability to make inferences based on the information in the text – these include items 2, 3, 5, 6, 8 and 10. Grade 5 learners obtained higher percentage correct for these items. However, some of the increases were less than five percent, for instance item 5 completed by English and isiZulu learners. In one case, a lower percentage of correct responses (item 10) was recorded by the Grade 5 isiZulu learners. This item asked how Granny Gunn felt about her new home by the end of the story – the Grade 5 isiZulu learners did not infer that by the end of the story the granny felt happy about her new home. This item was problematic and indicated DIF.

In terms of the higher order comprehension skills, learners were tested based on their ability to interpret and integrate information as well as to evaluate and critique the text. Based on the results in Table 7.33, there is a mixed reaction; for instance, although learners performed poorly on item 1, the Grade 5 learners were able to perform slightly better than the Grade 4 learners. Yet, item 12 proved difficult for both Grade 4 and 5 learners. This item required second order perspective taking, which is always more difficult than first order. The learners were required to describe the little boy's feelings at the beginning of the story when the granny moved in, and how they had changed by the end of the story. Z2 had explained that if this item was divided into two parts, it might have yielded a better response from the learners. In terms of the most cognitively demanding items, items 11 and 13, the Grade 5 learners obtained a higher percentage correct for both these items. However, some of the increases were less than five percent; for example, item 11 for the Afrikaans Grade 5 learners and item 13 for the Grade 5 isiZulu learners. Nevertheless, it would appear that there is some growth between Grade 4 and 5 in terms of their reading comprehension. At Grade 5 level, it seems that learners have started to understand typical literary devices used in the text and could differentiate between different characters in the story.

7.5 ANALYSIS OF *MACY AND THE RED HEN*

The source text (ST) as well as the Afrikaans and isiZulu versions of this story can be found in Appendix F. In South Africa, 861[61,62] Grade 5 learners completed the items for this text.

[61] Four Grade 5 learners were removed from the analysis due to extreme scores.
[62] As PIRLS made use of a matrix design, 16 booklets (including the Reader) were used during testing. Learners did not all receive the same booklet with the same texts. See Chapter Two for more details.

7.5.1 Overview of Macy and the Red Hen

The text entitled *Macy and the Red Hen* is realistic fiction and was developed for PIRLS 2016. The story follows two main characters, one a girl and one an animal, namely Macy and the red hen. The story follows Macy and how she struggles to deal with the bossy red hen. At first, Macy is introduced to the reader while feeding the hens. At night, Macy's job was to make sure all the hens were safe inside their cage. However, the red hen was bossy and wilful and did not want to go into the cage. This led to Macy becoming frustrated and discussing the issue with her parents about what to do next. In the end, she had an idea from an owl and tricked the red hen into thinking that Macy was protecting her from the owl – it was in fact some cloth on a pole. Thereafter, Macy ascended to the top of the pecking order and the red hen went into the cage without a fuss. It is important that the Grade 5 learners understand the psychology of the two main characters as well as the nature of the pecking order amongst fowls. The pecking order is explained, in dialogue form, by Macy's mother in paragraph 2 on page 1 of the story.

Macy and the Red Hen takes place at Macy's home. The explicit theme for this story is how Macy struggled with the bossy red hen and how she outsmarted the red hen in the end. Macy then moved to the top of the pecking order. The abstract theme is that even the most frustrating problems can be solved by being clever (using your head) and creative; for example, Macy created an owl made of cloth, some wire and a pole so that it looked like an owl trying to catch the red hen and Macy, being brave, 'saved' the red hen from 'the owl'. The story relies on what Macy thinks the red hen will think (second order perspective taking). Table 7.34 shows Macy and the red hen's traits, attributes and functions.

Table 7.34: Character Description of Macy and the Red Hen

Character Name	Traits or Attributes	Functions
Macy	She is a clever and creative child but one who is also stubborn and frustrated.	She shows that even though you have a frustrating problem, you should not run away or avoid it but rather solve it in a creative manner.
Red Hen	She is a bossy hen and at the top of the pecking order. She is also stubborn and quite mischievous.	She is an agitator for Macy and keeps finding ways not to obey her and to play tricks on her. She also shows how a pecking order functions and that it is important in the animal world.

In *Macy and the Red Hen*, literary and/or rhetorical devices such as tactile imagery, imagery and dialogue are used:

- *Tactile imagery* is a sub-group of imagery focusing specifically on the sense of touch (Masterclass, 2020). By using tactile imagery, the author describes something that the reader is able to feel.
- *Imagery*: as previously discussed, see page 238.
- *Dialogue*: as previously discussed, see page 238.

Table 7.35 depicts the literary and/or rhetorical devices identified by the IEA.

Table 7.35: Literary Devices for Macy and the Red Hen across Languages

Literary Devices	English	Afrikaans	isiZulu
Tactile Imagery	"Macy could feel the hen's little heart pounding in her feathery chest"	"Macy kan voel hoe die hen se hartjie in haar geveerde borskas klop"	"U-Macy wayizwa inhliziyo encane yesikhukhukazi ishaya esifubeni saso esinezimpaphe."
Imagery	"a cloud of hens exploded into the yard"	"'n wolk van henne die werf binneplof"	"inqwaba yezikhukhukazi igcwala ebaleni"
Imagery	"flapping and fussing this way and that"	"klap haar vlerke en skarrel dié kant toe en daai kant toe"	"siphakamisa amaphiko aso futh siqholosha"
Imagery	"her little red eyes gleaming in triumph as Macy slammed the cage door behind her"	"haar klein rooi ogies glinsterend van oorwinning wanneer Macy die hok se deur agter haar toeklap"	"amehlo aso amancane abomvana ekhazimula ngenxa yokunqoba njengoba u-Macy eshaya umnyango wehhoko esivalela"
Imagery	"The owl swooped across the grass on its huge white wings, grabbed a mouse in its claws, and soared back into the shadows."	"Die uil duik met sy yslike wit vlerke oor die gras, gryp 'n muis in sy kloue, en vlieg terug in die skaduwees in."	"Isikhova sashona phansi otshanini ngamaphiko aso amhlophe amakhulu, sabamba iqgundae

Literary Devices	English	Afrikaans	isiZulu
			ngezidlavu zaso, saphinde sandiza."
Imagery	"The hen stopped running and fluffed up her feathers, squawked furiously and beat her wings"	"Die hen hou op hardloop en pof haar vere, kekkel woedend en slaan met haar vlerke"	"Isikhukhukazi sayeka ukugijima futhi savula amaphiko aso, sakhala ngolaka sabe sesishaya amaphiko"
Dialogue	Dialogue between Macy, her mother, her father, and Sam.	Dialoog tussen Macy, haar ma, haar pa en Sam.	Inkulumo-mpendulwano phakathi kuka-Macy, unina, uyise kanye no-Sam.

All three literary devices were used in the Afrikaans and isiZulu versions of *Macy and the Red Hen*. The literary devices did not change from the ST to the translated versions, which means that even though the story was translated, the translated versions were able to maintain the literary devices without loss of meaning.

The plot of *Macy and the Red Hen* focuses on a problem that Macy encounters but is resolved at the end of the story. In this story, Macy is tasked to look after the hens, which includes feeding them and securing them in the cage at night so that predators cannot get to them. However, the red hen, who is at the top of the pecking order, consistently plays tricks with Macy to show that she (the hen) is the boss. Macy speaks to her parents but they explained that she could not change her job or leave the hen out at night as they need the eggs. The solution to Macy's problem lies in Macy playing a trick on the red hen so that it would be obedient and go into the cage without making a fuss. In doing so, the red hen relinquishes her top pecking order spot to Macy. The following list provides the major events during the story:

- Macy opens the cage door to let the hens out in order to feed them dinner.
- The red hen bosses the other hens by pecking, flapping and fussing.
- Macy asks why the other hens allow the red hen to boss them.
- Her mother explains that the red hen behaves as such because she is at the top of the pecking order. She further explains the pecking order to Macy where each hen can peck the hens below it, but they cannot peck the hens above them.

- Macy has trouble with the red hen as she refuses to go into the cage at night. The red hen plays tricks with Macy every night until she decides that Macy has chased her enough.
- Macy tries unsuccessfully to get the red hen to behave by luring her with food and banging pot lids together to scare her.
- Macy is frustrated and threatens to leave the red hen outside at night so that predators could catch her. Her father, however, tells her that the family requires the hens to lay eggs.
- Macy's father explains that even if the red hen were no longer the top hen, another hen would simply take her place as the next top hen (meaning that Macy's problem will not be resolved).
- Macy angrily goes to her mother and asked if she can have a new job, as she no longer wants to take care of the hens. Her mother suggests a couple of new jobs; however, Macy does not like any of them.
- One night, Macy sees an owl swoop across the grass to catch a mouse and it gives her an idea.
- Macy goes to her brother for help to construct a fake owl using cloth, wire and a pole. She wants to use this fake owl to scare the red hen.
- That night while Macy is chasing the red hen, an owl appears and tries to attack the red hen. Macy then steps in to save her.
- By saving the red hen from the fake owl, Macy secures her position at the top of the pecking order.
- Macy picks up the red hen and cuddles her to comfort her until she calms down.
- Her brother declares that Macy is now at the top of the pecking order (IEA, n.d.-c).

The events as listed above, describe the major events that take place during *Macy and the Red Hen*. These events appear in the three versions of the text. The events specifically focus on Macy, her frustration and ultimately her cleverness and creativity that allow her to become the top of the pecking order. The key vocabulary and expressions, which were identified by the IEA, are captured in Table 7.36.

Table 7.36: Key Vocabulary for Macy and the Red Hen across Languages

IEA Vocabulary	Frequency English	Frequency Afrikaans	Frequency isiZulu
squawking [b]	2	2	1
pecking	5	5	2
bossy	1	1	1
bribed [b]	1	1	1
beady [a b]	1	1	1
lure	1	1	1
stomped [a b]	1	1	1
crouched	2	2	1
took charge	1	1	1
pecking order [a]	4	4	4

[a] Words have a different structure in Afrikaans.

[b] Two words are used for each of these words in isiZulu.

Table 7.36 presents the frequency of the key vocabulary of *Macy and the Red Hen*. The IEA selected these key vocabularies based on the functions of the word. Since language structure is different, some of the key words in the translated versions become compound words. For example, the English word '*squawking'* in isiZulu is '*nokubanga umsindo*' (to make a sound); and the word '*beady'* in Afrikaans is '*kraal*'; however, adjectives in English become compound nouns in Afrikaans, thus the word '*kraal*' is added to '*ogies*' (small eyes) to form the compound '*kraalogies*' (beady eyes). Similarly, the word *stomped* in Afrikaans is '*boos stap'* (angry walk) as English verbs become a verb + adverb in Afrikaans. In isiZulu, stomped is expressed as '*egxoba izinyawo'* (foot stomping). The word '*pecking'* was used five times in the English and Afrikaans versions. However, the isiZulu word for pecking is '*singqofa*' but as the language makes use of prefixes and subject concords, '*singqofa*' was used with different prefixes such as '*esingasingqofa*' which added *es* + *nga* + *si* to the root word, which indicates the tense and that it 'cannot peck'. The word '*esingasingqofa*' is also a kind of relative clause construction as there are few adjectives in isiZulu. The translation used was appropriate as the same meaning can be drawn from the word. Thus, the root word remains the same but the meaning can change depending on the prefix added to the root word (Land, 2015). Table 7.37 shows the word profile of *Macy and the Red Hen*.

Table 7.37: Word Profile of Macy and the Red Hen

Word frequency level (English)	Occurrence	% of Tokens	Cumulative Tokens (%)	Examples
K1	748	82.5	82.5	smiled, dinner, decided
K2	68	7.5	90.0	boss, cage, feather
K3	6	0.7	90.7	faded, solve, triumph
K4	19	2.1	92.8	beady, owl, soared
K5	43	4.7	97.5	bribed, fussing, hens
K6	3	0.3	97.8	fluffed, swooped, swoop
K7	12	1.3	99.1	beak, cuddled, peck
K8	1	0.1	99.2	Stomped
K9	3	0.3	99.5	squawked, squawking(2)
Off list	0	0	≈100	

As indicated, research has shown that if the reader knows at least 95% of the words, they could score 60% on a reading comprehension test (*cf*. Laufer, 1989; Masrai, 2019). However, as this story only reaches the 95% threshold at K5, the word profile of *Macy and the Red Hen* is slightly different from the other two texts in this chapter. This text includes a greater variety of words from the mid-frequency range, making it lexically richer. For example, words such as '*stomped*' and '*squawked*' which are at the K8 and K9 range, respectively, were used. *Macy and the Red Hen* is not appropriate for beginner readers - it was developed for Grade 4 readers internationally but was completed by Grade 5 learners in South Africa. Notably 83% of the words used in *Macy and the Red Hen* is within the K1 (0 – 1000 word) frequency level. Grade 5 learners should know the words between K1 and K5, as they should have a larger vocabulary than Grade 4 learners. It confirms the research done on vocabulary and its association with reading comprehension and serves as an indicator of text difficulty (Laufer, 2013; Laufer & Ravenhorst–Kalovski, 2010). As this text is a literary text, it contains a wider range of vocabulary than an informational piece. As such, it includes low, mid and high frequency words, although most of the vocabulary lies between the high and mid frequency ranges, which indicates that lexically, this text is balanced. Table 7.38 indicates the readability analysis of *Macy and the Red Hen* for each of the languages.

Table 7.38: Readability of Macy and the Red Hen per Language

Readability Features	English	Afrikaans	isiZulu
No of words	905	962	602
Pages/paragraphs	6	6	6
Sentences per paragraph	11.0	11.0	11.0
Words per sentence	13.7	14.5	9.1
Mean word length	3.9	3.8	7.6
Passives [a]	3%		
Flesch reading ease [b]	88.8		
Grade level [c]	4.2		

[a, b, c] These features can only be calculated for English.

The English text consists 905 words, whereas the Afrikaans has 57 more words. As expected, the isiZulu word count was much lower in comparison to the English and Afrikaans versions. The orthography[63] of isiZulu is conjunctive in nature, which means that even though the sentences are shorter than English sentences, the words are more condensed and require more time to read (Land, 2015). Furthermore, *Macy and the Red Hen* (*cf.* Appendix F) should take HL English learners, who completed the test in English, about 7 minutes to read, based on the norms of Hasbrouck and Tindal (2017). Grade 5 learners at the 50th percentile should be able to read 127 words correctly per minute (WCPM) by the middle of the year and about 13 additional words by the end of the year. Sub-section 7.4.5 provides a discussion of the items that functioned differently for the English, Afrikaans and isiZulu Grade 5 learners.

7.5.2 Processes of Comprehension of Macy and the Red Hen Items

This section provides a summary of the *Processes of Comprehension* that indicates item difficulty, for the text *Macy and the Red Hen*. The four processes are used to show whether the item difficulty changed during translation. As such, this section provides partial evidence of whether non-equivalence is present for *Macy and the Red Hen*. This text has 16 items with a total mark allocation of 19. It has an equal number of CR (8) and MC (8) items. Item 6 counted two marks whereas item 13 counted three marks – both these items are CR. As this text was developed for PIRLS 2016, just over a third of the items tested learners' ability to *Interpret*

[63] *cf.* Page 235 for explanation about orthographies about the different languages.

and Integrate Ideas and Information. Table 7.39 depicts each process as well as the corresponding items per languages.

Table 7.39: Item Difficulty of Macy and the Red Hen across Languages

Processes of Comprehension	English Items	Afrikaans Items	isiZulu Items
Focus on and Retrieve Explicitly Stated Information	1, 6, 7, 10	1, 6, 7, 10	1, 6, 7, 10
Make Straightforward Inferences	3, 5, 9, 11	3, 5, 9, 11	3, 5, 9, 11
Interpret and Integrate Ideas and Information	4, 12, 13, 14, 15	4, 12, 13, 14, 15	4, 12, 13, 14, 15
Evaluate and Critique Content and Textual Elements	2, 8, 16	2, 8, 16	2, 8, 16

Macy and the Red Hen has eight items that make up lower order items. These items specifically focused on the Grade 5 learners' basic reading comprehension skills. Half of these items tested whether learners could *Focus on and Retrieve Explicitly Stated Information* and the remaining half tested whether learners could *Make Straightforward Inferences*. These items account for nine marks. Five of the items, worth seven marks, tested how well learners could *Interpret and Integrate Ideas and Information*. The remaining three items were pitched at the most difficult level namely, *Evaluate and Critique Content and Textual Elements*. Across the three languages, it seems that the items tested the same process of comprehension indicating that there were no discrepancies across the languages in terms of difficulty.

7.5.3 Professional Opinion on Translated Texts

According to majority of the experts, the translated versions of *Macy and the Red Hen* had low incidences of low frequency words. isiZulu expert 2 (Z2) mentioned that "there's an English word 'unclipped' which was translated as 'remove the clip'. There is no Zulu word for 'unclipped' that I know of". Z3 indicated that '*umshudo*' should rather have been used for 'unclipped'. Z2 further explained that when unknown words are used in the isiZulu text, it must be accompanied by the English words in brackets. Both Afrikaans expert 1 (A1) and isiZulu expert 1 (Z1) agreed that this text is written simply and is easy to understand, while both Z3 and A2 commented that the text is on the whole plainly written. A2 provided the following excerpt from the text as evidence "sy moet elke aand die henne in hulle hok toemaak sodat

jakkalse en uile hulle nie beetkry nie" (every night she had to lock the hens in their cage so that the jackals and owls cannot get them). A2 refers to the word '*beetkry*' (to get something) as she identified it as a potentially low frequency word. Z1 felt that "this was a very goo [sic] and interesting story and mind boggling, with some playful tricks which will of course excite the learner as needed by young learners in this grade". While the translation of the ST into Afrikaans and isiZulu was generally regarded as good, Z2 mentioned that question item 5, distractor B in isiZulu was translated as '*the door does not close*', yet in the English it is written as '*the door is hard to close*'. As such, the correct isiZulu should have been '*umnyango uvaleka kanzima*'. It should be noted that even though there was a slight error in the translation of distractor B, the item did not show any non-equivalence. Z3 did however mention that the text included "too many basic errors". For example, '*umama wakhe*' (my mother) is a direct translation and the isiZulu version should have used '*unina*' (her mother). Another example is the use of '*futhi*' (but), it should only be used between two clauses.

The cultural equivalence of the translated versions of the ST remained the same. Z1 explained that "there is no need for cultural understanding which can be wrongly interpreted, as the story related to the reality which the learners can associate with". The story of Macy and her struggle with the red hen is a reality that South African learners may have encountered, especially learners who live rural areas and have to look after the livestock. A2 explained that the English version might be easier to understand than the Afrikaans version.

Overall, the experts agreed that *Macy and the Red Hen* is construct relevant as nothing in the text poses as a hindrance for learners nor are there any irrelevant features in the text. South African Grade 5 learners should be able to cope with the text, including textual aspects such as the content, topic, genre, themes and the language used in the story. Z3 stated "Chickens and their behaviour are fairly well-known to children of the grade appropriate age." The experts also agree that the pictures are relevant to the text and accurately portray main events. A2 mentioned that the first illustration where Macy scatters the dinner scraps for the hens could have been better illustrated to clearly show bits of food. The story looks like a typical story that can be found in textbooks, with "illustrations that are relevant to the text and length of the text are appropriate for Grade 5 learners" (Z2). In addition, the layout of *Macy and the Red Hen* is the same across the three languages. The story consists of six pages of text followed by six pages with question items. Both the text and question items are slightly smaller in font size in comparison to PIRLS literacy texts. Although the font size is smaller, the text is still easy to

read as the layout makes use of paragraphs and illustrations to guide the reader. Furthermore, the layout of the text in relation to the illustrations is the same – they are at the same places across the ST and TT, which adds to the face validity of *Macy and the Red Hen*. Note that the text and the question items were in two different booklets – the reason for the split is so that the Grade 5 learners are be able to keep the PIRLS 2016 booklet as a classroom resource.

In terms of the *Processes of Comprehension*, the experts agreed that this text adequately tested the South African Grade 5 learners' reading comprehension. Based on Table 7.35, the translated versions of *Macy and the Red Hen* remained the same as the ST, and as such the item difficulty remained the same. Z2 further elaborates, "all questions asked are about the text and what is happening in the text. Questions force learners to read the text with understanding in order to answer the questions asked". A1 and A2 differ in opinion regarding the *Processes of Comprehension* – A2 commented that *Macy and the Red Hen* only partially tests the four processes as Afrikaans learners might have found difficulty with the text in terms of the use of low frequency words and sentence length. A1 rebutted this argument as most of the words used in the text should be known to Grade 5 learners and the story explains in detail the importance of the pecking order.

7.5.4 Per-item Analysis for Macy and the Red Hen

The following sub-section presents a summary of the items, which includes the percentage of incorrect responses per language, missing percentage per item, items that displayed non-equivalence among the languages and a selection of learner responses for constructed response items that displayed differential functioning. Table 7.40 (following page) shows the number of learners who incorrectly answered items for *Macy and the Red Hen*.

Table 7.40: Macy and the Red Hen: Percentage of Grade 5 Learners who Incorrectly Answered Items per Language

Item No	Processes of Comprehension	English			Afrikaans			isiZulu		
		N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing
1*	***Focus on and Retrieve Explicitly Stated Information***	***348***	***28***	***2***	***173***	***14***	***4***	***342***	***21***	***3***
6*	***Focus on and Retrieve Explicitly Stated Information***	***348***	***72***	***3***	***173***	***64***	***3***	***342***	***92***	***4***
7	Focus on and Retrieve Explicitly Stated Information	348	65	3	173	71	4	342	70	6
10*	***Focus on and Retrieve Explicitly Stated Information***	***348***	***48***	***5***	***173***	***52***	***5***	***342***	***4***	***6***
3	Make Straightforward Inferences	348	81	2	173	80	3	342	90	4
5	Make Straightforward Inferences	348	43	3	173	40	5	342	54	4
9	Make Straightforward Inferences	348	68	4	173	75	5	342	81	7
11	Make Straightforward Inferences	348	36	5	173	41	4	342	41	8
4*	***Interpret and Integrate Ideas and Information***	***348***	***91***	***3***	***173***	***82***	***2***	***342***	***93***	***4***
12*	***Interpret and Integrate Ideas and Information***	***348***	***39***	***5***	***173***	***57***	***5***	***342***	***60***	***8***
13*	***Interpret and Integrate Ideas and Information***	***348***	***89***	***9***	***173***	***94***	***5***	***342***	***86***	***13***
14	Interpret and Integrate Ideas and Information	348	74	11	173	75	8	342	79	18
15*	***Interpret and Integrate Ideas and Information***	***348***	***51***	***12***	***173***	***51***	***7***	***342***	***69***	***18***
2	Evaluate and Critique Content and Textual Elements	348	40	2	173	42	5	342	56	6
8	Evaluate and Critique Content and Textual Elements	348	85	6	173	77	5	342	88	7
16	Evaluate and Critique Content and Textual Elements	348	69	14	173	72	10	342	73	19

*Indicates items that displayed DIF.

Of the 16 items, seven displayed non-equivalence amongst English, Afrikaans and isiZulu learners. These include items 1, 4, 6, 10, 12, 13 and 15 of which items 1, 10 and 12 are MC. The remainder of the items took the form of CR. Only two of the four *Processes of Comprehension* were highlighted as part of the Rasch analysis results (*cf*. Chapter Six).

Items 1, 6 and 10 comprise literal question items where learners had to find the correct answer in the text. Item 1 was relatively easy as most learners obtained correct responses; however, the difference in incorrect responses were varied between the three language groups – Afrikaans (14%) had the lowest percentage of incorrect responses followed by isiZulu (21%) and English (28%). The missing values for each language was below 5%. Although item 6 was a literal type item, the majority of the learners struggled with it – the isiZulu learners found this item somewhat more difficult as 92% provided incorrect responses compared to the English (72%) and Afrikaans (64%) learners. In terms of item 10, approximately half of the learners provided an incorrect response with 52% of Afrikaans learners having the highest percentage of incorrect responses, followed by English (48%) and isiZulu (44%). Both English and Afrikaans had missing value of 5% whereas isiZulu had a slightly higher missing value of 6%.

Items 4, 12, 13 and 15 are part of the higher order comprehension skills and tested Grade 5 learners' ability to interpret and integrate information from across the text. Most of the learners struggled with item 4, with isiZulu (93%) learners scoring the highest percentage of incorrect answers, followed by English (91%) and Afrikaans (82%) learners. The missing values were relatively low – 4% of the isiZulu did not attempt to answer the item compared to 3% of English learners and 2% of Afrikaans learners.

Item 12 showed a different pattern with item 4 as 39% of English learners provided an incorrect response; however, more than half Afrikaans (57%) and isiZulu (60%) had a higher percentage of incorrect responses. Five percent of the English and Afrikaans learners did not attempt to answer the item, whereas 8% of isiZulu learners did not attempt it.

Item 13 was slightly more difficult for Afrikaans learners as 94% provided an incorrect response, whereas 89% and 86% of English and isiZulu learners, respectively, provided an incorrect answer. This item showed moderate missing values – 13% of isiZulu learners did not attempt the item while 9% and 5% of English and Afrikaans learners, respectively, did not answer the item. This finding shows that very few learners were able to provide a correct response worth three marks.

Item 15 was somewhat less difficult for English and Afrikaans learners than for the isiZulu learners. Just over half (51%) of English and Afrikaans learners provided incorrect responses, whereas 69% of isiZulu learners provided incorrect answers. This item showed the highest missing values of the DIF items for *Macy and the Red Hen*. Almost a fifth of isiZulu (18%) learners did not attempt item 15, compared to 12% of English learners and 7% of Afrikaans learners. Each of these items displaying non-equivalence are discussed below.

Macy and the Red Hen item 1 is a MC type item, which tested whether learners could *Focus on and Retrieve Explicitly Stated Information* and counted one point. This item was less difficult as the highest percentage of incorrect responses were made by English learners (29%) followed by isiZulu (22%) and Afrikaans (15%). It would appear that this item was more difficult for English learners. To recapitulate, *Macy and the Red Hen* item 1 posed the following question:

What is Macy doing at the start of the story?

a. catching a hen
b. feeding the hens* (correct answer)
c. looking for eggs
d. collecting feathers

Table 7.41 depicts the number of persons who selected each of the distractors.

Table 7.41: Macy and the Red Hen Item 1: Grade 5 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	29	8	10	6	33	10
B*	242	70	141	82	260	76
C	34	10	7	4	32	9
D	35	10	8	5	7	2
9 (not attempted)	8	2	7	4	10	3
Total	348	100	173	100	342	100

*Correct response (distractor B)

Eighty-two percent of Afrikaans learners selected the correct distractor (B), while 76% and 70% of isiZulu and English learners, respectively, selected it. Very few learners selected the remaining three distractors, which means that learners were able to find the correct answer from the text. In addition, the picture may have also helped the learners as it shows Macy feeding the chickens.

The next item that displayed differential functioning was item 4. This item required learners to *Interpret and Integrate Ideas and Information* from across the text in order to give a reason for the red hen playing tricks on Macy. Most of the learners were not able to obtain the mark, with 93% of isiZulu learners and 91% and 82% of English and Afrikaans, respectively (*cf.* Table 7.40). The item asked the following question:

Why does the red hen play tricks on Macy?

Table 7.42 (following page) presents a selection of ten Grade 5 learners' responses to *Macy and the Red Hen* item 4.

Based on the different pieces of information in the text, learners had to give a reason why the red hen played tricks on Macy[64] – this required taking some perspective of the situation. Their answer should have included, for instance, that the red hen is at the top of the pecking order, or that she (the red hen) wants to show Macy who is the boss. The notion of the red hen being at the top of the pecking order was introduced at the beginning of the story where Macy's mother had explained to her that chickens have a pecking order.

[64] *cf.* Appendix C for scoring guide.

Table 7.42: Macy and the Red Hen Item 4: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	Because the red hen was sitting in the middle of the yard	Not attempted	Not attempted
2	She plays tricks so that Macy can't put her in when she wants	Omdat macy haar nie haar gunsteling kos gee nie	Sasingafuni
3	The red hen plays tricks on Macy because the red hen thinks she's the boss of everyone	Want sy is die baas	Ngoba sasifuna ukuthi asijahe aze akhathale futhi ahlekwe abantu.
4	Macy has asked her mother	Wand die rooi hen was die baas van die ander henne	yingoba sasixuma kakhula futhi sasiphaphile
5	But Macy had a different view	van is gewoonlik vat due rooi hen	naso singazingasa zonke ezinye izikhukhukazi
6	The red hen plays tricks on Macy to make her tired	Want macy wil haar 'n les leer.	Ukuphakela isikhukhukazi kungani ngobasikhukhukazi esibomvu
7	She does not want to go into the cage	want rooi hen wil met die jakkels en uil bakklei.	Izikhukhukazi zinohlelo lokung gofana" Kuchaza umama wakhe
8	So that she won't get into the cage	Om dat macy haar moet jaag en macy kry haar nie.	Ingoba kwaphela izinsuku ezimbili zingaza leli amaqanda
9	He is used to playing tricks on other hens	Sy hou nie daar van as sy oor haar pad loop nie.	Yi ngoba sasa ukuthi uma umacy esho kumama wakhe umamaka macy uzasizwela
10	Because she doesn’t want to listen to macy and doesn't want to get in the cage	Sy is baas en niemand mag naby haar kom	Yingo khona into abazihlelile izikhukhukazi doku lokugqofana

Some of the learners, based on their responses in Table 7.42, provided incomplete answers such as Afrikaans learner 1 and isiZulu learner 1. Some learners also provided incomplete sentences, such as isiZulu learner 2 who only wrote the word '*sasingafuni*' (he/she/it not wanting/searching). Learners also copied words from the text to serve as their answer; for example, English learner 4 and isiZulu learner 7. The copied texts do not provide a reason why the red hen played tricks on Macy. There were also learners who misunderstood the reason why Macy wanted to leave the red hen out at night – it was to punish her by making her vulnerable to predators, but the learner interpreted it that the red hen wanted to fight these animals. In addition to these types of answers provided by learners, some also focused on the fact that the red hen did not want to go into the cage. As this is a higher cognitive item, it required learners to make use of higher order skills, namely interpreting and integrating why the red hen played tricks, to be able to provide a suitable answer.

Item 6 also displayed differential functioning between the languages. It tested whether learners could *Focus on and Retrieve Explicitly Stated Information* in order to answer the question. This item counted two points[65] – learners needed to provide two things Macy did to get the red hen into the cage, but these did not work. Item 6 asks the following question:

Macy wanted the red hen to go into the cage. What are two things Macy does that do not work?

Table 7.43 (following page) presents a selection of ten Grade 5 learner responses to item 6 of the *Macy and the Red Hen* text.

[65] *cf.* Appendix C for scoring guide.

Table 7.43: Macy and the Red Hen Item 6: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	Macy had tried to lure the red hen by putting her favourite food Macy had tried banging pot lids together to frighten the red hen	Macy het vir die rooi hen te haar gunsteling kos saam binne in die hok te sit. Macy het probeer om potdeksels teen mekaar te slaan om die rooi hen skrik te maak.	Elinyelhi
2	She tried lay putting the hens favourite food She tried banging pot lids together to frighten the red hen	Macy probeer die rooi hen to probeer vang Macy het probeer om potdeksels teem mekaar te slaan sodat die hen kan skrik	Not attempted
3	She bangs the pot lids together She bribes the red hen with her favourite food	Jaag die hoender in die hok Die skorelgoed te was	Wabhaklaza umnyango wehhoko Ukutshela ubaba wakhe ukuthi ufuna ukufundisa isikhukhukazo isifundo.
4	He smiled "Besides, it would not solve your problem. Macy went to find her father	Macy jag die rooi hen Sy gooi die rooi hen	U-Macy wenza into ethandekayo ngokuthi zazingamuhlaphi U-Macy. Umama ka U-Macy wayethi akayophakela izikhukhukazi zakhe
5	I would like your job" her mother called after her "Mum, I don't like my job I need a different	sy het al har gunstelingkos te gee. Sy het al potdeksels teen mekaar te slaan.	Izikhukhukazi ziyamtanda umlzoli oziqinevayo kodwa u-macy wayenombone ohuukile
6	She punts in her favourite food in the cage. Banged pit lids to frighten the red hen.	Sy wil nie hê die rooihen moet in die hok in gaan nie.	ukuthiizinye izito eziningi ekhukhukazi izinto eqyisaba kuzohamba egoli okuningi iziphlahla

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
		Sy wil hê die jakkels en die uile moet haar vang.	
7	Chasing the hen Leting it go in lant	sy hou nie van haar werk nie sy wil nie die skottelgoed was nie	Uyise wamamatheka "kanti-ke lokhonge ke kuyixazulute inkinga "macy" husho uBaba ubaba wakhe" Ngifuna
8	Chases the hen. Sit in die middle of the yard when macy get's closer she flaps her wings so macy can't get her	Macy jaag het in die hok wat reg is Die hen wil nie in die hok gaan nie	Ukufundisa isifundo isikhukhukazi uzosishiya ngaphandle
9	She puts the hens favrite food She tried putting pot lids to frighten the hen	leer oor haar nad Nacy wil haar uitgooi	Ukuthi umacy angatheli ukudla kwe sikhukhukazi ngaphandle Ukuthi umacy anga sishikizi ngoba sizophuma ngaphandle
10	She chases the hen She banged the pot lids together to frighten the red hen but she scared the like hens	Sy probeer die rooi hen lok deur haar gunstelingkos saan in die hok te sit Potdeksels teen mekaar te slaan	Ukufakela izinkukhu ukudla ukupheka ageze itsha kade bedlangazo

Note. The mark allocation was taken verbatim from the booklets.

Learners had to provide two things Macy had done, but which did not work, to get the red hen into the cage. Some learners were able to provide two things Macy had done, such as English learner 1 and Afrikaans learner 1. Learners also provided sentences from the text that do not include things Macy had done, for instance English learner 4 and isiZulu learner 7. One learner also provided an incomplete answer, see isiZulu learner 2. Most learners did attempt to answer this question, even if by only providing a partial answer such as isiZulu learner 1.

Table 7.44 depicts the breakdown of the partial credit item per language.

Table 7.44: Macy and the Red Hen Item 6: Partial Credit Breakdown

Points	English	%	Afrikaans	%	isiZulu	%
	Persons	Obtained	Persons	Obtained	Persons	Obtained
0	211	61	76	44	270	79
1	40	11	35	20	44	13
2	87	25	57	33	13	4
9 (not attempted)	10	3	5	3	15	4
Total	348	100	173	100	342	100

Of the 342 isiZulu learners who completed this test, 79% received a zero compared to 61% of English learners. Based on Table 7.40, the Afrikaans learners found the item less difficult as only 44% had obtained zero scores. Moreover, 33% of the Afrikaans learners obtained full marks for this item compared to 25% and 4% of English and isiZulu learners, respectively. From the learner responses provided in Table 7.43, learners were unable to list two things Macy had done that did not work. Even though this is a lower order comprehension item and required learners to find the answers from the text, learners still seem to have struggled with this item. It could be that learners are not used to searching for answers in the text, but rather think up an answer, or that they are not used to providing a more extended response to a question.

Item 10 of *Macy and the Red Hen* is a MC item, which tested whether learners could *Focus on and Retrieve Explicitly Stated Information*. The percentages of incorrect responses were relatively close to one another (*cf.* Table 7.40). To reiterate, the item asked the following question:

How does Macy get her idea?

a. Macy's brother, Sam, tells her the plan.
b. Macy sees an owl catching a mouse.* (correct answer)
c. Macy's Dad tells her about owls.
d. Macy sees some wire and some white cloth.

Table 7.45 depicts the number of learners who selected each of the distractors.

Table 7.45: Macy and the Red Hen Item 10: Grade 5 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	64	18	37	21	73	21
B*	163	47	75	43	170	50
C	45	13	28	16	40	12
D	58	17	25	14	37	11
9 (not attempted)	18	5	8	5	22	6
Total	348	100	173	100	342	100

*Correct response (distractor B)

Most learners across the languages attempted to answer this question. Approximately half of the learners were able to select the correct distractor (B); isiZulu learners performed somewhat better on this item compared to the English and Afrikaans learners, with 50% of isiZulu learners selecting distractor B. In the text, Macy witnesses how an owl hunts and catches a mouse, which gives her the idea of making an owl to scare the red hen and make her think that she is protecting the hen from the fake owl. There is also a picture of an owl on the hunt. However, it seems that not all learners were able to find the answer in the text and provide it as their answer. This could possibly mean that learners are not used to searching for answers in the text and perhaps relied on their memory regarding what they have read. About a fifth of learners selected distractor A, possibly indicating that the learners remembered that Macy had spoken with her brother, Sam, but instead of checking the text, the learners selected it even though the text says that Macy told her brother about her plan. It would seem that learners relied on their memory to answer this question as the questions are only available at the end of the story, which requires learners to page back in order to find the answer.

Macy and the Red Hen item 12 was the next item that displayed non-equivalence amongst the languages. It is also a MC type item and tested the learners' ability to *Interpret and Integrate Ideas and Information* from across the text in order to answer the question. Moreover, the item required taking some perspective of the situation. This item asked the following question:

Macy "hit the wings with her hands and pushed the wings away." What does Macy want the hen to think?

a. that Macy is saving the hen* (correct answer)
b. that Macy is angry with the hen
c. that Macy is terrified of the owl
d. that Macy is playing with the owl

Table 7.46 shows the number of learners who selected each of the distractors.

Table 7.46: Macy and the Red Hen Item 12: Grade 5 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A*	193	55	66	38	108	32
B	72	21	53	31	111	32
C	39	11	34	20	71	21
D	25	7	12	7	24	7
9 (not attempted)	19	5	8	5	28	8
Total	348	100	173	100	342	100

*Correct response (distractor A)

As with the previous item, most learners attempted to answer this item as there is less than 10% missing data per language. Just over half (55%) of the English learners selected the correct distractor whilst less than 40% of the Afrikaans and isiZulu learners selected it. These learners were tempted by distractor B which indicates that these learners struggled to interpret and integrate ideas from the text, specifically about what Macy wanted the red hen to think when she saved it from the fake owl. This item requires second order perspective taking which is more difficult than first order perspective taking. The answer was not explicitly available in the text, but on the page where the story tells how Macy stepped in to save the hen is a picture of Macy saving the hen, which should have provided sufficient information for the learners to

answer this question. It is possible that the Afrikaans and isiZulu learners did not pick up the nuances of the story as they only focused on the fact that Macy was frustrated with the red hen, as it did not want to go into the cage like the other hens. As such, it may be possible that learners merely accessed their memory rather than referring to the text for evidence.

The next item that indicated non-equivalence was item 13. The majority of the learners found this item difficult as 94% of those who completed the test in Afrikaans gave an incorrect response, while 89% and 86% of the English and isiZulu learners, respectively, gave incorrect responses. To reiterate, the item asked he following:

You learn what Macy is like by the things she does. Describe what Macy is like and give two examples from the story that show this.

As this item counts three points, it required learners to show extensive comprehension[66] by interpreting and integrating information from across the text. Learners obtained three marks if they showed an understanding of what Macy is like (for example, a trait, feeling or attitude) and supporting it with evidence from the text. Very few of the learners were able to provide such an extensive answer, with some only being able to provide a partially correct answer with either listing one trait, feeling or attitude. Table 7.47 shows 10 Grade 5 learners' responses for *Macy and the Red Hen* item 13.

[66] *cf.* Appendix C for scoring guide.

Table 7.47: Macy and the Red Hen Item 13: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	Macy is like quite learner Macy always want to make a plan Macy is a good child	Macy moet elke dag die rooi han en die ander hane in te kry en sy was baie vies gewees.	Not attempted
2	Macy is kind. She does not like seeing hens suffer. She also doesn't like people or animals being mean to one another. She also really love hens.	Macy wie die hen vang. Macy se pa wil nie so he nie. Die hen het self net die hok gestap. Macy is al moeg vir die werk. Haar ma wil he sy moet skottelgoed was. Haar ma wil he sy die baba opas.	Not attempted
3	Macy is doesn't tolerate the red hen, she's smart and straight forward.	Haar pa het haar nuwe voorbeelde gegee het	Umasi wayezithanda isikhukhukazi ezinakekela wathukuthela ngenxa yesikhukhukazi esibomvu. Izibonelo: Ngokukti asisabise ngesikhoya ngokuthi asijahe mesibaleka ukuze singeke sidliwe ezinye izilwane.
4	Macy stomped to the kitchen. "Mum, I don't like my job, I need a different job."	Macy is 'n sort mens wat baie laat is en sy hou nie van laat.	U-Macy wezeka ebukeka edumele kakhula futhi wayephakela izinkukhu umqhazula futhi ebumele kakhu ngoba ubaba uthe azibulawe lezizinkuku eziningi kakhulu nezikhukhukazi zami engizitholile futhi engizithengile.
5	Macy shouted at the white wings. She hit the wings with her hands and pushed the wings away. Her borther lifted the wing upwards	Macy is goeie mens. Sy het har hen geret en syt 'n plan gemak	esibomvu kodwa wethusa exinye izikhukhukazi kwaphela izinsubzu ezimbili izinye khukhukazi

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
			u-macy wahamba wayfuna ubaba wakhe ngifuna
6	Macy like the hends and feeds the hens. Macy cuddled the hen until its heart slowed down, then gently put it in the cage and smiled at Sam.	Sy is 'n mens wat baie kwaad en streng is.	Uyafundisia isizulu goqobol uthukile ungana amicane kule ndaba ezibonisalokho eziningi izikhukhukazi inditkazi emnyama iyamyundisa ecikoleni yima amadonda amonengi
7	Trying to make red hen see she was saving her because she is tired of her playing tricks on her	Macyhet al probeer om ie rooi hen te lok. Macy stap boos die kombuis binne. "Ma ek hou nie van my werk nie, ek wil 'n ander werk hê." Macy vlug by die agterdeur uit.	Kodwa macy wayenombona ohlukile njalo ebasakukwakadingeka avakle iziknukhazi enhokweni lazoukuza izimpisi nezikhova u-macy ukuba asijane esibomue kodwa iesi sikhukhukazi
8	She is not lazy, she is a very good care taker	Macy is hartseer dat sy ma en pa vir haar nie om gee nie en dat die hen nie hok toe sal kom en die ander hen se pikordee sal hoom af vat nie	ususishiya ngaphandle isifundisa isifundo usikahe nokwanele
9	She is used to helping other animals to do everything she is a care giver and does a lot more things	Sy is 'n diere tipe mens. Sy gee vir die henne	Umacy uke ajahe sikhukhukazi esibomvu afune ukusiphakela ukudla kwasa umacy uze afune ukusijaha hafu singene e wokweni laso uyasi thanda umacy isikhukhukazi esibomvu
10	She likes all the hens to be fearly treated. She wants the red hen to listen to her. She doesn’t want any chicken rulling another	Sy is 'n wrede mens. Sy wil die hen skrik maak met potdeksels. Sy vang die hen met selfgemaakde vere wat aan 'n paal vas gemaak is.	Kumele apheke ekhaya agezeizitsha uhambe futhi uyonakela izinkukhukaza ukundla izimvuthu

From the learner responses captured in Table 7.47, it would seem that some learners provided incomplete answers, such as isiZulu learners 1 and 2, while isiZulu learner 8 provided incomplete sentences. There are also learners who copied words or sentences from the text and provided these as their answer; however, these are not related to Macy's demeanour nor do they provide examples thereof, see for instance Afrikaans learner 7 who wrote about Macy trying to lure the hen, complaining about her job and running out the door. Some learners provided answers that were considered too generic or vague, such as Afrikaans learner 5 who only said that Macy was a good person. This item required learners to write one trait or attitude relating to Macy and supporting it with an example from the text, which required learners to be able to interpret and integrate information across the text. Yet, learners struggled to provide one of Macy's feelings, traits or attitudes. Several of Macy's characteristics are highlighted throughout the text; for example, Macy is frustrated or angry and she slams the cage door shut and stomps to the kitchen. Across the text, Macy's feelings, traits and attitudes are shown with different examples, as such learners could have selected any section of the text to find one of her characteristics. However, as this item counts three points and was at the end of the test, learners may not have been motivated to complete the question even though they could see that it counted three points, which required more of them than a literal type question. Learners may have also experienced test fatigue as they had to read the entire text before starting with the question items. Table 7.48 shows the breakdown of the partial credit item per language.

Table 7.48: Macy and the Red Hen Item 12: Partial Credit Breakdown

Points	English Persons	% Obtained	Afrikaans Persons	% Obtained	isiZulu Persons	% Obtained
0	239	69	101	58	266	78
1	54	16	48	28	22	6
2	18	5	13	8	7	2
3	6	2	2	1	1	0
9 (not attempted)	31	9	9	5	46	13
Total	348	100	173	100	342	100

Very few learners were able to show extensive comprehension of Macy's feelings, traits and attitudes with an example thereof, with only six English learners, two Afrikaans and only one isiZulu being able to do so. Eight percent of Afrikaans learners were able to provide at least

one feeling, trait or attitude with one example, while only 5% and 2% of English and isiZulu learners, respectively, were able to provide similar answers. It is important to note that isiZulu had 13% missing data, which means learners provided incomplete answers or left the answer line blank. Learners may not be used to gathering information from across the text to answer a question and may rely on what they can remember from reading the text.

The last item that showed non-equivalence between the languages is *Macy and the Red Hen* item 15 and required learners to *Interpret and Integrate Ideas and Information*. The item required the reader to take perspective of the situation, as the reader needs to think what will X do next time. The item asked the following question:

What do you think the red hen will do next time Macy puts the hens in their cage?

This item only counted one point[67] and required learners to produce an answer that indicates that the red hen will behave the next time Macy puts the hens in their cage, which requires an understanding of how the pecking order works. Few isiZulu learners were able to say that they thought that the red hen would behave. Table 7.49 (following page) depicts a selection of Grade 5 learners' responses for *Macy and the Red Hen* item 15.

[67] *cf.* Appendix C for scoring guide.

Table 7.49: Macy and the Red Hen Item 15: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	The red hen will fight with Macy	Die rooi het gaan in die gevolg in die hok gaan.	Not attempted
2	I think she will go into the cage	Sy het weg gehardloop want sy wil nie in die hok is nie	Not attempted
3	I think the red hen will also go back into the cage	Die hen gaan inbeskof wees	Ngicabanga ukuthi sizongena ngoba sibone isikhova.
4	Because the other hens let the red hen boss them like that?	Macy gaan die hoe toe trek	Yingoba zifunwa zidliwe ngonyalga ozayo
5	Macy bent down and put out her hands	die ander sê kos afvat.	wabheka inqwaba yhokudla okwabuse bhentshini
6	She will get in the cage	Hy gaan volgende keer hom pik en nie meer so ombeskof met macy is nie	lapha egoli adonda oningi ayafunda okhule amadondanu amaeyahambaegoli eyefunda
7	She will never play tricks again	sy sal ook in die hok gaan.	kungani lezi ezinye izikhukazi zivama ukufine lwayilesi
8	She will go with her into the cage	Daan sal die hen weer vir macy rond sit	ukuthi sizophuma
9	They will all follow Macy when she tells them to	Sy gaan haar pik	Not attempted
10	Ran away and hide	Sy sal stil staan dat Macy haar kan vat	Ngicabanga ukuthi umesi uzobonisa izikhukhukazi ezibovu

As with most items discussed in this chapter, learners provided incomplete answers; for example, isiZulu learners 1 and 2. Learner responses captured in Table 7.45 show that some learners thought that the red hen would not change her behaviour or would harm Macy, see English learner 1 and Afrikaans learner 9. Other learners gave incomplete sentences, such as isiZulu learner 8, who wrote that she/he thought it (presumably the red hen) would go outside. Learners who provided vague responses also received a zero – for example Afrikaans learner 8. Some learners also provided generic answers, such as isiZulu learner 10 who wrote that she/he thought that Macy would look after the red hen. Most of the answers did not relate to the red hen's change in behaviour. This finding could mean that the learners did not fully understand what had transpired in the text – Macy moved to the top of the pecking order, and as a result, the red hen became obedient and went into the cage. The pecking order is explained at the beginning of the story; however, it was not explicitly stated or explained in the text at the end, which means that the learners should have inferred with Macy moving to the top of the pecking order, the red hen would have had to obey Macy. The explanation of the pecking order requires a global inference and weak readers, who are better at making local connections, may have struggled with making global connections. As with the previous item, learners may be used to more literal type questions that do not require them to read the text and interpret and integrate information from it. This item was also the second last item of the text and required learners to write quite extensive answers. As a result, learners may not have had sufficient test-stamina to complete the test.

7.6 CHAPTER SUMMARY

After analysing the PIRLS Literacy 2016 and PIRLS 2016 released literary texts namely, *The Pearl*, *Flowers on the Roof* and *Macy and the Red Hen* it is evident that neither the South African Grade 4 learners nor the Grade 5 learners performed well on these texts. All three of these texts displayed levels of differential item functioning (*cf.* Chapters Five and Six) amongst English, Afrikaans and isiZulu learners and were further explored to find possible reasons for the non-equivalence. This chapter specifically focused on providing answers to two subsidiary research questions. Sub-question 3 posed the following:

To what extent do the PIRLS 2016 and PIRLS Literacy 2016 assessment instruments have face validity?

All three of the released literary texts were examined to determine whether each text has face validity. Based on the experts' knowledge and their review of these texts, it was found that *The Pearl*, *Flowers on the Roof* and *Macy and the Red Hen* show sufficient evidence to be considered having face validity. The three literary texts were deemed by the experts to look like typical Grade 4 or 5 stories that learners may encounter. Even though A2, Z2 and Z3 made specific reference to certain words being uncommon or not used in everyday communication, these words formed a small constitution of the words used in the text. Furthermore, a main goal of narrative or literary texts is to use a rich lexical repertoire which includes mid and low frequency words. As such, the texts are considered to be appropriate for South African learners. The layout for each of the texts is exactly the same for each text whereby a piece of text is located at the same place in the ST as well as in TT. The illustrations used throughout each text are also the same across ST and TT and in the same position. Moreover, the items asked for each text are also located at the same place; for example, the answer to item 4 of *Macy and the Red Hen* is located on the second page in the same paragraph (in the booklet the page number is 5).

The sub-question 4 asked the following:

How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?

In order to answer sub-question 4, I evaluated the PIRLS Literacy and PIRLS texts and items by *macro*, *meso* and *micro* levels of equivalence. The *macro level* included looking at the genre, themes, main characters, layout of the text and length of the text, whereas the *meso level* focused on the narrative events as identified by the IEA, the different question items as well as the comprehension processes that accompanied each of the question items. The *micro level* entailed looking at the vocabulary used in each text, key expressions, literary devices used as well as readability measures such as average word length and words per sentence.

It is evident that even though the literary texts had a high number of items which displayed differential functioning (*cf.* Chapters Five and Six), the item non-equivalence did not seem to be due to poor translations per se, as some differences between experts tended to revolve around the translation of particular phrases or words and not paragraphs or larger units of text. Moreover, it did not appear to be the difficulty of the texts selected for the assessment, different cultural experiences nor overall reliability of the PIRLS Literacy and PIRLS assessments. The

non-equivalence of the tests may rather have been due to factors related to the knowledge and skills that learners bring to the reading process. These may include not properly engaging with the text, unfamiliarity with responding to texts at a higher level of comprehension, unfamiliarity with locating relevant information in the text that matches the question and problems with perspective taking. The aforementioned could be indicative of what is happening in the classroom (*cf.* Zimmerman, 2017).

In terms of pedagogy, it difficult to ascertain how teachers prepare and organise their lessons. There are many projects and government initiatives such as the Drop All and Read campaign, the Reading to Lead campaign, the Primary School Reading Improvement Programme as well as the Ramaphosa Reading Plan. More recently, South Africa conducted the Early Grade Reading Study (EGRS) that focused on improving reading outcomes in the Foundation Phase – this was done by techniques such as training and coaching. Even though these initiatives were launched, there is no clear indication of what happens in the classroom. Of concern is the quality of pre-service and in-service teacher training and their knowledge and ability to teach learners how to read (Spaull & Pretorius, 2019). Based on the results from ILSAs such as PIRLS and SACMEQ, the quality of teaching and learning is debatable. This study did not investigate the classroom environment or teacher training; however, it was important to refer to these as teacher knowledge, skill and experience in teaching reading is an important aspect in learner reading literacy (*cf.* Zimmerman, 2017).

Even though this study did not look at vocabulary challenges or decoding skills, other studies (*cf.* Hemphill & Tivnan, 2008; Netten, Droop & Verhoeven, 2011; Pretorius, 2016; Pretorius & Stoffelsma, 2017; Sibanda & Baxon, 2018; Snow, Burns & Griffin, 1998) have found that vocabulary is reliable predictor of a learner's future reading ability.

Any assessment, such as PIRLS, has specific aims and objectives that it has to fulfil. However, factors such as test fatigue, or cognitive fatigue may occur for reasons such as the time of testing, length of the test or the testing skill of the learners. A study found that test fatigue may negatively hamper the learners' ability to perform well on a test and by using breaks in between the test, it could lead to better performance on the test (Sievertsen, Gino, & Piovesan, 2016). Both PIRLS Literacy 2016 and PIRLS 2016 include a break between passages; however, the test fatigue for South African learners may rather be due to their inexperience with longer texts and providing long answers.

In terms of quality of translation, the IEA strives to conduct quality large-scale assessments that can be used to inform local or national education policies. As the IEA's studies, such as PIRLS, is conducted in numerous countries, it has stringent translation procedures in place that are updated for each cycle. During the translation process initial incidences of translation non-equivalence may be picked-up by the international language verifier and indicated on the NAF. However, this study found that both TTs included errors, some of which were minor such as where a word is cut, or the use of punctuation. However, one panellist felt that other translation errors were extensive, specifically the isiZulu translation of *Flowers on the Roof*. Based on the information gathered during this study, it would appear that the translator during PIRLS 2006 freely translated the ST and more stringent back-translation checking should have been employed.

-⌘-

CHAPTER 8

Phase Two: Informational Texts Qualitative Data Presentation and Results

8.1 ORIENTATION

This chapter presents the qualitative findings of Phase Two, specifically the PIRLS Literacy 2016 and PIRLS 2016 informational texts. As this study takes the form of a sequential explanatory mixed methods design, this chapter extends the Rasch analysis presented in Chapters Five and Six. For the purposes of this chapter, the findings of research sub-questions 3 and 5 are discussed. Sub-questions 3 and 5 ask the following:

3. *To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?*
5. *How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?*

This chapter follows the same structure used during Chapter Seven. It presents a multipronged analysis of the informational texts used during PIRLS Literacy 2016 and PIRLS 2016. Informational texts may use several different organisational elements to convey their message. Readers need a different set of skills in order to comprehend informational texts, as these texts are different from literary or narrative type texts (Hebert, Bohaty, Nelson & Brown, 2016). An information text is made up of five general organisational elements, namely description, problem and solution, compare and contrast, cause and effect, as well as sequence (Hebert, Bohaty, Nelson & Roehling, 2018). A text can comprise a combination of these elements, for instance, problem and solution usually assumes cause and effect as well as some sequence. As such, informational texts require learners to access their knowledge and understanding of text structure (structural awareness). Learners equipped with these skills are able to understand the organisation of the text and apply them to the text being read (Arfé, Mason & Fajardo, 2018). According to Ray and Meyer (2011) “structural awareness improves comprehension because it facilitates the construction of a coherent mental representation of text” (p.68). As coherence refers to the clear relationship among different textual ideas, it is deemed essential to text comprehension (Kintsch, 2004). Moreover, informational texts are usually more complex than narrative texts due to the “additional constraints on cognitive resources” (Hebert et al., 2018, p. 2166). It has been found that the comprehension skills of a reader are associated with the reader’s structural awareness (Ray & Meyer, 2011). As such, readers who possess stronger

comprehension skills usually have greater structural awareness and comprehend informational texts better than those who struggle with reading comprehension. Both narrative and expository texts make use of inferences and problem-solving skills; however, expository texts may focus on less familiar topics and are more complex.

This chapter makes use of various sources of evidence to answer the abovementioned sub-questions: passage mapping of each text; item difficulties; text analysis; workshop notes and questionnaires that deal with the equivalence of the translated texts. The questionnaires were completed by Afrikaans and isiZulu experts (see Section 7.2 for more detail).

Section 8.2 presents the analysis conducted of *African Rhinos and the Oxpecker Birds* followed by *The Green Sea Turtle's Journey of a Lifetime* (8.3). Both informational texts are explored by way of genre, purpose or message, summary and organisational elements. In addition to these textual features, this chapter also provides the word count, key vocabulary, average word and sentence length as well as the number of items for each *Processes of Comprehension*. Finally (8.4), a conclusion is provided whereby I answer the two sub-questions posed in this chapter.

8.2 ANALYSIS OF *AFRICAN RHINOS AND OXPECKER BIRDS*

This section presents an overview of the text and composition of *African Rhinos and Oxpecker Birds* (8.2.1). Thereafter, a summary of the four reading comprehension processes (8.2.2) is provided which indicates the level of difficulty of each item for the *African Rhinos and Oxpecker Birds* text. Sub-section 8.2.3 deals with the professional opinions of the quality as well as veracity of the translated instruments. The next sub-section (8.2.4) provides an in-depth look at the Grade 4 English, Afrikaans and isiZulu learner responses of the *African Rhinos and Oxpecker Birds* items. The ST as well as both TTs are presented in Appendix G. The above outlined structure will apply to the entire chapter. *African Rhinos and Oxpecker Birds* was developed specifically for PIRLS Literacy 2016 and was completed by South African Grade 4 learners (n = 832[68]).

[68] Four learners were removed from the analysis due to extreme scores. Extreme scores occur when a person has either 0% or 100% chance of getting all the items correct.

8.2.1 Overview of African Rhinos and Oxpecker Birds

African Rhinos and Oxpecker Birds is an expository text dealing with animals, used during PIRLS Literacy 2016. The purpose of this text is to inform readers about the mutually beneficial relationship between the African rhinoceros and the oxpecker bird. The mutually beneficial relationship revolves around the rhinos who have their ticks removed by the oxpeckers and the ticks, in turn, provide sustenance to the birds. The text also provides a description of their unique relationship as well as information about the animals. These aspects include what they look like, how much they weigh, what they eat, as well as strengths and weaknesses of each animal.

Two of the organisational elements were used in *African Rhinos and Oxpecker Birds* namely description, and compare and contrast.

- *Description* is used when the author wishes to provide information about a specific topic (Hebert et al., 2016). The topic of interest is then described in order for the reader to become aware and understand the specific topic or certain attributes thereof. Description is also less structured than other organisational elements such as compare and contrast (Ray & Meyer, 2011).
- *Compare and contrast* is employed when the author wants to compare and contrast different perspectives, concepts or ideas (Hebert et al., 2016). It is also used to draw attention to similarities and differences (Arfé et al., 2018). Compare and contrast is more structured, as it is usually organised into two logical arguments.

In terms of *description*, information is provided about two main concepts or ideas contained in the text. During description, the author also uses explicit signalling of information through the use of headings. The first main concept is to provide information about rhinoceros. Supporting ideas include:

- Rhinos are famous for their horns.
- There are people who think that rhino horns can be used as a medicine.
- Consequently, rhinos are often illegally killed for their horns.
- Rhinos eat grass and leaves.
- Sometimes when the rhinos are eating, ticks jump onto the rhino's skin.
- Ticks irritate the rhinos' skin when they feed on the rhinos' blood (IEA, n.d.-d).

The second main concept, as part of *description*, is to provide information about the oxpecker bird. Supporting ideas include:

- Oxpeckers are small brown birds.
- They have a wide bill, stiff tail and sharp claws.
- Oxpecker birds eat ticks that live on rhinos (on their skin).
- They prefer to eat ticks that are full of blood.
- They can eat nearly 100 blood-filled ticks each day (IEA, n.d.-d).

The main idea of the *compare and contrast* element of the text is to show the different strengths of each animal, which allow them to depend on each other – indicating a mutually beneficial relationship. Supporting ideas include:

- The oxpecker birds benefit from the rhinos as the rhinos provide them with food (ticks).
- The rhinos benefit from the oxpecker birds as the birds remove the ticks from the rhinos.
- The rhinos have poor vision, consequently they cannot easily spot danger.
- The oxpecker birds have sharper vision and when they see danger, they make loud noises to alert the rhinos (IEA, n.d.-d).

The above concepts and ideas provide an overview of what was included in the text. These concepts and ideas occur in each language. The IEA also indicated a limited set of key vocabulary, which was used in the text, as shown in Table 8.1, which presents key vocabulary as well as the frequency with which these words occurred in the ST and TT.

Table 8.1: IEA Key Vocabulary for African Rhinos and Oxpecker Birds across Languages

IEA Vocabulary	**Frequency English**	**Frequency Afrikaans**	**Frequency isiZulu**
valuable	1	1	1
horn	4 (horn, horns)	4 (horing, horings, neushoring)	4 (uphondo, ngezimpondo, yezimpondo)
ticks	8	10	8
enemies	2	2	2

The IEA selected these words as they indicate specific animals, insects, behaviours or descriptions in *African Rhinos and Oxpecker Birds*. Remarkably, the words '*African*', '*rhino*'

and '*oxpecker*' were not used in the body of the text. The word '*African*' was only used in the title of the text. It could be that the author focused specifically on rhinos[69] in Africa and their relationship with the oxpeckers. What is more, three of the key words identified by the IEA refer to peripheral issues that do not focus on the main topic, which is the mutual beneficial relationship between two animals. Only the word '*ticks*' is included which forms part of the mutual relationship. The word '*valuable*' was only used once in the ST as well as in Afrikaans ('*waardevol*') and isiZulu ('*obalulekile*'). It was used to describe the worth of rhino horn as it is used as an ingredient in medicine in some countries. The word '*horn*' was used twice in the text, but when the plural form ('*horns*') is included, the word appears four times in the ST. In the Afrikaans and isiZulu versions, the singular form '*horing*' and '*uphondo*', respectively, appears twice. The plural of horn was also used in both Afrikaans and isiZulu but due to the differences in orthography[70] as explained in Section 7.2, the root word is accompanied by a prefix in isiZulu, '*ngezimpondo*' (with/for their/its horns) and '*yezimpondo*' (of its horns) were used. Vowel coalescence '*nga*' + '*e*' occurs as the morpheme becomes '*nge*' when it is added to the plural of rhino '*mpondo*' to indicate possession. In Afrikaans, '*horings*' (horns) and the compound noun '*neushoring*' (nose + horn) were used.

The *LexTutor VocabProfile* software tool was used to analyse the vocabulary used in the corpus. Table 8.2 presents the English word profile of *African Rhinos and Oxpecker Birds*. As mentioned in Chapter Seven, the K-levels identify word frequency levels which in turn serve as an indicator of text difficulty (Laufer, 2013; Laufer & Ravenhorst–Kalovski, 2010).

Table 8.2: Word Profile of African Rhinos and Oxpecker Birds

Word frequency level (English)	Occurrence	% of Tokens	Cumulative Tokens (%)	Examples
K1	358	80.6	80.6	animals, colour, weight
K2	24	5.4	86.0	enemies, sharp, tiny
K3	1	0.2	86.2	sensitive
K4	20	4.5	90.7	horn, insects, kilograms

[69] There are five species of rhino, two of which can be found in Africa with the remaining three species being located in Asia (World Wildlife Fund, 2020).
[70] Due to the agglutinative nature of isiZulu, some key words do not occur as individual words but are embedded in complex morpho-grammatical structures.

Word frequency level (English)	Occurrence	% of Tokens	Cumulative Tokens (%)	Examples
K5	3	0.7	91.4	claws, grams, hissing
K6	1	0.2	91.6	itchy
K8	25	5.6	97.2	bloated, rhino, rhinos
K11	3	0.7	97.9	rhinoceros
Off list	0	0	≈100	

The word profile of *African Rhinos and Oxpecker Birds* only reached the 95% threshold[71] at K8 – meaning that most of the words used in this text range between high and mid frequency words (Laufer, 2013; Laufer & Ravenhorst–Kalovski, 2010). In contrast, the majority of the narrative texts comprised words between K1 to K3 (*cf*. Chapter Seven). As this text is expository in nature, it contains diverse and technical vocabulary not often used in narrative texts. This type of text enables the teacher and learner to engage in more abstract language use in order to make connections between ideas or concepts in a text (Pentimonti, Zucker, Justice & Kaderavek, 2010).

The word frequency level shown in Table 8.2 indicates the lexical threshold, which refers to the words known by readers in order for them to understand the text (Laufer; 1989; Nguyen & Nation, 2011; Schmitt et al., 2017). Words that fall between K1 and K3 are considered high frequency words, and if 95% of the words fall between these levels, the text is then considered to be easy to comprehend (Laufer, 2013; Laufer & Ravenhorst–Kalovski, 2010). As 95% of the words used in this text fall between the K1 and K8 levels, in other words between the 1 000- and 8 000-word frequency levels, some of the words may not be known to Grade 4 learners.

Words such as '*bloated*' and '*rhinoceros*' were identified at the K8 and K11 frequency levels, respectively. The adverb '*bloated*' was used once in the text where it referred to the tick after it has fed on the rhino. The noun '*rhinoceros*' was used three times in the text, but the author made use of '*rhino*' sixteen times in the ST text (15 times in Afrikaans and 17 times in isiZulu). The use of '*rhino*' so often should provide the reader with enough exposure to learn the word.

[71] *cf*. Page 242 for an explanation of the ratio of low, mid and high frequency words required for reading with comprehension.

Although '*rhino*' is located at K8, based to the BNC-COCA[72], within the South African context, the word '*rhino*' would not be such an unfamiliar word as it forms part of the big five animal group, there is rhino awareness in the country due to rhino poaching and it is sometimes included in the news. *African Rhinos and Oxpecker Birds* includes a higher percentage of mid frequency (7%) words, but it also includes a large percentage (86%) of high frequency words.

The next table depicts the readability analysis of *African Rhinos and Oxpecker Birds* for each of the selected languages.

Table 8.3: Readability of African Rhinos and Oxpecker Birds per Language

Readability Features	English	Afrikaans	isiZulu
No of words	444	448	294
Pages/paragraphs	19*	19*	19*
Sentences per paragraph	3,4	3,5	3,5
Words per sentence	13,4	13,1	8,4
Mean word length	4,3	4,6	7,8
Passives [a]	6%		
Flesch reading ease [b]	77,7		
Grade level [c]	5,7		

[a, b, c] These features can only be calculated for English.

*The number appears to be higher as it includes the words used in a table.

Table 8.3 presents an overview of the text length of the ST and TT. As anticipated, the isiZulu[73] word count was much lower compared to the English and Afrikaans versions. Even though *African Rhinos and Oxpecker Birds* might seem long for South African Grade 4 learners as it exceeds the national recommendation of 150 – 200 words (DBE, 2011), it is shorter than the narrative texts. However, according to the reading norms established by Hasbrouck and Tindal (2017), English Grade 4 learners, at the 50th percentile, should be able to read 112 words correctly per minute (WCPM) by the middle of the year. By the end of the formal school year,

[72] The BNC-COCA is a combination of the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA).

[73] *cf.* See section 7.2 for explanation about orthographies about the different languages.

they should be able to read 11 additional WCPM. Therefore, *African Rhinos and Oxpecker Birds* should take the average English Grade 4 learner approximately four minutes to read. However, due to the layout of the text leaners may take longer to read and answer the questions, as the text is divided into sections, indicating that learners need to read a portion of the text and answer the accompanying items before continuing to the next section of the text (*cf.* Appendix G). Sub-section 8.2.4 provides a discussion of the items that showed non-equivalence between the ST and TT.

Note that due to the table information extracted and included in the readability formula, the number of paragraphs were increased. Even so, the number of paragraphs and pages were the same across the different languages. Furthermore, the number of words per sentence ranged between the languages, with English and Afrikaans reaching 13.4 and 13.1 words per sentence respectively while isiZulu reached 8.4. The mean word length showed a similar pattern whereby Afrikaans (4.6) and English (4.3) had a lower mean word length compared to isiZulu (7.8), which confirms the occurrence of longer words in the conjunctive orthography of isiZulu. In all three texts, there is evidence of equivalence based on the overall paragraph and page structure.

8.2.2 Processes of Comprehension of African Rhinos and Oxpecker Birds Items

This section explores the four *Processes of Comprehension* in terms of question item difficulty for *African Rhinos and Oxpecker Birds* and provides partial evidence whether non-equivalence is present for this informational text. This text included 17 items, mostly worth one point. The total mark allocation for this text is 19. Of the 17 items, ten items were CR and the remaining items were MC. The CR items counted 12 points while the MC items counted seven points. The following table depicts the item difficulty of *African Rhinos and Oxpecker Birds* per language keeping in mind the four *Processes of Comprehension*.

Table 8.4: Item Difficulty of African Rhinos and Oxpecker Birds across Languages

Processes of Comprehension	English Items	Afrikaans Items	isiZulu Items
Focus on and Retrieve Explicitly Stated Information	1, 4, 5, 7, 10, 11, 12, 14, 16	1, 4, 5, 7, 10, 11, 12, 14, 16	1, 4, 5, 7, 10, 11, 12, 14, 16
Make Straightforward Inferences	6, 8, 13, 15	6, 8, 13, 15	6, 8, 13, 15

Processes of Comprehension	English Items	Afrikaans Items	isiZulu Items
Interpret and Integrate Ideas and Information	9, 17	9, 17	9, 17
Evaluate and Critique Content and Textual Elements	2, 3	2, 3	2, 3

More than half (9) of the items are located at the literal level, *Focus on and Retrieve Explicitly Stated Information*. Four of the items tested learners' ability to *Make Straightforward Inferences*. The remaining four items were at the inferential level, of which two items tested how well learners could *Make Straightforward Inferences* and two items tested whether learners could *Evaluate and Critique Content and Textual Elements*. The division of the items according to the four *Processes of Comprehension* conforms to the guidelines set out by the IEA, where majority of the items could be at the *Focus on and Retrieve Explicitly Stated Information* level (Mullis & Martin, 2015). It is clear, based on Table 8.4 that all the question items remained at the same comprehension level for the ST and the TT, so none of items' difficulty changed from the ST to the TT.

8.2.3 Professional Opinion on Translated Texts

The majority of the Afrikaans (A1 and A2) and isiZulu (Z1 and Z3) experts agreed that the text is mostly written in a simple and straightforward style with evidence of some complex language structures in places. All of the experts agreed that the translation was overall well done but with a few errors. Z2 and Z3 noted that the isiZulu version used two different words for '*sucks*', namely '*-munca*' and '*-ncela*'. Z2 suggested that only one of these words be used in the text while Z3 explained that "there may be dialectal preferences". The text made use of high and mid-frequency words, with only one word at the K11 frequency level. However, both A1 and A2 as well as Z2 and Z3 indicated that some words used in the text may not necessarily be known to Grade 4 learners. A2 specifically mentioned the following words: '*renoster*' (rhinoceros); '*renostervoël*' (oxpecker bird); '*afhangend*' (depending on); '*bosluise*' (ticks) and '*ontslae*' (get rid of). The word '*rhinoceros*' is deemed a K11 and '*rhino*' a K8[74] frequency word, both A1 and A2 agreed that the Afrikaans version, '*renoster*' (rhinoceros), could also be

[74] Based on the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) lists.

at a similar level as '*rhino*'. The isiZulu word for *rhinoceros* is '*obhejane*'. Even though the words '*ox'* and *'peck'* are considered a K1 words, meaning that it is part of the first thousand words, in Afrikaans the translation becomes a compound noun ('*renostervoël*') that is made up of rhinoceros ('*renoster*') and bird ('*voël*'). In isiZulu *oxpecker* is known as '*ihlalankomo*'. Z3 stated that "the Zulu word for oxpecker '*ihlalanyathi/ihlalankomo*' is not well known, but that should not pose any problem since the illustrations and descriptions are adequate." The K-levels indicate that '*tick(s)*' is within the 4 000 (K4) range and should be known to English Grade 4 learners whereas A2 stated '*bosluise*' might not be generally known. However, A1 disagreed as '*bosluise*' is a word that learners would have picked up early on and domestic pets, such as cats and dogs, can be seen almost everywhere and are carriers of ticks. Furthermore, learners in more rural areas may have more exposure to the word 'tick' as opposed those living in urban areas. In addition, words such as '*afhangend*' and '*ontslae*' are correctly used in the text and should not pose a hindrance to the Afrikaans Grade 4 learners. A2 further commented that although the text itself is not complicated, the register of the text is unfamiliar and "dit is beslis nie hoe kinders sal praat nie" (children would not talk that way). Learners at the Grade 4 level may still be developing conversational fluency or BICS (Cummins, 2008), especially if they have transitioned from the home language to the LoLT. With this foundation of basic interpersonal communication skills, learners during the Intermediate Phase will begin to develop CALP, which deals with skills essential to academic literacy such as listening, reading, writing and speaking in the educational context and relies strongly on skills of written language (Cummins, 2000). Thus, the learners have to learn how to read and interpret a variety of registers in different contexts. As such, learners are exposed to new words that are scientific and technical. *African Rhinos and the Oxpecker Birds* make use of mid and higher-level frequency words, but this is expected due to the nature of the text. Z3 noted that "the language register is at the desired level and more or less corresponds with the Enlish [sic] version".

With respect to cultural equivalence, the ST and TT should be similarly interpreted and understood by the English, Afrikaans and isiZulu learners. Z2 mentioned that "the text will be equally understood in the three languages"; however, Z1 cautioned that some learners may not "interpret [the] information in the same way as some who have no knowledge of the chosen animals/birds". A2 agreed and went on to say that the English version might be easier due to the easier vocabulary. Although both Z1 and A2 make good points, the purpose of the text is to inform the reader about the mutually beneficial relationship between the rhinos and the

oxpecker birds. Z3 explains, "the speakers of the different languages should have an equal understanding of the text." The author introduces the two animals, namely '*rhinoceros'* and '*oxpecker birds'* at the beginning of the text and repeatedly uses the two animals in order for the reader to gain exposure to them. The word '*rhinoceros*' is used 3 times while '*rhino*' was used 16 times and '*oxpecker*' was used 5 times.

In terms of the construct relevance of *African Rhinos and Oxpecker Birds*, the majority of the Afrikaans and isiZulu experts regard the text as being relevant. Z3 argued that it is "not entirely" construct relevant as the text includes factual inaccuracies; for example "ticks do not jump, they climb onto their host from the [g]rass blades or tips of the branches or leaves." Nevertheless, the South African Grade 4 learners should be able to cope with the structure and information contained in the text as the language is "clear and simple" (Z2). A2 explained that even though the text is construct relevant, attention should be paid to the use of double diminutives in the Afrikaans text; for example, "*klein diertjie*" (tiny small animal). It should be noted that even though the text focuses on two animals that live in Africa, you may not always see the oxpecker bird on the rhino and that most South African learners may not have seen a rhino or oxpecker in real life due to the large geographical differences of the country. However, the rhinoceros is iconic in South Africa and learners would have seen pictures. A2 also mentioned that the poaching of rhinos is viewed in a negative light as black and white rhinos are endangered. All experts agreed that the illustrations used in the text are helpful and informative and would likely help the Grade 4 learners understand the text more clearly. Z2 commented, "the text is about rhinos, oxpeckers and ticks. Illustrations were also about the three animals". A1 and A2 agreed that the illustrations are of good quality – A2 further mentioned that it is because photographs were used of the animals. Even though the photographs were only available in black and white, it should have no negative bearing on the learners' understanding of what each animal looks like, since the photographs were still clear.

The experts were also asked whether the *African Rhinos and the Oxpecker Bird* text was a typical informational text that Grade 4 learners may encounter. The experts were divided on this matter as Z1 relies on learners' possible lack of prior knowledge when it comes to the two animals, yet Z2 commented "illustrations, the language used and questions asked were suitable for Grade 4 learners". It should be noted that there are not many information texts available in African languages, as such, the learners may not be familiar with the isiZulu informational texts but rather with the English ones. A1 also believed that the text is interesting and focuses

on one of South Africa's big five animals, the rhinoceros. The layout of *African Rhinos and Oxpecker Birds* is the same across the ST and TT. The text consists six pages of text that include headings as well as six pages with questions. As with the literary PIRLS Literacy 2016 text, a large, easy to read font was used. Each page dealt with a specific heading and text and these can be found at the same places in the English, Afrikaans and isiZulu versions. The photographs of the rhino, oxpecker and tick were positioned in a specific section of the text. The layout and photographs used add to the face validity of *African Rhinos and Oxpecker Birds*. The length of the text exceeds the word limit stipulated by the national curriculum (*cf.* DBE, 2011) yet this should not be an obstacle for the learners as the layout of the text and question items is designed to help lessen the cognitive load.

Based on the experts' opinions, the *Processes of Comprehension* did not change from the ST to the TT and the text adequately covered each of the four processes. Z1 explained that "comprehension takes place and answers of questions will be answered as is from the story without additional information of their own to add". Z2 and Z3 commented that the isiZulu version of the text may be misleading and item 13 may confuse the learners. Item 13 asked: "Why do oxpeckers especially like to eat ticks?". The distractors include: a) because the ticks have been sucking blood; b) because there are many ticks on each rhino; c) because the ticks have been eating leaves; and d) because the ticks are tiny and easy to eat. Note that this item required learners to make a straightforward inference from the text – the oxpeckers like to eat ticks because they (the ticks) have been sucking blood. Z2 argues that the correct answer was not provided as the text stated 'but their favourite food is blood so they prefer the ticks that are full of the rhino's blood'; as such, the learners would have selected distractor A as the most plausible answer. Most of the experts agreed that the question items tested what they intended to measure. Z1 highlighted the fact that the lack of information about wildlife in isiZulu may have some influence on the learners. A2 noted that some Afrikaans items could have been more clearly phrased as they could be confusing for the learners. For example, question 2 and 3 could have used sentence starters in order to avoid repetition of words in the distractors.

8.2.4 Per-item Analysis for African Rhinos and Oxpecker Birds

The responses to the *African Rhinos and Oxpecker Birds* items are presented in this sub-section. It includes the percentage of incorrect responses; missing percentage per item as well as the items that displayed differential functioning between English, Afrikaans and isiZulu languages.

Table 8.5: African Rhinos and Oxpecker Birds: Percentage of Grade 4 Learners who Incorrectly Answered Items per Language

Item No	Processes of Comprehension	English			Afrikaans			isiZulu		
		N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing
1	Focus on and Retrieve Explicitly Stated Information	336	41	4	198	50	3	287	64	2
4*	***Focus on and Retrieve Explicitly Stated Information***	***336***	***27***	***3***	***198***	***54***	***2***	***287***	***69***	***2***
5*	***Focus on and Retrieve Explicitly Stated Information***	***336***	***36***	***3***	***198***	***51***	***2***	***287***	***74***	***1***
7	Focus on and Retrieve Explicitly Stated Information	336	18	2	198	21	1	287	29	1
10*	***Focus on and Retrieve Explicitly Stated Information***	***336***	***37***	***2***	***198***	***24***	***7***	***287***	***33***	***3***
11	Focus on and Retrieve Explicitly Stated Information	336	47	6	198	49	6	287	52	6
12*	***Focus on and Retrieve Explicitly Stated Information***	***336***	***46***	***5***	***198***	***63***	***4***	***287***	***53***	***3***
14	Focus on and Retrieve Explicitly Stated Information	336	27	8	198	36	7	287	44	9
16*	***Focus on and Retrieve Explicitly Stated Information***	***336***	***47***	***8***	***198***	***67***	***5***	***287***	***73***	***8***
6*	***Make Straightforward Inferences***	***336***	***38***	***3***	***198***	***33***	***6***	***287***	***33***	***1***
8	Make Straightforward Inferences	336	44	2	198	46	6	287	53	1
13*	***Make Straightforward Inferences***	***336***	***41***	***7***	***198***	***53***	***8***	***287***	***35***	***8***
15	Make Straightforward Inferences	336	49	6	198	54	7	287	57	6
9	Interpret and Integrate Ideas and Information	336	65	3	198	74	3	287	77	2
17	Interpret and Integrate Ideas and Information	336	91	9	198	90	10	287	91	9
2*	***Evaluate and Critique Content and Textual Elements***	***336***	***36***	***2***	***198***	***40***	***6***	***287***	***62***	***2***
3*	***Evaluate and Critique Content and Textual Elements***	***336***	***62***	***4***	***198***	***57***	***6***	***287***	***62***	***3***

*Indicates items that displayed DIF.

A sample of the learner responses for each constructed response item is also included that indicates not attempted, partially correct, correct and incorrect responses. Table 8.5, sorted according to the four *Processes of Comprehension*, depicts the total number of learners who incorrectly responded to the items. Only items that displayed non-equivalence between the ST and TT is explored. Of the 17 items, nine indicated non-equivalence, which include items 2, 3, 4, 5, 6, 10, 12, 13 and 16. Of the nine items, items 2, 3, 6, 10 and 13 are MC. The remaining four items are CR.

Of these nine items, five were surprisingly literal type items (4, 5, 10, 12 and 16) that required learners to obtain the answer from the text. Item 4 provided a mixed outcome as the learners who completed the test in English found the item relatively easy with less than a third (27%) providing incorrect responses, whereas just over half (54%) of Afrikaans learners and almost two-thirds (69%) of isiZulu learners provided incorrect responses. The missing values for each language ranged between two and three percent. A similar pattern is observed for item 5 where 36% of the English learners provided an incorrect response while more than half of the Afrikaans (51%) and isiZulu (74%) learners provided incorrect answers. Item 10 showed a different pattern as both English (37%) and isiZulu (31%) learners obtained a higher percentage of incorrect answers than those who completed the test in Afrikaans (24%). Although the missing values of item 10 differ, seven percent of Afrikaans learners did not attempt the question while three and two percent of isiZulu and English learners, respectively, did not attempt the item. In terms of item 12, the majority (63%) of Afrikaans learners struggled with this item, followed by isiZulu (53%) and English (46%) learners. Missing values range between three and five percent. Item 16 was more difficult for isiZulu learners as most (73%) provided incorrect responses, followed by Afrikaans (67%) learners. The item was somewhat easier for English learners as less than half (47%) provided incorrect responses.

Two items (6, 13) tested learners' ability to make inferences based on what they had read. Item 6 was relatively easy for all languages, with 38% of learners who completed the test in English providing incorrect responses. Interestingly, this item was slightly easier for both Afrikaans (33%) and isiZulu (33%) learners as they provided fewer incorrect responses. Item 13 was somewhat more difficult for Afrikaans learners as 53% provided incorrect responses while 41% and 35% of English and isiZulu learners, respectively, provided incorrect answers.

The remaining two items (2, 3) tested a higher order skill, where learners had to evaluate and critique information contained in the text. Most of the isiZulu learners (62%) found item 2 difficult as they provided incorrect responses, followed by Afrikaans (40%) and English (36%) learners. The missing values were low; seven percent of Afrikaans learners did not attempt the question while only two percent of English and isiZulu learners did not answer the question. In terms of item 3, the learners' percentage of incorrect responses was similar to Afrikaans learners who found the item slightly easier as 57% provided incorrect responses compared to the English (62%) and isiZulu (62%) learners.

Rasch analyses (*cf.* Chapter Six) of the *African Rhinos and Oxpecker Birds* items highlighted non-equivalence across the three languages. As already explained in Chapter Seven, a selection of learner responses in each language was captured in order to further investigate possible reasons for the non-equivalence. If the selected item was MC, the entire sample's responses were included in this sub-section.

To reiterate, *African Rhinos and Oxpecker Bird* item 2 asks the following question:

What does the picture in the big circle help you understand?

a. how oxpeckers fly
b. what oxpeckers sound like
c. an oxpecker's nest
d. what oxpeckers look like* (correct answer)

Table 8.6 presents the number of persons who selected each of the distractors.

Table 8.6: African Rhinos and Oxpecker Birds Item 2: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	41	12	25	13	86	30
B	50	15	22	11	39	14
C	29	9	33	17	52	18
D*	209	62	107	54	104	36
9 (not attempted)	7	2	11	6	6	2
Total	336	100	198	100	287	100

*Correct response (distractor D)

Almost two-thirds (62%) of learners who completed the test in English selected distractor D, while 54% of Afrikaans learners selected it. Slightly more than one-third (36%) of isiZulu learners selected the correct option. Even though this item appears to be more difficult for isiZulu learners, the text prompts an answer for item 2 as it refers to the picture next to the item, specifically the picture with the circle around it. Furthermore, the text explains that the oxpecker bird and the rhino look very different to one another and that the oxpecker is a tiny bird. The bird can be seen sitting on the back of a rhino. Learners could refer to the picture to determine why focus was placed on the oxpecker bird, which shows what the oxpecker bird looks like. Based on the results, 30% of isiZulu learners found distractor A very tempting. The isiZulu learners may not have understood the question or what the picture with the 'big circle' referred to. Z3 commented, that '*ihlalankomo*' (oxpecker) "is not well known, but that should not pose any problem since the illustrations and descriptions are adequate". However, the same argument could be made for the other two languages, as '*oxpecker bird*' is not generally well known to learners. In essence, the text and illustrations provide sufficient information for the learners to link '*renostervoël*' or '*ihlalankomo*' to the bird and that the picture in the 'big circle' shows what the bird looks like. Learners who are less skilled may have selected the most salient feature associated with birds – that they fly.

Item 3 of *African Rhinos and Oxpecker Birds* was the next item that displayed differential functioning and required learners to *Evaluate and Critique Content and Textual Elements*. To recapitulate, the item asked the following question:

Why does the writer tell you about the elephant?

a. to show that elephants live near rhinos
b. to show that the rhino is very big* (correct answer)
c. to show that elephants have oxpeckers
d. to show that rhinos and elephants eat the same food

Table 8.7 presents the number of persons who selected each of the distractors.

Table 8.7: African Rhinos and Oxpecker Birds Item 3: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	92	27	50	25	70	24

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
B*	116	35	74	37	101	35
C	55	16	18	9	51	18
D	61	18	45	23	56	20
9 (not attempted)	12	4	11	6	9	3
Total	336	100	198	100	287	100

*Correct response (distractor B)

This item appears to be equally difficult for the Afrikaans (37%), English (35%) and isiZulu (35%) learners as only slightly more than a third of each language selected the correct option (distractor B). However, a large proportion of learners found distractor A tempting. Nonetheless, the text explains that the rhino is a very large land animal and only the elephant is larger (than the rhino). Based on the question posed to the learners, the learners were supposed to come to the conclusion that the reason for including the elephant was to indicate how big the rhino is, in other words to provide the readers with a comparison of size. It did not mention the proximity in which these two large land animals lived, merely that the elephant is larger than the rhino. Learners who selected distractor A might have relied on their own background knowledge of these two animals in order to answer this question.

Item 4 is a CR type item which showed differential functioning between the three languages. To reiterate, it asked the following question:

Look at the table. How much does an oxpecker weigh?

African Rhinos and Oxpecker Birds item 4 is a literal item that tested learners' ability to *Focus on and Retrieve Explicitly Stated Information*. Since this is a CR item, the South African Grade 4 learners had to look at the table provided in the text and find the answer within the table and provide it as the answer. This item is worth one point and the only acceptable answer[75] is 60 grams. Learners had to indicate the unit of measurement, that is, grams, in order to obtain the mark. Table 8.8 presents a selection of ten Grade 4 learners' responses to *African Rhinos and Oxpecker Birds* item 4.

[75] *cf.* Appendix C for scoring guide.

Table 8.8: African Rhinos and Oxpecker Birds Item 4: Grade 4 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	live to ge ther an help each other	60 gm.	idloizi mbuzane nemikaza
2	Not attempted	Renoster in die beroemdste vir hulle groot	isisindo -60
3	They are still hunted for their horns	Die renoservoel dyt seer	amagremu angama 60
4	Still, rhinos ar great dager from hunters	300 kilogram.	amagremu angama-60
5	60 grams	60 gram	Eqinisweni igama likabhejane ngesilungu elithi rhinoceros
6	60 grams	60 gram	isisindo: amagremu-angama-60
7	An oxpecker weighs 60 grams	Renosters is die beroemdste vir hulle groot horings	ayi-3600.
8	it's is great at hente an it's is danger	Ealemebomedieisemebi?	isisindo esincane
9	Rhinos are most famous for	Renosters is die beroemdste vir	obhejane badume kakhulu ngezimpondo zabo ezinkulu.
10	Rhinos are most famous for their large horns	60 gram.	Eqinisweni

Note. The colours indicate the correctness of each answer. Incorrect reponses are red; correct responses are green; and nonsensical responses are in purple. The scorers were instructed not to focus on spelling or grammar during the scoring process but rather the content thereof.

Item 4 is one of five literal type items that displayed differential functioning between the languages. From the learner responses in Table 8.8, only one learner (English learner 2) did not attempt the item meaning that the learner left the answer line blank. As this question merely wanted the weight of the oxpecker, it did not require learners to provide full sentences; however, English learner 7 provided a full sentence. A few learners provided other pieces of irrelevant text as their answer, such as Afrikaans learner 7 and isiZulu learner 9, while others referred to the weight of the rhino (isiZulu learner 7). There is also one instance where a learner wrote a nonsensical response (Afrikaans learner 8). This learner copied letters randomly and added a question mark at the end of the 'word'.

Although the answer is explicitly stated in the text, the majority of learners found this item difficult. The majority of Afrikaans (54%) and isiZulu (69%) learners struggled with this item while most English (27%) learners were able to copy the correct answer from the text (*cf.* Table 8.5). A possible reason for this finding could be partly due to learner familiarity with tables – in other words, how to read a table. Learners may have looked elsewhere for a possible answer in the text. Another possible reason for the differential functioning could be that Afrikaans and isiZulu learners may not be as familiar with expository texts as the English learners. There are very few isiZulu informational texts available as most texts available are literary based texts (Makalela & Fakude, 2014; van Staden et al., 2016). Furthermore, learners who provided incorrect responses might not be skilled in reading and writing, such as the Afrikaans learner who scribbled letters as to mask their lack of reading literacy skill from their peers. Afrikaans learner 3 wrote '*dyt seer*' instead of '*byt seer*', which is an indication of confusion with the directionality of letters b and d; however this mirror writing phase usually passes around Grade 1 or 2.

The next item that showed differential functioning during the Rasch analysis, is item 5. To recapitulate, item 5 asked the following question:

What is the height of a rhino?

Item 5 is a literal CR type item which tested whether the learners could retrieve explicitly stated information from the text. The learners were required to look at the table, as with item 4, and find the correct answer within the table. This item counted one point and the only acceptable

answer is 2 metres – the unit of measurement must be included in order to obtain the point (*cf.* Appendix C for the scoring guide).

Table 8.9 (following page) provides a selection of Grade 4 learner responses to item 5. Similar patterns are observed for item 4 and 5, for instance English learner 2 did not attempt to answer the question while Afrikaans learner 8 provided a nonsensical response with a question mark. English learner 7 provided a sentence as an answer even though it was not required. Some learners did not indicate the unit of measurement and forfeited the point, see for instance isiZulu learners 2 and 7.

Table 8.9: African Rhinos and Oxpecker Birds Item 5: Grade 4 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	are most famous for their large horns.	20 sentimeter.	untsihani namaqqbunga
2	Not attempted	Hoewel renosters deur	ubude-2
3	2 metres	Dit eet sy vlys	amamitha angu-2
4	Rinos are most famous for their large horns	2 meter	amamitha ama 2
5	2 metres	2 meter	kod wa lokno akulona iqiniso
6	2 metres	2 meter	Ubude amamitha ama 2
7	The height of a rhino is 2 metres	Renosters is egter steeds in groot gevaar as gevolg van jagters	ama-2
8	they are still hunted their horns	yoamebemegetgneget?	mude kakhulu
9	They are still hunted	horings in Engels beteken die	ngokukhuluphala komzimba
10	In fact then ame rhinoceros mean nose horns	3600 kilogram	obhejdlo

This item focuses on the same table as item 4; however, it required learners to find the height of the rhino. As the word *rhinoceros* is classified as a K11 level word, it might be that learners did not make the connection between *rhino* and *rhinoceros*. Nevertheless, *rhino* and *rhinoceros* were introduced early in the text and learners should have become familiar with the words. Similar to item 4, the answer was explicitly stated in the text, albeit in a table, and it might have been difficult for the learners to navigate the table. More than half of the Afrikaans (51%) and isiZulu (74%) learners found this item difficult whereas 36% of English learners were not able to provide the correct response (*cf.* Table 8.5). As this item is similar in terms of *Processes of Comprehension* to item 4 and also makes use of the same table, the possible reasons for the incorrect responses might be the same (refer to item 4).

African Rhinos and Oxpecker Birds item 6 was the next item that displayed differential functioning between the languages. The item required learners to *Make Straightforward Inferences* based on the information contained in the text. To repeat, item 6 asked the following question:

Why do hunters want to kill rhinos?

a. Rhinos are too dangerous.
b. Hunters want rhino meat.
c. Hunters want rhino horns.* (correct answer)
d. There are too many rhinos.

Table 8.10 presents the number of persons who selected each of the distractors.

Table 8.10: African Rhinos and Oxpecker Birds Item 6: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	66	20	29	15	57	20
B	38	11	18	9	26	9
C*	199	59	121	61	189	66
D	22	7	19	10	11	4
9 (not attempted)	11	3	11	6	4	1
Total	336	100	198	100	287	100

*Correct response (distractor C)

The majority of the learners were able to select the correct answer from the four distractors. Interestingly, 66% of learners who completed the test in isiZulu selected correct distractor C, followed by Afrikaans (61%) and English (59%) learners. Distractor A attracted some of the learners' attention – learners who selected this option may have relied on their prior knowledge of rhinos or could have fixated on the illustrations that shows large, pointy horns and deduced that rhinos are dangerous. However, the text clearly states that rhino horns are considered by some to be valuable and that rhinos are hunted for their horns. Based on this information, learners were required to make the inference that the reason that hunters kill rhinos is because of their horns. Furthermore, most learners attempted to answer this question – only 1% of isiZulu learners did not attempt to answer this question while 3% and 6% of English and Afrikaans learners, respectively, did not attempt this question.

Item 10 of *African Rhinos and Oxpecker Birds* also showed differential functioning between the three languages. As this is a literal question item, learners were required to read the text and find the answer in the text. To recapitulate, item 10 asked the following question:

What do ticks need to live?

a. trees
b. rhinos
c. bushes
d. blood* (correct answer)

Table 8.11 presents the number of persons who selected each of the distractors.

Table 8.11: African Rhinos and Oxpecker Birds Item 10: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	42	13	13	7	24	8
B	54	16	15	8	47	16
C	28	8	19	10	24	8
D*	204	61	137	69	184	64
9 (not attempted)	8	2	14	7	8	3
Total	336	100	198	100	287	100

*Correct response (distractor D)

Based on Table 8.11, the Afrikaans (69%) learners found this item less difficult than the other two languages. Notably, 64% of isiZulu learners and 61% of English learners selected the correct distractor. Although more than half of the learners were able to select the correct answer, both Afrikaans and isiZulu learners outperformed the English learners. Distractor B drew the attention of some of the English and isiZulu learners while distractor C did the same for the Afrikaans learners. Moreover, the reasons learners selected distractor B may include learners not being familiar with MC type items and that options could appear to be similar to the learner. The Afrikaans (7%) group had the highest percent of learners who did not attempt to answer the item, followed by isiZulu (3%) and English (2%).

Item 12 also displayed non-equivalence between the languages. It is a literal type item that tested whether learners could retrieve explicitly stated information from the text in answer to the question. This item counted two marks; as such, learners were required to provide two answers in order to obtain full marks. Item 12 asks the following question:

What are these parts of an oxpecker like?

*Its tail is*_____.
*Its claws are*_____.

Table 8.12 (following page) presents a selection of Grade 4 learner responses to item 12.

Table 8.12: African Rhinos and Oxpecker Birds Item 12: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	sharp claws	sat.	mondiumosilowgoni
	the oxpecker sitch the rhino's back	wit.	
2	Not attempted	Die renostervoël is 'n klein bruin	mude
		Die renostervoel sit op die rug	ifishane
3	Not attempted	ten lsys.	uqinile
		mont.	zicijile
4	the oxpecker is a small brown bird with a wide	lank	oqinile
	bill.	hard	ezicijile
	stiff tail and sharp claws.		
5	Not attempted	lank	amahlalankomo ahlala
		skerp	kanye nezipho ezicijile
6	stiff	Stywe stert	uqinile
	sharp	Skerp	ezicijile
7	stiff	Skerp	uqinile
	sharp	100 van hierdie opgeswelde bosluise eet	ezicijile
8	wide bil, stiff and sharp claws	Kgebedemebege	yinyoni encane
	the oxpecker sit on the rhino's back.	tegebemedie ted	nezinzipho eziqinile
9	Mameds yame?	Die renostervoël is n klein	uqinile
	Hame sysyds	bek stywe stert en	zicijile
10	and sharp claws	stywe stert.	zinetenye zenluldukoma
	the oxpeck	kloue skerp.	lezi zingxenye

In order to obtain two marks, the learners had to indicate that the tail is stiff and claws are sharp. The memorandum makes provision for using appropriate synonyms such as rigid and pointy. Some of the learners were able to provide two correct responses, such as English learner 6, Afrikaans learner 6 and isiZulu learner 3. If the learners only provided one correct response, the learners would receive only one mark; see for example, English learner 8, Afrikaans learner 5 and isiZulu learner 5. A few learners provided sentences from the text as their answer, but they do not refer to the stiff tail and sharp claws of the oxpecker, see for instance Afrikaans learner 2.

Table 8.13 shows the breakdown of the partial credit item per language.

Table 8.13: African Rhinos and Oxpecker Birds Item 12: Partial Credit Breakdown

Points	English Persons	% Obtained	Afrikaans Persons	% Obtained	isiZulu Persons	% Obtained
0	114	34	95	48	133	46
1	42	13	30	15	20	7
2	164	49	66	33	126	44
9 (not attempted)	16	5	7	4	8	3
Total	336	100	198	100	287	100

It is evident from the table that most learners attempted to answer this item. A total of 198 Afrikaans learners completed this test and only 33% were able to provide two correct responses. Almost half of these learners provided incorrect responses. This item was less difficult for both English and isiZulu learners: 49% and 44% of English and isiZulu learners, respectively, obtained full marks. However, 46% of isiZulu learners provided incorrect responses. It would appear that most of the isiZulu learners obtained either zero marks or full marks for this item. Even though this is a literal item that required learners to find the answers in the text, a large percentage of learners found the item difficult. Based on the learner responses captured in Table 8.12, it would seem that some learners were confused by the layout of the item as they provided sentences and did not identify the sentence starters. Learners had to complete the sentences by filling in the missing word.

The penultimate item that displayed differential functioning between the languages, is item 13. To repeat, *African Rhinos and Oxpecker Bird* item 13 asks the following question:

Why do oxpeckers especially like to eat ticks?

a. because the ticks have been sucking blood* (correct answer)
b. because there are many ticks on each rhino
c. because the ticks have been eating leaves
d. because the ticks are tiny and easy to eat

Table 8.14 shows the number of persons who selected each of the distractors.

Table 8.14: African Rhinos and Oxpecker Birds Item 13: Grade 4 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A*	174	52	79	40	163	57
B	72	21	51	26	52	18
C	28	8	21	11	27	9
D	39	12	32	16	22	8
9 (not attempted)	23	7	15	8	23	8
Total	336	100	198	100	287	100

*Correct response (distractor A)

More than half the learners who completed the test in isiZulu (57%) and English (52%) selected the correct distractor, whilst only 40% of Afrikaans learners selected it. Based on this result, the item seemed somewhat more difficult for the Afrikaans learners as they might have missed the emphasis "*veral*" (especially) in the item. This item required learners to make an inference based on what they had read and it would appear that a number of learners struggled to do so. Learners found distractor B to be tempting as it refers to many ticks on a rhino. These learners may not have properly read the question as it asks why the oxpecker birds prefer ("especially like") to eat ticks – the question's emphasis itself implies that there is a reason why oxpeckers like to eat ticks. The text explains that the oxpecker bird eats insects but that it prefers eating ticks as it contains their (the oxpeckers') favourite food – blood. The learners who selected distractor B might have relied on what they had read earlier in the paragraph which explains why oxpeckers sit on the rhino's back in order to eat the ticks that live in the rhino's skin.

The last item of *African Rhinos and Oxpecker Birds* that showed differential functioning, is item 16. It asked the following question:

What does the oxpecker do to warn the rhino of danger?

This item tested learners' ability to *Focus on and Retrieve Explicitly Stated Information* and counted one mark. Learners were required to write what the oxpecker does to warn the rhino of danger. The text describes how the rhino and oxpecker depend on each other, referring to their mutually beneficial relationship. The rhino has poor eyesight and cannot spot enemies from far while the oxpeckers have sharp eyes and if they spot a potential threat, they make loud noises (including hissing). Learners had to indicate that the oxpecker makes a noise or hisses/hissing sound[76].

Table 8.15 presents a selection of ten Grade 4 learner responses to *African Rhinos and Oxpecker Birds* item 16. This is the last of the five literal items that showed non-equivalence between the languages. The isiZulu learners found this item relatively more difficult than the Afrikaans and English learners as 73% provided an incorrect answer, followed by Afrikaans (67%) and then English (47%) (*cf.* Table 8.5).

Based on the responses, some learners did not attempt to answer the question, see English learner 2 and isiZulu learner 3. The missing percentage for each language is slightly higher in comparison to some of the other items mentioned in this section: eight percent of English and isiZulu leaners did not attempt the question whereas five percent of Afrikaans did not answer the question. Learners were not required to write a full sentence but it must be clear that the oxpecker bird will make a noise when it spots danger. Afrikaans learner 10 provided a full sentence answer that states the oxpecker makes a noise if there is danger. Most of the learners who provided the correct answer only indicated that it makes a noise, see for example English learner 1, Afrikaans learner 5 and isiZulu learner 6. A few learners indicated that the two animals depend on each other but did not state the oxpecker's role, see for instance English learner 4 and Afrikaans learner 1. Some learners deviated from the question topic, indicating that the oxpecker removes the ticks from the rhino, such as isiZulu learner 7. Afrikaans learner 8 provided a nonsensical answer where letters were randomly added.

[76] *cf.* Appendix C for scoring guide.

Table 8.15: African Rhinos and Oxpecker Birds Item 16: Grade 4 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	by making loud noises and hissing	Die renostervoël beskerm ook die renostervoëls hou vyunde.	bayawudumula
2	Not attempted	Die renosters en renvoels	yebo
3	Not attempted	Hulle doen nesie	Not attempted
4	The rhinos and the oxpeckers help each other	Hulle help mekaar.	Not attempted
5	of niose	Raas	Amanlalankomo athela ukudla kwawo
6	loud noises or hissing sounds	deer om hard te raas of te blaas.	knza umsindo omkulu lindize
7	I makes loud noises and hisses	Die renosters en renstervoëls help mekaar.	basuswa imikhaza
8	because there is lots of ticks on a each rhino	syemebetgebeme	basuswa imikhaza
9	By making loud noises;	Die renosters en renost	liwasasa umukhaza emhlane
10	Depending on each other	Die skerpsiende renostervoels hou om en waarsku die renostervoëls teen gevaar deur hard to raas en blaas	amanlalala nkhome

Even though the text clearly stated that the oxpecker makes a loud noise or hissing, Afrikaans and isiZulu still found this item difficult. One of the reasons why learners struggled with this item could be test fatigue – this item was the second last item of the PIRLS Literacy 2016 Reader booklet. Another possible reason could be that these learners, especially the isiZulu learners, might not be familiar with informational type texts or familiar with the questioning and answering of such a text. It appears that learners do not read the question properly as their answers are off the topic and could be indicative of classroom practice. As the structure differs to a narrative text, they are required to make use of their structural awareness in order to answer the questions, that is, to refer to the different headings (layout) and subsequent paragraphs to find the answers.

8.3 ANALYSIS OF *THE GREEN SEA TURTLE'S JOURNEY OF A LIFETIME*

The source text (ST) as well as the Afrikaans and isiZulu versions of this story can be found in Appendix H. *The Green Sea Turtle's Journey of a Lifetime* was developed for PIRLS 2016 – it was internationally completed by Grade 4 learners; however, in South Africa, it was completed by Grade 5 learners (n = 861[77]).

8.3.1 Overview of The Green Sea Turtle's Journey of a Lifetime

The Green Sea Turtle's Journey of a Lifetime is an informational (expository) text dealing with animals. This text was used during the PIRLS 2016 cycle. South African Grade 5 learners participated in PIRLS 2016; as such, these learners completed the items for this text. This text was developed to explain the migratory patterns of a green sea turtle throughout her life, in other words, the life cycle of a green sea turtle.

In terms of the organisational elements, *The Green Sea Turtle's Journey of a Lifetime* utilises two elements that form part of the broader set of organisational sequence elements, namely chronological and spatial.

- *Chronological*: when the information in the text is arranged according to a progression of time or when the events that occurred are organised in sequence, the text is

[77] Four Grade 5 learners were removed from the analysis due to extreme scores.

considered to be chronologically organised (Dymock, 2005; Ghaith & Harkouss, 2003; Roehling et al, 2017). A writer makes use of chronological order as each section refers to a main event or next step in a process.

- *Spatial*: a writer makes use of a spatial pattern or sequence to fit information together in a physical space or location. Spatial is used when it is necessary to create a mental schema of the topic (Eitel & Scheiter, 2015). It could have links with a geographical location such as a temperate forest. Texts that utilise spatial elements such as illustrations or drawings, maps and diagrams.

With regards to *chronological*, the main idea is to provide information about the green sea turtle's life cycle in chronological order. The text begins with the first part of the life cycle, namely a baby turtle (hatchling) hatching from her egg on the beach, then crawling to the ocean where she spends most of her life, and then she returns to the beach where she was born to lay eggs of her own. Supporting ideas throughout the text include:

- The baby green sea turtle hatches from her shell (takes place on a beach in Costa Rica).
- The next day, the baby turtle crawls to the surface of the sand.
- The baby turtle then moves towards the ocean – avoiding predators.
- The baby turtle reaches the ocean and continues swimming for two days.
- Scientists do not really know what happens during the first few years after the baby turtle enters the ocean.
- The turtle, after several years, starts her trek to the Florida coast.
- As the turtle grows and eventually becomes an adult turtle, she moves further off the coast.
- Roughly at 26 years, the turtle begins her return to the beach where she was born to lay eggs.
- The turtle lays eggs in four nests over the course of two months.
- After laying her eggs, the turtle returns once more to the ocean.
- The turtle repeats this cycle for the next 80 years (IEA, n.d.-e).

The main idea of the *spatial* element of *The Green Sea Turtle's Journey of a Lifetime* is to illustrate the turtle's life in three main areas: the beach where she is born (Costa Rica), being in the open ocean, and the coast of Florida. A map is provided that clearly shows the two spatial points as part of the turtle's journey. The supporting ideas include:

- The adult turtle is able to find the beach where she is born.
- She lays her eggs (on the same beach where she is born).
- Scientists are uncertain how turtles are able to find the beach where they are born.
- Scientists hypothesise that turtles might be able to sense magnetic fields or even chemicals in the water (IEA, n.d.-e).

The abovementioned ideas provide an overview of the text by structuring them according to the two organisational elements. The text structure and organisational elements are the same across all texts. In addition, the IEA listed key vocabulary that were used in the text. The vocabulary is presented in Table 8.16.

Table 8.16: Key Vocabulary for The Green Sea Turtle's Journey of a Lifetime across Languages

IEA Vocabulary	Frequency English	Frequency Afrikaans	Frequency isiZulu
hatchling	6	12 (babaseeskilpadjie)	13 (ichwane lofudu; ichwane)
frothy surf	1	1 (skuimerige branders)	1 (amaza anamagwebu)
predators	1	1	1
juvenile	1	1	1
slurps	1	1	1
algae	2	2	2
trek	1	2 (tog, tog)	1 (uhambo olude)

The table indicates the frequency of the selected key words in ST and TT. The key words point out animals, behaviours, food and descriptions used in *The Green Sea Turtle's Journey of a Lifetime*. However, as with the previous released texts, it is not exactly clear why these key words were selected, since '*turtle*' was not included as one of the key words even though it (turtle, turtles and turtle's) appeared 26 times in the ST. In the ST, the word '*hatchling*' appeared six times whereas in Afrikaans it was used 12 times; and in isiZulu it was used 13 times. The ST used '*hatchling*' and '*baby (sea) turtle*' interchangeably, whereas in Afrikaans only '*babaseeskilpadjie*' (baby sea turtle) was used. In isiZulu, '*ichwane lofudu*' (baby turtle) and '*ichwane*' (baby) were used. The IEA also selected '*frothy surf*' – this phrase includes an adjective + noun in all three languages: '*skuimerige branders*' (frothy surf) in Afrikaans and '*amaza anamagwebu*' (bubbly waves) in isiZulu. The noun '*trek*' was used only once in

English. It is defined as "a long arduous journey" (Lexico, 2020b). In isiZulu '*uhambo olude*' was used as it refers to a long journey or trip. In both English and isiZulu, this key word was only used once and '*journey*' was used as a synonym in the next paragraph. However, in Afrikaans '*tog*' (trek) was used twice – in the text both times referred to a long journey. Note that '*tog*' is a homonym[78] and could also be used as a conjunction (but, however). The first time, the text stated '*haar lang tog terug*' (her long trek back) and included an adjective to disambiguate meaning.

The *LexTutor VocabProfile* software tool was used to analyse the vocabulary used in the corpus. Table 8.17 depicts the English word profile of *The Green Sea Turtle's Journey of a Lifetime*.

Table 8.17: Word Profile of The Green Sea Turtle's Journey of a Lifetime

Word frequency level (English)	Occurrence	% of Tokens	Cumulative Tokens (%)	Examples
K1	771	84.0	84.0	her, swim, green
K2	68	7.4	91.4	journey, sand, shell
K3	15	1.6	93.0	ocean, pit, reflecting
K4	7	0.8	93.8	juvenile, predators, snatch
K5	33	3.6	97.4	crabs, sharks, turtle
K6	7	0.8	98.2	algae, trek, walnut
K7	3	0.3	98.5	beak, herons, snails
K8	1	0.1	98.6	frothy
K11	1	0.1	98.7	slurps
K12	4	0.4	99.1	flippers
Off list	0	0	≈100	

[78] Has more than one meaning but has similar spelling and is pronounced the same (Merriam-Webster, 2020).

The word profile of *The Green Sea Turtle's Journey of a Lifetime* reached the 95% threshold[79] at K5. This finding means that majority of the words used in this expository text fall within the high (90%) to mid frequency (7%) range, while only 2% are low frequency words which corroborates research done on vocabulary and its relationship with learners' reading comprehension (*cf.* Laufer, 1989; Nguyen & Nation, 2011; Schmitt et al., 2017). Expository texts use vocabulary not often used in literary texts or in everyday discourse, but both narrative and expository texts in this study consisted of words found predominantly between the mid and high frequency levels. This finding partly explains why the 95% threshold was only reached at K5 level. As mentioned in the previous text in this chapter, readers use a different set of reading literacy skills when reading an expository text (Hebert et al., 2016). This text also made use of structure, signalled specifically by headings, as well as non-textual elements such as a map and diagram. Learners engaged with these non-textual elements to assist in the meaning-making process.

Most words used in this text fall between the K1 and K5 levels (that is, first five thousand words). Some of the words identified in this text are classified as mid or low frequency words, although this is expected for an informational text. Approximately seven percent (7.1%) of the words are considered mid frequency words. Examples include '*juvenile*', '*algae*' and '*frothy*'. Some of the low frequency words, those above the K9 range, include '*slurps*' and '*flippers*'. The verb '*slurps*' was used once in the text and describes how the juvenile green sea turtle fed on jellyfish. The noun '*flippers*' was used four times in the text and refers to the turtle's limbs used for climbing out of the nest, swimming in the ocean and digging in the sand. In both Afrikaans ('*swempote*') and isiZulu ('*amaphiko*'), the word '*flippers*' was used four times. Although '*flippers*' is a low frequency word, the text is accompanied by pictures of the sea turtle showing her flippers. Table 8.18 shows the readability analysis of *The Green Sea Turtle's Journey of a Lifetime* per language.

Table 8.18: Readability of The Green Sea Turtle's Journey of a Lifetime per Language

Readability Features	English	Afrikaans	isiZulu
No of words	907	913	633
Pages/paragraphs	7	7	7

[79] *cf.* Page 242 for an explanation of the ratio of low, mid and high frequency words required for reading with comprehension.

Readability Features	English	Afrikaans	isiZulu
Sentences per paragraph	10.1	10.0	10.1
Words per sentence	12.7	13.0	8.9
Mean word length	4.1	4.3	7.7
Passives [a]	2%		
Flesch reading ease [b]	84.1		
Grade level [c]	4.6		

[a, b, c] These features can only be calculated for English

Table 8.18 compares the text length between the ST and TT. The word count of the isiZulu version was, as expected, lower than the English and Afrikaans versions. South African Grade 5 learners completed this test even though it exceeds the 200 – 250 words limit recommended by the national curriculum, (DBE, 2011). Based on Hasbrouck and Tindal's (2017) reading norms, English Grade 5 learners, at the 50th percentile, should be able to read 127 WCPM by the middle of the year. At the end of the year, the same group should be able to read 140 WCPM. Therefore, it should take the average English Grade 5 learner approximately seven minutes to read. As this text was developed for PIRLS 2016, the layout differed from the previous expository text in this chapter – the entire text is provided which is followed by 16 question items (see Appendix H). The items that showed non-equivalence between the three languages are explored in sub-section 8.3.4.

In terms of the number of sentences per paragraph, it remained the same across the ST and TT. When looking at the number of words per sentence, it appears that the Afrikaans (13.0) version had the longest sentences, closely followed by the English (12.7) version. The isiZulu (8.9) version had somewhat shorter sentences. This finding is mirrored by the mean word length where Afrikaans (4.3) and English (4.1) were similar although isiZulu (7.7) had somewhat longer words. This finding is similar to the previous text that had similar results that confirm the conjunctive orthography of isiZulu.

8.3.2 Processes of Comprehension of The Green Sea Turtle's Journey of a Lifetime Items

The following section deals with the *Processes of Comprehension*, specifically question item difficulty for *The Green Sea Turtle's Journey of a Lifetime*. It also provides partial evidence whether non-equivalence is present in this text. The text was accompanied by 16 items worth

21 marks. Nine of the items were CR and counted 14 points whereas the MC items counted seven points. Table 8.19 presents the item difficulty of *The Green Sea Turtle's Journey of a Lifetime* per language.

Table 8.19: Item Difficulty of The Green Sea Turtle's Journey of a Lifetime across Languages

Processes of Comprehension	English Items	Afrikaans Items	isiZulu Items
Focus on and Retrieve Explicitly Stated Information	2, 5, 8, 10	2, 5, 8, 10	2, 5, 8, 10
Make Straightforward Inferences	1, 3, 4, 6, 9, 12, 13	1, 3, 4, 6, 9, 12, 13	1, 3, 4, 6, 9, 12, 13
Interpret and Integrate Ideas and Information	7, 11	7, 11	7, 11
Evaluate and Critique Content and Textual Elements	14, 15, 16	14, 15, 16	14, 15, 16

Four items are located at the *Focus on and Retrieve Explicitly Stated Information* level and forms the group of easiest items of this text. Most (7) of the items were developed to determine whether learners could *Make Straightforward Inferences*. The remaining five items comprise inferential type items, two of which tested learners' ability to *Interpret and Integrate Ideas and Information*. The remaining three items assess how well learners could *Evaluate and Critique Content and Textual Elements* from the text. Based on Table 8.19, it appears that all the question items remained at the same comprehension level across the three languages; it further indicates that none of the aforementioned TT items' difficulty changed during the translation process.

8.3.3 Professional Opinion on Translated Texts

The majority of the experts agreed that *The Green Sea Turtle's Journey of a Lifetime* is mostly written in a simple and straightforward style but does contain complex language structures; only Z3 indicated that the text is complex due to translation infelicities. For example, Z3 indicated that the heading '*Luya Olwandle Oluvulekile*' is problematic as it is a direct translation which refers to 'open seas' whereas the heading in English is '*Out to the Open Sea*'. This translation is misleading for isiZulu speakers and the translator could rather have used 'in the deep sea water'. Overall, the experts believe that the translation was adequately translated, with Z2 commenting that "the translators have done their best but were challenged but some

of the words that are not available in Zulu [sic]". Some of these words include '*wolnathi'* (walnut) and '*ebhishi*' (beach). These are also known as loan words[80]. In terms of the number of lower frequency words, the experts agree there is a tolerable incidence of low frequency words used in this text. A1 and Z2 believe that Grade 5 learners may not necessarily know some of the words used in this text. A1 remarked that words such as '*nagreiers*' (night herons), '*skuimerige*' (frothy), '*garnale*' (shrimp), '*okkerneut*' (walnut) and '*lysies*' (ledges) could be considered less common words. Z2 identified the following as possible low frequency words '*namapulastiki*' (plastic), '*ojelifishi*' (jellyfish) and '*amakhemikhali*' (chemicals). Z2 continues by saying that when loan words or low frequency words are used in an isiZulu text, "it may be of benefit to learners if the English word [was included] in brackets". The word '*turtle*' and '*shrimp*' falls within the K5 range while '*walnut*' and '*ledges*' are considered to be K6 words. Based on these observations, the three languages share some commonalities in terms of the medium frequency words, such as '*walnut*' and '*frothy*'. The noun '*turtle*' was used together with '*baby*' when referring to the hatchling[81] in the text. However, adjectives in English become compound nouns in Afrikaans, for example, '*baby sea turtle*' becomes '*babaseeskilpad*'. The words '*baby*' and 'sea' are classified as K1 words (part of the first 1 000). Interestingly the word '*turtle*' falls within the K5 range of frequency words – in Afrikaans it is '*skilpad*' and in isiZulu it is '*ufudu*'. A1 believes that '*skilpad*' matches the level of frequency as the English version '*turtle*'. A1 further commented that even though the (English) text made use of K1 – K11 frequency words, the Grade 5 learners should be able to read the text if they make use of their structural awareness knowledge. As it is an expository text, it follows certain steps, such as the use of headings and other non-textual elements (that is, illustrations and maps) to guide the reader through the text. Therefore, the reader should apply a different set of reading skills (Arfé et al., 2018; Hebert et al., 2016). *The Green Sea Turtle's Journey of a Lifetime* exposes learners to possibly new and scientific vocabulary and informs them about the lifecycle of a green sea turtle thus taking the learner beyond the known and informs them of things that are not necessarily familiar to them in their everyday lives.

In terms of cultural equivalence, there are a few aspects to take into consideration. Z1 explained that the text should be equally understood by the learners "but I think those who have no

[80] New lexical items added to a recipient language from a source language. isiZulu adds loan words from both English and Afrikaans sources (Ngcobo, 2013).
[81] Baby turtle was used as a synonym for hatchling in the ST.

experience of the seas will find it abstract". A1 has a similar thought regarding some of the animals used in the text, such as '*garnale*' (shrimp) and '*nagreiers*' (night herons) which were used only once in the text and the text does not provide an illustration or explanation of these animals. Even though the text does not expand on these issues, it should not pose a barrier to the learners but it might require the learners to reread a small portion of the text in order to make sense of what is being read. Z2 also agreed that the text should be similarly understood by the different learners. Z3 explained "the different children should be able to interpret the text with the same effort. It is an information text and the exposition is fairly logic. So, no problem". Furthermore, as this expository text aimed to inform the readers about the lifecycle and migratory patterns of a green sea turtle, the writer used relevant words and terminology, which can be seen throughout the ST and TT.

The experts also considered the construct relevance of *The Green Sea Turtle's Journey of a Lifetime*. The majority of the experts agreed that the text is construct relevant. However, Z1 explained that learners living next to the coast would probably know something about the life cycle of turtles. As a large portion of South Africans do not live near the coast, some learners might find the text abstract. Z2 commented that this text includes "factual information which some learners may find difficult to remember" – however, the purpose of an expository text is to provide readers with knowledge about a specific topic. Yet, Z3 explained that due to the translation errors, "it makes it extremely difficult to access the meaning of parts of the text". Nevertheless, the learners are able to page back to a piece of text and reread if necessary. A1 and Z2 both noted that this text links with the natural science curriculum. A1 further mentioned that the diminutive form was used in the Afrikaans version whenever the text referred to the hatchling or baby sea turtle: '*babaseeskilpadjie*' (tiny baby sea turtle). When looking at the illustrations, the majority of the experts agreed that the illustrations (pictures of the turtle, migratory map and life cycle of the turtle) are relevant to the text, which Z1 verified by stating that "pictures will clarify and give more understanding".

The majority of the experts indicated that *The Green Sea Turtle's Journey of a Lifetime* looks like a typical informational text that Grade 5 learners may come across. The experts explained that because of the structure, along with the illustrations and headings, it is similar to a text found in a textbook – with Z1 mentioning, "the way it is written" makes it look like an informational typical text. However, the text exceeds the word limit specified by the national curriculum (DBE, 2011). South African Grade 5 learners should read between 200 and 250

words for a reading comprehension exercise, nevertheless, the Grade 5 learners should not find the text too taxing considering the layout of the text. The layout and structure of *The Green Sea Turtle's Journey of a Lifetime* is the same across the ST and TT. The text comprises five pages; each page focuses on a specific part(s) of the turtle's life cycle. As this is a PIRLS 2016 text, the items were in a separate booklet and consisted of seven pages. The font used was slightly smaller than the PIRLS Literacy 2016 texts but included enough spacing so that it still reads easily. The photographs of the green sea turtle as a hatchling, juvenile and adult clearly shows what the turtle looked like during each part of the life cycle. The other non-textual elements such as the map and life cycle also add to the learners' understanding of the text. As such, the layout and illustrations used adds to the face validity of *The Green Sea Turtle's Journey of a Lifetime*.

In terms of the *Process of Comprehension*, the experts agreed that these processes did not change from the ST to the TT. The text adequately covers the four processes, as Z2 indicated that the "answers to questions can be found in the text". A1 agreed that answers can be found in the text but learners should make use of the "right tools". Z1 added that "learners have to think deeply and use the picture to support their understanding". It should be noted that Z2 highlighted item 11, which asks: What information does the article provide about the sea turtle's size and food at each stage of its life? Algae is one of the answers, however, in the isiZulu text the word '*izithalo*' (plants) was used as a synonym for algae – there is no conventional word for algae in isiZulu. Z3 argued that '*izithombo*' would have been more suitable for algae as it refers to plants in general, including sea plants. Although Z3 makes a good point, it should not have a negative bearing on the learners' understanding of the text. Most of the experts agreed that after examining the items, the items measure what they intended to measure. Moreover, the experts continued to say that the questions asked were fair. Z1 explained that she thought that "when learners are able to answer questions which show understanding they would have". Z1 thought that stories should be localised as an abstract storyline could be difficult for the learners to understand. However, this argument is often applied to narrative texts, where it is easier to localise stories, but informational texts are not readily amenable to localisation. It should be noted that PIRLS texts are selected by over 50 participating countries and each text was vigorously examined to ensure that there is no

discrimination, bias or prejudice of any kind and that the texts meet criteria as set out by the IEA which includes exposing learners to new information[82].

8.3.4 Per-item Analysis for The Green Sea Turtle's Journey of a Lifetime

The following sub-section presents the per-item analysis of *The Green Sea Turtle's Journey of a Lifetime*, which includes the percentage of incorrect responses, missing percent per item as well as the item which displayed differential functioning among the languages. In addition, a sample of the English, Afrikaans and isiZulu learner responses is included for each constructed response (CR) item. Table 8.20[83] (following page) shows the number of learners who had incorrectly responded to the question items.

Only three of the 16 items displayed non-equivalence between the ST and TT. This text had the lowest incidence of DIF items among the two informational texts. These include items 1, 10 and 11. Item 1 is a MC item while both items 10 and 11 are CR type items that required learners to provide a written answer.

Item 10 is a literal item which required learners to find the answer in the text. The learners who completed this test in isiZulu obtained the highest incorrect percentage (79%), followed by Afrikaans (68%) and English (50%) learners. In terms of missing data, only 5% of Afrikaans learners did not attempt the item, whereas 10% and 11% of isiZulu and English learners, respectively, did not attempt to answer the item.

Item 1 tested learners' ability to make inferences from the text. Learners who completed the test in English found this item relatively easy with about a third (31%) providing incorrect responses, followed by 34% of Afrikaans learners providing incorrect responses. The isiZulu learners found this item somewhat more difficult with more than half (56%) of the learners providing an incorrect answer. Missing values ranged between 1% and 2% across languages.

[82] *cf.* Chapter Two; and Mullis & Martin (2015).
[83] Table is sorted according to the four *Processes of Comprehension*.

Table 8.20: The Green Sea Turtle's Journey of a Lifetime: Percentage of Grade 5 Learners who Incorrectly Answered Items per Language

Item No	Processes of Comprehension	English			Afrikaans			isiZulu		
		N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing	N Completed	% Incorrect	% Missing
2	Focus on and Retrieve Explicitly Stated Information	343	74	2	172	74	2	339	94	3
5	Focus on and Retrieve Explicitly Stated Information	343	49	3	172	56	2	339	62	3
8	Focus on and Retrieve Explicitly Stated Information	343	64	10	172	69	6	339	81	13
10*	***Focus on and Retrieve Explicitly Stated Information***	***343***	***50***	***11***	***172***	***68***	***5***	***339***	***79***	***10***
1*	***Make Straightforward Inferences***	***343***	***31***	***2***	***172***	***34***	***1***	***339***	***56***	***2***
3	Make Straightforward Inferences	343	84	2	172	87	2	339	94	3
4	Make Straightforward Inferences	343	79	4	172	85	2	339	92	5
6	Make Straightforward Inferences	343	58	5	172	71	4	339	74	6
9	Make Straightforward Inferences	343	62	9	172	66	5	339	74	8
12	Make Straightforward Inferences	343	36	14	172	40	6	339	45	12
13	Make Straightforward Inferences	343	58	14	172	70	7	339	71	13
7	Interpret and Integrate Ideas and Information	343	87	8	172	91	3	339	94	6
11*	***Interpret and Integrate Ideas and Information***	***343***	***78***	***15***	***172***	***83***	***11***	***339***	***84***	***16***
14	Evaluate and Critique Content and Textual Elements	343	61	18	172	63	10	339	68	19
15	Evaluate and Critique Content and Textual Elements	343	42	23	172	53	17	339	55	22
16	Evaluate and Critique Content and Textual Elements	343	33	24	172	48	16	339	47	24

*Indicates items that displayed DIF.

The last item that displayed DIF is item 11, which tested the higher order skill to interpret and integrate information across the text. This item was challenging for learners across all three languages as most of the isiZulu (84%), Afrikaans (83%) and English (78%) learners gave incorrect responses. The missing values for this item are quite high: 16% of isiZulu learners did not attempt to answer this item, closely followed by English (15%) learners with 11% of Afrikaans learners who did not attempt the question.

The Rasch analysis (*cf.* Chapter Six) of *The Green Sea Turtle's Journey of a Lifetime* indicated items that displayed non-equivalence across the languages. As two of the DIF items are CR type items, a selection[84] of learner responses[85] in each language was captured to further examine possible reasons for the non-equivalence. In terms of the MC item, the entire sample's responses were included for discussion.

To repeat, *The Green Sea Turtle's Journey of a Lifetime* item 1 asks the following question:

What is the first section "Out From the Sand" about?

a. what different sea turtles look like
b. how sea turtles learn to swim
c. what sea turtles like to eat
d. how sea turtles' eggs hatch * (correct answer)

This item tested the comprehension process *Make Straightforward Inferences*. Table 8.21 depicts the number of persons who selected each of the distractors.

Table 8.21: The Green Sea Turtle's Journey of a Lifetime Item 1: Grade 5 Learner Responses

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
A	50	15	19	11	70	21
B	41	12	30	17	85	25
C	17	5	9	5	34	10

[84] The Centre for Evaluation and Assessment's (CEA) data manager randomly selected the booklets.
[85] Constructed Response (CR) items only.

Distractor	English Persons	% Selected	Afrikaans Persons	% Selected	isiZulu Persons	% Selected
D*	228	66	112	65	142	42
9 (not attempted)	7	2	2	1	8	2
Total	343	100	172	100	339	100

*Correct response (distractor D)

Two-thirds of English (66%) and Afrikaans (65%) learners selected the correct distractor while less than half of isiZulu (42%) learners selected distractor D. The item seemed somewhat less difficult for the English and Afrikaans learners. Even though the isiZulu learners struggled with this item, the item asks specifically about the first section of the story with the heading '*Out From the Sand*'. In this section, the writer introduces the reader to a sea turtle's nest in the sand and then moves to how the sea turtles start to hatch from their eggs. From the information contained in this section, learners were required to make the inference that the section was about how the sea turtles' eggs hatch. Both distractors A and B were tempting for the learners, with only a few selecting distractor C. Learners who selected distractor A might have focused on the picture of a baby sea turtle which shows what the turtle looks like. The section did not mention how sea turtles learn to swim nor what they like to eat. As such, learners who selected distractors B and C might have relied on what they have read about the turtles further along in the text.

The next item that showed non-equivalence between the languages is item 10:

Why does a sea turtle's body fat become green?

Table 8.22 shows a selection of ten Grade 5 learner responses to *The Green Sea Turtle's Journey of a Lifetime* item 10. It is a literal question item and required learners to *Focus on and Retrieve Explicitly Stated Information* from the text. This item takes the form of CR which counts one mark. The Grade 5 learners had to identify the (green) food the green sea turtle eats. Examples include (sea) grass, algae, seaweed[86] or food that is green.

[86] *cf.* Appendix C for scoring guide.

Table 8.22: The Green Sea Turtle's Journey of a Lifetime Item 10: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	Because it became to grow	As sy baie groen gras eet	Not attempted
2	Because the turtle eat green grass	Omdat die skilpad eet die gras onder die water.	Not attempted
3	It's body becomes green because of eating the sea grass	Wan hulle groei	Not attempted
4	The baby turtle's journey through the open sea is often called the "lost years."	Wand hulle is in die sand gebore	yingoba ufudu amafutha umziba ayashibilika
5	The nest holds more than 100 green sea turtle eggs, each about the size of a golf ball	van dit is 'n seeskilpad.	konke kofudu
6	Eating sea grass and algae turns her body fat a green colour	Wanneer hulle volwasse word.	Ufudu dohlaza lubikakhulu laba angani bomi aboningi abafundanga
7	It is becoming an adult	hulle raak groter en groter.	angaba ngama-26 130 manje luqala
8	because it eats alot of green grass	As sy uit die dop kom en dan raak tan grour	Ingoba luhlala emanzini
9	It eats the grass	Dat hy lyk soos die gras	yingoba ufudu lufana
10	Be cause eating sea grass and algae turns her body fat a green colour	Die seegras en alge wat sy eet veroorsaak dat haar vel groen word.	Yingoba angene emanzini asolwandle libaleka

As this is a literal item, the Grade 5 learners had to find the answer in the text; however, the learner should only provide answers that relate to green food. The different types of green food (algae, sea grass and seaweed) are mentioned under the section '*Growing Up Green*'. The text explains that the juvenile sea turtle mostly eats algae and sea grass. A picture of the turtle eating sea grass accompanies the text. Table 8.22 shows that only a few learners did not attempt to answer the item, see isiZulu learner 1 and 2. Some learners referred to the sea turtle growing up or becoming an adult such as English learner 1 and Afrikaans learner 3. Others referred to the size or age of the turtle, see English learner 5 and isiZulu learner 7. A few learners provided incomplete sentences, for instance isiZulu learner 5. It is evident that both Afrikaans and isiZulu learners found this item more challenging compared to the English learners (see also Table 8.20). The majority of Afrikaans and isiZulu learners' responses did not include green food but rather about other aspects of the turtle. These learners might not have understood that the question was about the (green) food that the turtle ate, which turned the turtle's body fat green. As these learners struggled to find explicitly stated information in the text, it might possibly relate to difficulty with reading and writing skills.

Item 11 of *The Green Sea Turtle's Journey of a Lifetime* is the last item of this text that displayed differential functioning between the languages. It asks the following:

What information does the article provide about the sea turtle's size and food at each stage of its life?

Complete the table below.
Three have been done for you.

Stage of life	***Size***	***Food***
egg	*(i)*	*The egg has its own food*
hatchling	*(ii)*	*(iii)*
juvenile	*dinner plate*	*(iv)*
mature	*(v)*	*algae and sea grass*

This item counted three marks and in order to obtain full marks, the Grade 5 learners were required to complete the answer table by interpreting and integrating information from across the text. The answer table focuses on each stage of the sea turtle's life, its size and the type of

food it consumes during each stage. The table provides learners with three answers meaning that the learners should complete the remaining five spaces (see *i* to *v*). Table 8.23 provides the correct answers for item 11.

Table 8.23: The Green Sea Turtle's Journey of a Lifetime Item 11: Correct Responses

Stage of life	Size	Food
egg	golf ball	The egg has its own food
hatchling	walnut	shrimp; jellyfish; snails; sea snails
juvenile	dinner plate	algae; grass; sea grass; turtle grass; seaweed; jellyfish
mature	1 metre; 130 kilograms	algae and sea grass

To obtain full marks, learners had to provide a correct answer for each of the five missing answers. The marking guide also stipulates that learners should not be credited if they have included any incorrect pieces alongside the correct responses. If learners correctly completed four out of five answers, they were awarded two marks and learners who correctly completed three out of five spaces obtained one mark. Those learners who provided two or fewer correct responses, received an incorrect (zero mark). Table 8.24 presents ten Grade 5 learners' responses to *The Green Sea Turtle's Journey of a Lifetime* item 11. The learners' responses were captured in order of the blank spaces in the answer table (see also Table 8.23).

Table 8.24: The Green Sea Turtle's Journey of a Lifetime Item 11: Grade 5 Learner Constructed Responses

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
1	of a golf ball; 100; sand; grass; 150	Not attempted	Not attempted
2	tennis ball; ball; green grass; jellyfish; 130 kg's	Not attempted	Not attempted
3	golf ball; walnut; shrimp, small jellyfish and snails; jellyfish; rocks (sea)	skryf bord; waterbord; die babas maak ver hulle kos; soek ver hulle kos; swembord	Not attempted
4	egg food; dinner plate; few week later; the juvenile has good food; eat sea grass	volwassene; jongskilpad; die baba moet vir hom kos gaan soek; hulle soek sla kos; babaskilpadjie	libukuda nstukuzonke; lisakane icwane; icwane linempilo; ukudla okuningi; luyahlupha ngendlela eyenza ngayo emazini
5	Green; egg; the eggs has hatchling; the has its dinner; grass	babaskilpadjie; lê eiers; swem lang afstande om kos te kry; dieselfde plek toe om te eet; jongskilpaaie	lokuphila; uhamba; okofudu; esihlabathini; lokhelo
6	golf ball; no bigger than a walnut; shrimp, small jellyfish, and snails; she some times slurps jellyfish; fat and green colour	seegras; seeskilpad; jellievisies; gras; seegras	liqandalelo ekilekuphla; abaningiabaha mbiesheshi; mina Ngahamba esontweni; ipulethilibi kakhulu; oludala uthango solugugile
7	golf ball; golf ball; grass; grass; 130 kg.	eet; skilpad; hulle eie babaskilpadjies; jonger word; broei	not attempted; qanda lazalela ama; zasalwandle; ne amachwa; angama

Learner Number	English Answer	Afrikaans Answer	isiZulu Answer
8	golf ball; walnut; not attempted; jellyfish, algae, sea grass; 1 meter	groot eier; dop; ma skilpad; en kos; groot mense	ichwane; asekhulile; oludala; liphumela n qaphandle; zidla
9	golf ball; golf ball; sand; jelly fish; dinner plate	grasse; basse	likhule; ayaqandal iqandaliyokhu; liyafa; layakhula kakhulu
10	golf ball; bigger than a walnut; small jellyfish and snails; algae and sea grass; 1 metre long	gholfbal; okkerneut; lappie seegras; lappie see gras; 1 meter lank	amachwane; lile egezikhulakhu; amaqada luzalela; oludala; aluncane

Based on the learner responses provided in Table 8.24, some learners provided incomplete answers, see Afrikaans learner 1 and isiZulu learner 1. Some learners, such as Afrikaans learner 9, only provided a partial response. In this case, the learner provided two possible answers, leaving the remaining three spaces empty. Learners also gave answers not relevant to the text, see for example English learner 2 who referred to a tennis ball. A few learners also repeated words from the table such as isiZulu learner 8. Others were able to list at least two correct answers, see English learner 7; however, even though those two answers were correct the learner scored a zero[87]. Four learners were able to obtain a partially correct score, see English learner 2 and Afrikaans learner 10. Both these learners scored two marks as they provided four correct answers. English learner 10 was able to provide five correct answers and obtained full marks for this item.

This item required learners to make use of their visual literacy skills as it includes a graphical device as part of the item. Graphical devices or non-textual elements such as graphs, pictures and tables are often used in expository texts (Roberts, Norman & Cocco, 2015). A reader is required to comprehend these non-textual elements to assist in making sense of what he or she is reading (Van den Broek & Kendeou, 2017). In terms of this item, learners were required to complete a table. However, based on the above learner responses as well as Table 8.20, it seems this item was somewhat difficult for the learners as very few were able to obtain three marks. Table 8.25 shows the breakdown of the partial credit item for each language.

Table 8.25: The Green Sea Turtle's Journey of a Lifetime Item 11: Partial Credit Breakdown

Points	English Persons	% Obtained	Afrikaans Persons	% Obtained	isiZulu Persons	% Obtained
0	224	65	127	74	282	83
1	14	4	6	3	2	1
2	29	8	10	6	0	0
3	26	8	10	6	0	0
9 (not attempted)	50	15	19	11	55	16
Total	343	100	172	100	339	100

[87] *cf.* Appendix C for scoring guide.

The majority of isiZulu (83%), Afrikaans (74%) and English (65%) learners were not able to obtain a mark, indicating the difficulty of this item. None of the learners who completed the test in isiZulu obtained full marks, while only 8% and 6% of English and Afrikaans learners, respectively, were able to provide five correct responses and received full marks. The missing percentage was relatively high for this item, with 16% of isiZulu learners who did not attempt an answer to this item, closely followed by the English learners (15%). Eleven percent of Afrikaans learners did not attempt to answer this item.

Evidence from both *African Rhinos and the Oxpecker Birds* and *The Journey of the Green Sea Turtle* texts (*cf.* Tables 8.8, 8.20, 8.24 and 8.25) showed that learners have difficulty in reading and completing a table in relation to information provided in the text. The difficulty of this particular item might be threefold: a) it tests a higher order comprehension skill, namely *Interpret and Integrate Ideas and Information*; b) it is an expository text which makes use of a different set of literacy skills (Arfé et al., 2018; Hebert et al., 2016); and c) requires understanding of how to transfer information from the text into table form. Even though the South African home language curriculum for Grade 5 learners indicates that learners should be taught how to read visuals as part of reading comprehension (*cf.* DBE, 2011), it would appear that these learners still struggle with reading and completing tables. As non-textual elements such as tables and diagrams, amongst others, are ubiquitous with expository texts, it is important for learners to use, decipher and complete these non-textual elements (Roberts, Norman, Duke, Morsink, Martin & Knight, 2013).

8.4 CHAPTER SUMMARY

This chapter provided the findings of the qualitative analysis of the PIRLS Literacy 2016 and PIRLS 2016 released texts. These include *African Rhinos and the Oxpecker Birds* and *The Green Sea Turtle's Journey of a Lifetime*. Evidence suggests that both the South African Grades 4 and 5 learners did not perform well on the abovementioned texts as both texts displayed differential functioning across English, Afrikaans and isiZulu (*cf.* Chapters Five and Six). Thus, this study further explored these two texts and items for possible reasons for the non-equivalence in an attempt to answer two of the five sub-questions.

As this study wanted to determine whether the PIRLS assessments had face validity, sub-question 3 asked:

To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?

Both informational (expository) texts were examined in order to answer this research question. Based on the evidence presented in this chapter, the *African Rhinos and Oxpecker Birds* and *The Green Sea Turtle's Journey of a Lifetime* is considered to have face validity. The evidence was based on expert knowledge and their review of each text. Each text included headings which added to the overall structure of the texts and assisted in guiding the reader through the information. Both texts, as indicated by the experts, are similar to informational texts to which Grade 4 or 5 learners may be exposed at school. The texts were also accompanied by illustrations, specifically photographs that clearly give an indication of what each animal looks like, however the Grade 4 illustrations were only in black and white. During the examination of these texts, the experts found that each text made use of words not used every day; however, as with the narrative texts in Chapter Seven, there was a low percentage of low frequency words. Both informational texts focused on animals which were considered familiar to South African learners – one expert did mention that the informational texts should be localised; however as a main goal of informational texts is to expose learners to new topics such as the animals presented in each text, it is not appropriate to localise these texts. In terms of *African Rhinos and the Oxpecker Birds*, a portion of the text was followed by a few items and then continued with the information on the next page whereas the items of *The Green Sea Turtle's Journey of a Lifetime* were presented after the entire text. Note that across the ST and TTs, the above layout was the same. Thus, the layout of these texts provided further evidence for face validity as each text's content was located in the same place across the three versions, including the position of each illustration.

Next, sub-question 5 posed the following:

How does textual coherence as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?

This sub-question required the evaluation of the two informational texts by focusing on *macro, meso* and *micro* levels of equivalence. To recapitulate, *macro* level looked at the purpose of the text, length and layout of each text. The *meso* level included the main ideas or concepts of each text, question types and each item's process of comprehension. The *micro* level included the key words as identified by the IEA, vocabulary used in each text, non-textual features such

as the illustrations (such as photographs and maps) and overall readability of the texts such as the average word length.

Interestingly, the shortest of the Grade 4 texts, *African Rhinos and the Oxpecker Birds* had 9 items that displayed DIF (*cf.* Chapter Six) while the Grade 5 text, *The Green Sea Turtle's Journey of a Lifetime* only highlighted 3 items which behaved differently across the ST and TTs (*cf.* Chapter Five). The non-equivalence of each text can be attributed to several reasons; however, there is no compelling evidence from these analyses that the difficulty of the texts or different cultural experiences form part of these reasons. Overall, the panellists thought the translations were adequately conducted, however one panellist did specifically mention that the isiZulu translation of *The Green Sea Turtle's Journey of a Lifetime* was problematic. Interestingly, this text had very few items that functioned differently among the English, Afrikaans and isiZulu learners. Based on my evaluations, together with the panellists, it is argued that other factors may have posed a barrier to performance. These factors include the learners' lack of knowledge of structural awareness of informational or expository texts, exposure to this type of text especially in isiZulu, learners' lack of engagement with these texts, understanding the purpose of each text as well as reading and completing tables. Even though informational texts form part of the national curriculum, these factors could be due to what is happening in the classroom.

As with sub-question 4 (*cf.* Chapter Seven), it is not always known how or to what extent teachers prepare and organise their lessons, in other words, the classroom environment. Spaull and Pretorius (2019) in their review of the Early Grade Reading Assessment (EGRA) found that there are several reasons why South African learners still lag behind other countries in terms of reading. One in particular is the quality of pre-service teacher training and their knowledge and skill to teach reading to learners. As such, the quality of pedagogy in South Africa is questionable, as based on the PIRLS results, learners cannot read for meaning. It should be noted that this study did not specifically look at the classroom environment, but it was nevertheless important to mention these as teachers' skill and experience in teaching language and literacy is an important factor in learner reading comprehension (*cf.* Zimmerman, 2017).

During the exploration of the PIRLS Literacy 2016 and PIRLS 2016 informational texts and items, it was found that learners may have experienced test fatigue as the texts were longer than what they were used to (*cf.* Sievertsen et al., 2016). The PIRLS assessments did include a

break after 40 minutes in order for learners to refresh before continuing with the second part of the assessment. The learners may also have limited experience with completing multiple choice type items and providing reasons for their answers. Another point to consider is the number of literal items which displayed differential functioning across the ST and TTs. The learners may be under a misconception that they need to provide their own answers for a question rather than search in the text – this could be indicative of how the teachers are teaching reading comprehension as well as learners' inability to read the question before attempting to answer it.

When considering the translation of each text into Afrikaans and isiZulu, it was found that despite some differences of opinion between the panellists, particularly in isiZulu, the texts were felt to be adequately translated. However, as mentioned above, one panellist found the isiZulu translation of *The Green Sea Turtle's Journey of a Lifetime* to be problematic and possibly hamper the learners' ability to answer the questions associated with this text. As there are few informational texts available in African languages, such as isiZulu, it may have been difficult for the translator and back translator to ensure a faithful translation of the ST at the appropriate grade level. As mentioned in Chapters Two and Seven, the IEA endeavours to provide quality assessments to inform policy and practice. Even though the IEA has a set of guidelines to be followed when translating the test instruments, this study found that both texts in Afrikaans and isiZulu included translation errors, some of which were extensive and may be a possible reason for the learners' poor performance.

The final chapter of this thesis discusses the data presented and research findings of this chapter as well as Chapters Five, Six and Seven. It incorporates and interprets these findings to answer this study's main research question. In addition, Chapter Nine also provides reflections, implications and limitations of this study.

-⌘-

CHAPTER 9

Conclusions and Recommendations

9.1 ORIENTATION

This study examined the equivalence of the released PIRLS Literacy 2016 and PIRLS 2016 assessment instruments across three languages. As such, a sequential explanatory mixed methods design was utilised to investigate the South African English, Afrikaans and isiZulu learners' performance on the PIRLS Literacy 2016 and PIRLS 2016 assessments. This mixed methods design examined the quantitative data (collection and analysis) before qualitative data collection and analysis (Creswell, 2013; Creswell & Plano Clark, 2011). As this study examined the PIRLS assessments, the quantitative phase of this study was a secondary analysis. During the quantitative phase, descriptive analysis such as means and percentages were calculated. Thereafter, Rasch analysis was conducted to determine whether any of the items functioned differently between the selected languages. The outcomes of the first phase informed the second phase of the study where I undertook content analysis of the literary and informational texts, and conducted workshops with experts in order to review texts and items as well as to complete questionnaires about the similarity, layout and other textual aspects of the texts and items across languages.

A synopsis of the preceding chapters is presented initially (9.2), followed by a summary of the research undertaken during this study to answer the research questions posed by the study (9.3). Then the research findings of the first quantitative phase are discussed in order to address sub-questions 1 and 2 (9.4). Thereafter, the research findings of the second qualitative phase are discussed in order to address sub-questions 3, 4 and 5 (9.5). A reflection is provided on the conceptual framework (9.6). Then, the main conclusions (9.7) as well as the research methodology (9.8) of this study are provided. As this study took the form of a mixed methods research design, the strengths and limitations of the study are discussed (9.9). Recommendations for policy, practice as well as research recommendations are provided in Section 9.10. This chapter also indicates the possible contributions of the study (9.11) and concludes with final thoughts (9.12).

9.2 SYNOPSIS OF THE PRECEDING CHAPTERS

Chapter One provided information on the South African Grade 4 and 5 learners' reading

literacy comprehension during the PIRLS Literacy 2016 and PIRLS 2016 studies – both Grade 4 and 5 learners obtained poor results. This chapter contained the problem statement and rationale for conducting this study which served as the reason and motivation for exploring possible reasons why South African Grades 4 and 5 learners performed poorly on the PIRLS assessments. This study investigated whether the PIRLS assessments are equivalent across three languages, namely English, Afrikaans and isiZulu. This chapter also presented the study's main research question, as well as the subsidiary research questions, which were answered by employing a sequential explanatory mixed methods design. That being said, I conceptualised this study based on the work of Peña (2007) who focused on translation equivalence by examining linguistic, functional, cultural and metric equivalences. Thereafter, a brief overview of the thesis was provided to guide the reader.

Chapter Two focused solely on PIRLS; to begin with, I provided background information on the IEA, the research centre responsible for conducting several ILSAs worldwide. The PIRLS reading framework was then discussed by considering what reading literacy is as well as the two purposes for reading literacy and four processes of comprehension. Thereafter I discussed the context questionnaire framework that included information about the learner, parent or primary caretaker, teacher and school principal. The PIRLS assessment instruments were discussed next which included the achievement booklets and contextual questionnaires. As PIRLS is an ILSA, it was important to report on the translation procedures adhered to during the test development process. These procedures included the translation of the instruments, selecting the target language, translators and reviewers, as well as international translation verification. Finally, the data collection, monitoring, scoring and analysis were considered as well as the ethical considerations for conducting the assessment in South Africa.

A review of the literature was presented in *Chapter Three*. The chapter began with an overview of the South African educational landscape post 1994 and in particular, examined the national curriculum followed in South Africa. Subsequently, I investigated the LiEP and provided some shortcomings thereof. As this study focused on the equivalence of text across different languages, I next considered issues related to translations and adaptations. In addition to the aforementioned, a discussion of face, content and construct validity was undertaken as it forms part of equivalence in this study. Thereafter, I considered what constitutes text and textual coherence. Next, I deliberated on Peña's (2007) work about translation equivalence by

specifically exploring linguistic, functional, cultural and metric equivalences. This chapter concluded with a presentation of the conceptual framework that guided the study.

Chapter Four presented the study's research design and methodology used during Phases One and Two of the study. This chapter began with my ontological and epistemological viewpoints, which was followed by providing complementarities of both quantitative and qualitative methods. Thereafter, I positioned this study within the pragmatic research paradigm. Next, a description of mixed methods design was given with justification for selecting a sequential explanatory mixed methods design for the study. Phase One of the study was a secondary analysis of the PIRLS Literacy 2016 and PIRLS 2016 South African data and as such, made use of the sample of Grade 4 and 5 learners who completed the released texts in English, Afrikaans and isiZulu. The first phase, the quantitative phase, sought evidence of DIF amongst the three languages for each of the PIRLS released texts. The information and findings gathered during Phase One informed the second, qualitative phase. Phase Two involved an in-depth look at the released texts and learner responses of problematic items identified during Phase One. The data analysis for this phase was also discussed. Thereafter the methodological considerations of the study were presented.

Chapters Five and *Six* presented the quantitative findings of the study for Phase One. *Chapter Five* focused on the PIRLS Literacy 2016 released texts whereas *Chapter Six* focused on the PIRLS 2016 released texts, with one text being shared between PIRLS Literacy 2016 and PIRLS 2016 as a link between the two studies. The overall learner achievement on both PIRLS assessments was examined. Subsequently, I conducted Rasch analysis to determine whether measurement invariance, also known as item bias, was present for each item for each of the texts. Both chapters concluded with a summary of the items which displayed differential functioning between the languages.

Chapters Seven and *Eight* presented the qualitative findings of the study for Phase Two. In order to enhance the readability and flow of these chapters, I presented the findings of the literary (narrative) texts in *Chapter Seven* while *Chapter Eight* dealt with the informational (expository) texts. A discussion of the different features of the multipronged analysis of the PIRLS texts was presented in *Chapter Seven* which included the macro, meso and micro levels of equivalence. As part of the content analysis, I provided an overview of each text, followed by an examination of the items by comparing each item's difficulty. Thereafter, the professional opinions on the Afrikaans and isiZulu PIRLS released texts and items are given.

As both *Chapters Seven* and *Eight* form part of the second, qualitative phase, I further examined the items which displayed differential functioning by examining a selection of learner responses for each DIF constructed response item to identify possible patterns and reasons as to why each of these items behaved differently amongst the English, Afrikaans and isiZulu learners. By examining the learner responses, it was possible to see how learners responded to each of these items.

9.3 SUMMARY OF THE RESEARCH

This study was prompted by South African Grade 4 and 5 learners' poor reading literacy comprehension during PIRLS Literacy 2016 (Howie et al., 2017a) and PIRLS 2016 (Howie et al., 2017b). In South Africa, there is a plethora of research focusing on language, literacy and translations (*cf*. Ngcobo, 2013; Pretorius & Stoffelsma, 2017; Sibanda & Baxen, 2018; Spaull et al., 2020; Taylor & von Fintel, 2016); however, little information is available about test equivalence across different languages. The aim of this study was to explore the equivalence across English, Afrikaans and isiZulu Grade 4 and 5 texts during PIRLS.

The South African Grade 4 and 5 learners performed poorly during the 2016 round of PIRLS. The Grade 4 learners obtained the lowest average (320, SE=4.4) amongst the PIRLS Literacy 2016 participating countries. Other PIRLS Literacy participants included Egypt (330, SE=5.6), Morocco (358, SE=3.9), Kuwait (393, SE=4.1), Iran (428, SE=4.0) and Denmark[88] (501, SE=2.7). South Africa's cohort of Grade 5[89] learners were included as a benchmark participant. Of the benchmark participants, South African Grade 5 learners obtained 406 score points (SE=6.0) which was also the lowest average amongst the benchmarking participants. Some of the other benchmarking participants included Buenos Aries, Argentina (480, SE=3.1), Norway (517, SE=2.0) and Moscow City, Russian Federation (612, SE=2.2).

Based on the above PIRLS results, it is evident that the South African learners performed poorly on reading comprehension with less than a quarter of Grade 4 (22%) learners and about half of the Grade 5 (51%) learners only reaching the low international benchmark (Howie et al., 2017a; 2017b). A myriad of factors could be associated with the learners' poor performance. Some of the main concerns, as identified by Mullis, Martin, Foy and Hooper

[88] Tested Grade 4 learners during PIRLS Literacy 2016 in all official languages.
[89] Only Grade 5 English, Afrikaans and isiZulu participated in PIRLS 2016.

(2017) include the home environment, school composition and resources, school climate, overall discipline and safety at schools, teacher preparation and instruction as well as the learners' own attitudes and beliefs towards reading.

Due to recent PIRLS results, the South African government as well as the Department of Basic Education have stressed the importance of reading literacy, with the President of South Africa stating in his state of the nation address in 2019 that every 10-year old should be able to read with meaning (South African Government, 2019). As such, the development of reading literacy has become a national aim for South African education.

Over the last two decades, the education system of South Africa has changed considerably. Where education was first demarcated along racial lines, post-1994, the democratic country scrapped education policies and curricula used during apartheid and developed a new education curriculum. The National Education Policy Act (NEPA), Act 27 of 1996 was one of the first new policies implemented to address the educational inequalities of the past. NEPA brought several different policies into effect – for instance, the LiEP in 1997. Curriculum 2005 (C2005) was also developed in 1997 which comprised Outcomes Based Education (OBE) (Chisholm, 2005). A few years later, the National Curriculum Statement (NCS) was developed and completed in 2002 (DoE, 2002; Chisholm, 2005). However, as South Africa was still refining its education curricula, a revised version of the NCS was introduced in 2006. In 2009, a ministerial committee report was submitted to the Minster of Education with concerns about and suggested changes to the-then current curriculum (DoE, 2009; du Plessis, 2013). Consequently, the Curriculum Assessment Policy Statement (CAPS) was introduced as an amendment to the NCS (du Plessis, 2013; Pinnock, 2011). The learners who participated in both PIRLS 2016 assessments had been following the newly implemented CAPS.

As this study examined the equivalence of the PIRLS assessments across different languages, it was important to gain an understanding of the translation and adaptions as both PIRLS assessments were adapted for the South African context. PIRLS Literacy 2016 translated into the ten different national languages while PIRLS 2016 only translated into Afrikaans and isiZulu. Translation is the act of moving linguistic discourse from a SL to a TL (Chan & So, 2017). It could be considered as an encounter between a SL and TL (*cf.* Vottonen, 2016). In his work, Nida (1964) argued that the message or content contained in the TL should be as close and equivalent as possible to the SL. As such, different levels of equivalence were investigated in this study.

In her work, Peña (2007) identified four main considerations to safeguard against threats to validity of large, cross-cultural assessments. These include linguistic, functional, cultural and metric equivalences. This study focused on these equivalences to determine whether the PIRLS assessments are equivalent across English, Afrikaans and isiZulu. Linguistic equivalence entails whether the ST and TT are similar (Chesterman, 2016) while functional equivalence concerns whether the TT behaves similarly to the ST (Greenfield et al., 2006). When conducting ILSAs, it is important to consider cultural equivalence – ensuring that the same meaning is understood in more than one culture (Chan & So, 2017; du Plessis et al., 2015). Metric equivalence comprises the item difficulty of assessments (Peña, 2007), particularly in ILSAs. In order to gain a comprehensive understanding of the South African PIRLS 2016 results, this study looked in to macro, meso and micro levels of equivalence in order to assist in determining whether the PIRLS 2016 released texts were equivalent in terms of Peña's four considerations (equivalences) to protect against possible validity threats against cross-cultural research. The *macro* level comprised an overview of each text including text characteristics such as genre, theme, purpose as well as the layout of each text. This level linked with functional equivalence. The *meso* level included passage mapping for each text, item difficulty as well as looking at the *Processes of Comprehension*. Therefore, the *meso* level connected with functional, metric and linguistic equivalences. The *micro* level involved looking at aspects such as vocabulary, word and sentence length for each text. This level related with linguistic and cultural equivalence.

Against this backdrop, the study posed the following main research question:

To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?

Five sub-questions were proposed in order to assist in addressing the main research question. These sub-questions include:

Phase One Sub-Questions

1. *How does learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?*
2. *How does learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016?*

Both sub-questions 1 and 2 deal with South African learners' reading literacy achievement during the PIRLS assessments. Sub-question 1 focuses on PIRLS Literacy 2016 (Grade 4) whereas sub-question 2 focuses on PIRLS 2016 (Grade 5). The main purpose of both questions was to determine whether the English, Afrikaans and isiZulu learners performed differently on the PIRLS assessments and whether these differences could then be reasonably attributed to possible translation equivalence issues. The findings of both sub-questions 1 and 2 raised additional questions about the validity and equivalence of PIRLS within the South African context.

Phase Two Sub-Questions

3. *To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?*

The reason for including this question for this study is twofold. Firstly, as the South African national curriculum prescribes a word count for reading comprehension texts, it was important to compare these with the word count of the PIRLS texts. Secondly, it was necessary to establish whether the released texts included in the PIRLS assessments, were similar to the texts prescribed for use in schools in South Africa.

4. *How does textual coherence, as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?*
5. *How does textual coherence, as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?*

The last two sub-questions are multifaceted and in order to ascertain whether the textual coherence of the texts differs amongst the source text (English) and translated texts (Afrikaans and isiZulu), three levels of equivalence were examined, namely *macro*, *meso* and *micro*. At the *macro* level, an overview of each text was presented; the *meso* level provided passage mapping for each text as well as an investigation of item level difficulty; and the *micro* level looked at the word and sentence level features of each text. Each level of equivalence was examined in order to partially assist in addressing the sub-question.

9.4 PHASE ONE: SUMMARY AND DISCUSSION OF THE RESEARCH FINDINGS

This section provides a summary of the research findings of the first, quantitative phase of the

study. Sub-questions 1 and 2 investigated South African learners' reading comprehension scores during the PIRLS assessments. Sub-question 1 examined the Grade 4 learners' reading comprehension by using the PIRLS Literacy 2016 data whereas sub-question 2 studied the Grade 5 learners' reading comprehension by examining the PIRLS 2016 data. The purpose of these two sub-questions was to determine if the learners who completed the assessments in English, Afrikaans and isiZulu performed differently on each of these assessments. Sub-section 9.4.1 deals with the findings of sub-question 1 and sub-section 9.4.2 discusses the findings of sub-question 2.

9.4.1 Research Sub-Question 1

The first sub-question of the study asked the following:

1. *How does learner performance differ across English, Afrikaans and isiZulu languages on PIRLS Literacy 2016?*

This sub-question referred specifically to three of the released texts used during PIRLS Literacy 2016, namely *The Pearl*, *African Rhinos and the Oxpecker Birds* and *Flowers on the Roof*. In order to address this question, descriptive analysis as well as Rasch analysis were used to analyse the South African Grade 4 learner results.

Descriptive analyses were conducted to partly assist in answering sub-question 1 by providing raw mean scores per text and percentage correct per item across English, Afrikaans and isiZulu. The mean scores were based on the total mark allocation for each text, for example, *The Pearl* counted 18 marks. Even though the *Flowers on the Roof* text was considered to contain easy language structures and familiar words, the Grade 4 learners struggled the most with this text. The *Flowers on the Roof* text was considerably more difficult than *The Pearl* or *African Rhinos and the Oxpecker Birds*. Further investigation of the latter two texts painted a curious picture where the Grade 4 learners performed slightly better on the informational text, *African Rhinos and the Oxpecker Birds* than the narrative text, *The Pearl*. In terms of language groups, the English learners obtained higher mean scores than the Afrikaans and isiZulu learners for *African Rhinos and the Oxpecker Birds* (*cf*. Table 5.3). For the narrative text, the English learners also obtained higher mean scores, with Afrikaans learners' scores being slightly lower (*cf*. Table 5.3).

After the initial exploration of the PIRLS Literacy 2016 data, Rasch analysis was conducted to determine whether the items from the three released texts displayed possible item bias or non-equivalence. Two Rasch analyses were conducted, the first for *The Pearl* and *African Rhinos and Oxpecker Birds* texts as these were completed by the same learners; a second was conducted on the text *Flowers on the Roof* which was completed by a different group of learners.

As part of the Rasch analysis, individual item-fit statistics were conducted to partially test the null hypothesis of this study, namely that the mean score of the English, Afrikaans and isiZulu learners who participated in PIRLS Literacy 2016, are the same (μEnglish = μAfrikaans = μisiZulu). The item-fit statistics indicated the chi-square of each item, which showed the invariance across the trait (Pallant & Tennant, 2007). In this case, the trait would be learner achievement. The item-fit statistics indicated the misfit of possible problematic items; these items were classified either as *overfit* or as *underfit* (*cf.* Tables 5.9, 5.10 and 5.11). After the item-fit statistics, DIF was conducted to highlight differential functioning amongst the three languages.

The DIF also assisted in testing the null hypothesis by specifically looking at the ANOVA results. The ANOVA tested for significant differences between the mean scores of English, Afrikaans and isiZulu (p-value <0.05). The null hypothesis was rejected (*cf.* Table 5.12) as the mean scores of the English, Afrikaans and isiZulu learners were not equal for *The Pearl* and *African Rhinos and the Oxpecker Birds*. In particular, the following items displayed differential functioning across the sub-groups:

- *The Pearl* items 2, 5, and 14
- *African Rhinos and the Oxpecker Birds* items 2, 3, 4, 5, 6, 10, 12, 13 and 16

The results of the ANOVA presented evidence that the above items functioned differently for the English, Afrikaans and isiZulu language sub-groups. Based on the DIF analysis, the problematic items were further explored by conducting item characteristic curves (ICC). In terms of the ICC of each item that displayed DIF, no clear pattern was observed as there was no universal discrimination against one language (*cf.* Figures 5.2 to 5.20).

For the *Flowers on the Roof* text, a similar examination of the English, Afrikaans and isiZulu results were conducted. The item-fit statistics can be found in Tables 5.15, 5.16 and 5.17. Based on the statistics, DIF was conducted on the *Flowers of the Roof* items (*cf.* Table 5.18). The

following items indicated DIF: 1, 5, 6, 9 and 12. The ANOVA analysis showed evidence that these items functioned differently for English, Afrikaans and isiZulu learners and as such the null hypothesis (μEnglish = μAfrikaans = μisiZulu) was rejected. Item curves were conducted for each of the *Flowers on the Roof* items that displayed DIF (*cf.* Figures 5.21 to 5.27), however here too, no clear pattern was observed between the three languages. It should be noted that overall, this text proved to be the most challenging, with few learners correctly responding to items.

9.4.2 Research Sub-Question 2

The second sub-question of the study posed the following:

2. *How does learner performance differ across English, Afrikaans and isiZulu languages on PIRLS 2016?*

This sub-question asked a similar question to sub-question 1, however, the focus is on the Grade 5 English, Afrikaans and isiZulu learner reading literacy achievement during PIRLS 2016. Therefore, a similar structure was used during the summary and discussion of research sub-question 2.

Sub-question 2 involved three released texts used during PIRLS 2016; these texts included *Macy and the Red Hen*, *The Green Sea Turtle's Journey of a Lifetime* and *Flowers on the Roof*. Both descriptive analysis and Rasch analysis were used to answer this research question.

During the descriptive analyses, the raw mean scores per text as well as the percentage correct for each item across the languages were calculated. The mean scores were calculated based on the number of correct responses for each text in relation to the total mark allocation for each text, for instance *Macy and the Red Hen* counted 19 marks. Overall, the South African Grade 5 learners found *The Green Sea Turtle's Journey of a Lifetime* somewhat more difficult compared to the narrative texts. Based on the raw mean scores, the learners who completed PIRLS 2016 in isiZulu obtained the lowest scores across the three texts (*cf.* Table 6.3).

Thereafter, Rasch analysis was conducted to ascertain whether any of the items of the three PIRLS 2016 released texts displayed non-equivalence. Two Rasch analyses were conducted, one for the *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime* (same learners completed both these texts) and the second for *Flowers on the Roof* (completed by different learners).

The Rasch analysis which included individual item-fit statistics conducted to partially test the null hypothesis, revealed no significant differences in performance between the Grade 5 English, Afrikaans and isiZulu learners during PIRLS 2016 (μEnglish = μAfrikaans = μisiZulu). Tables 6.9, 6.10 and 6.11 provide the item-fit statistics which indicated misfit. Based on the item-fit statistics, DIF was conducted to further examine the differential functioning of the items among the languages.

DIF also assisted in testing the null hypothesis as it provides ANOVA statistics. The ANOVA statistics tested for significant differences between the Grade 5 English, Afrikaans and isiZulu mean scores (p-value <0.05). The null hypothesis was rejected (*cf.* Table 6.12), as the mean scores of the different languages were not similar. The following items displayed DIF across English, Afrikaans and isiZulu:

- *Macy and the Red Hen* items 1, 4, 6, 10, 12, 13 and 15
- *The Green Sea Turtle's Journey of a Lifetime* items 1, 10 and 11

Based on the ANOVA statistics presented, it is evident that the above items functioned differently across the three languages. The DIF analysis indicated that these items were problematic requiring further investigation by considering each problematic item's ICC. As with sub-question 1, no clear pattern was observed as none of languages were universally discriminated against (*cf.* Figures 6.2 to 6.15).

Next, the *Flowers on the Roof* text was investigated with the same examination being conducted as with the previous texts. Tables 6.15, 6.16 and 6.17 present the item-fit statistics of *Flowers on the Roof* per language. From the results of the item-fit statistics, DIF was conducted on the *Flowers on the Roof* items (*cf.* Table 6.18). Seven items displayed DIF, namely 5, 7, 8, 9, 10, 12 and 13. The ANOVA indicated that each of these items functioned differently for the English, Afrikaans and isiZulu learners; thus, the null hypothesis (μEnglish = μAfrikaans = μisiZulu) was rejected. In order to gain a better understanding of these problematic items, item curves were conducted (*cf.* Figures 6.16 to 6.24). In terms of the DIF and ICC, no clear pattern was observed as there was no universal discrimination against one language.

It is evident from the findings of Phase One of this study that some items did function differently amongst English, Afrikaans and isiZulu speaking learners. However, the findings of sub-question 1 and 2 provided evidence of no particular response patterns that presented

bias for any of the narrative texts, against one language. This finding suggests that the PIRLS 2016 achievement results present no discernible pattern of discrimination against any of the languages. This study found variation across language texts in that some items were easier or more difficult for one language than for the other. For instance, the learners who completed the test in isiZulu had a much higher chance of correctly answering item 13 of *African Rhinos and the Oxpecker Birds* (*cf.* Figure 5.8) while Afrikaans learners had a higher probability of correctly answering item 5 of *The Pearl* (*cf.* Figure 5.17). As such, these Phase One findings add to the credibility of the PIRLS 2016 assessments in South Africa, in at least three of the official languages. Based on these results, an in-depth qualitative analysis was conducted to further explore this phenomenon.

9.5 PHASE TWO: SUMMARY AND DISCUSSION OF THE RESEARCH FINDINGS

The second, qualitative phase of this study was informed by the findings of the first quantitative phase of the study. The quantitative phase, as highlighted in Section 9.4, indicated that several items of each of the released texts functioned differently amongst the English, Afrikaans and isiZulu learners. Consequently, this study further investigated the possible reasons for the measurement invariance (as indicated by the DIF analysis) during Phase Two of the study. The third sub-question involved the face validity of the PIRLS assessments (Section 9.5.1). The two last sub-questions each dealt with the textual coherence of the literary (Section 9.5.2) and informational (Section 9.5.3) texts by focusing on macro, meso and micro levels of equivalence.

9.5.1 Research Sub-Question 3

The third sub-question asked the following:

3. *To what extent do the PIRLS Literacy 2016 and PIRLS 2016 assessment instruments have face validity?*

In order to address this question, the focus shifted to an examination of the five released texts. The majority of the five panellists, experts in either Afrikaans or isiZulu (*cf.* Chapters Seven and Eight) agreed that the texts were similar to typical texts suggested for use by South African Grade 4 or 5 learners. They also agreed that the items were appropriate for Grade 4 and 5 learners. However, there were some qualifications to this general consensus, reflecting

differing expectations about what Grade 4 and 5 learners were capable of. One of the panellists contended that *The Pearl* was perhaps too long for Grade 4 learners and it should rather have been used in Grade 5. A panellist mentioned that *African Rhinos and the Oxpecker Birds* might not be culturally relevant as some South African learners may not know about the wildlife used in the text. One panellist thought that even though *The Green Sea Turtle's Journey of a Lifetime* made use of headings and illustrations, the text itself contains too much information for South African Grade 5 learners.

To further investigate sub-question 3, each text was examined in terms of layout. The layout for each text across English, Afrikaans and isiZulu is exactly the same, with each section of text being located at exactly the same place on each page across each language. The illustrations were on the same pages across the languages and in the same position. Thus, the information pertaining to each question is located within the same paragraphs in each language. However, there is a slight difference in terms of layout between the PIRLS Literacy 2016 and PIRLS 2016 released texts. As part of the international format and guidelines set out by the IEA, the PIRLS Literacy 2016 texts contained one page of text which was directly followed by a few questions based on the information on the opposite page while the PIRLS 2016 texts included the entire text which was then followed by the set of questions. Based on the evidence provided, the five released texts and items are all considered to have similar face validity, thereby strengthening the functional equivalence of assessment across the three languages.

9.5.2 Research Sub-Question 4

The fourth sub-question asked the following:

4. *How does textual coherence, as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS literary texts?*

The PIRLS Literacy 2016 and PIRLS 2016 literary (narrative) texts include *The Pearl*, *Flowers on the Roof* and *Macy and the Red Hen*. This sub-question involved evaluating the literary texts and items by *macro*, *meso* and *micro* levels of equivalence. To recap, the *macro* level comprised an overview of the text by providing text characteristics such as genre, theme, main characters, layout as well as the length of the text. The *meso* level related to the narrative events of each text, and the different question items which tested the processes of comprehension. The *micro* level focused on the vocabulary used, key expressions, literary or rhetorical devices used in the text as well as frequency of words, average word length and average sentence length (*cf.*

Section 7.2 for a discussion of the multipronged analysis of the texts). This sub-section provides the findings of sub-question 4 by discussing the *macro*, *meso* and *micro* levels of equivalence of the narrative texts across the three languages. In addition to the above-mentioned levels of equivalence, this sub-section also provides the findings of the translations of the texts.

- *Macro Level Findings of the Narrative Texts*

The *macro* level of equivalence focused on the overview of each text which included the genre of each text, both explicit and abstract themes, plot as well as the main characters for each text. All three texts were realistic fiction texts with two main characters and entailed a typical problem and resolution plot.

The Pearl is a rags-to-riches text of a young boy who became a wealthy pearl merchant, only to realise that wealth alone does not bring happiness. *Flowers on the Roof* is a story of a young boy who became friends with a quirky and resourceful granny who moved to the city and who brought her animals with her in order to make her new house feel like home. *Macy and the Red Hen* tells a story about a girl who struggles to deal with a bossy hen and solves her problem by outsmarting the hen and becoming top of the pecking order. Each of these texts in terms of genre, main characters, setting, plot and themes remained the same across English, Afrikaans and isiZulu.

As part of *macro* level equivalence, the general layout of each text was evaluated. Of the three narrative texts, *The Pearl* was developed for PIRLS Literacy, meaning that the layout was different to the other two texts (*Flowers on the Roof*; *Macy and the Red Hen*) developed for PIRLS.

The Pearl contained six pages of text and six pages with questions. However, the text was divided into sections where each page focused on a specific part of the text which was then followed by a few questions before continuing with the next part of the text. The layout was the same across each language, such that the same information can be found at the same place across the English, Afrikaans and isiZulu versions. *The Pearl* also included illustrations that accompanied each section of the text.

Flowers on the Roof consists of four pages of text followed by five pages of questions, meaning that the learners had to read the entire text before answering the questions. This layout was

identical across English, Afrikaans and isiZulu. Thus, the learners could find the information pertaining to a question at the same location across the different versions. The font size was smaller than *The Pearl* but it was still easily readable. Illustrations were also included – some portrayed the quirkiness of the granny, while others captured the essence of the piece of text it accompanied.

Macy and the Red Hen involves six pages of text and six pages of questions. As with the above text, the learners had to read the entire text before starting with the question items. However, as *Macy and the Red Hen* was part of the Reader booklet, the text and questions were provided in two booklets – the reason for this split is so that the learners could keep the text as a classroom resource. The font size was similar to that of *Flowers on the Roof* and was easy to read. This text also included illustrations that pertained to a specific part of the text. The information and illustrations were located at the same location across the English, Afrikaans and isiZulu versions.

In terms of text length of each text, in line with the conjunctive nature of isiZulu orthography, the isiZulu texts were much shorter as they contained longer polymorphemic words, hence shorter sentences than the English and Afrikaans texts, while retaining the same content. The English and Afrikaans texts were relatively similar in terms of text length and average word length. This finding is not unexpected due to the orthography of isiZulu, resulting in longer word units.

Based on the above discussions, it is evident that based on the *macro* levels of equivalence, *The Pearl*, *Flowers on the Roof* and *Macy and the Red Hen* were the same across the three languages and added to the face validity of the PIRLS assessments. Furthermore, these findings support the importance of achieving functional equivalence across different languages.

- *Meso Level Findings of the Narrative Texts*

The *meso* level of equivalence entailed the narrative events of each text. The IEA generated passage maps for each text that included major or main events that occurred throughout each text. As part of the content analysis of each text, this study included the narrative events of *The Pearl*, *Flowers on the Roof* and *Macy and the Red Hen*. The same narrative events were evident across English, Afrikaans and isiZulu versions as each text followed the events in the same manner.

The PIRLS *Processes of Comprehension* were also considered during the *meso* level exploration of the narrative texts. Each item was developed to test a specific skill according to the four *Processes of Comprehension,* as identified in the PIRLS 2016 Framework (Mullis & Martin, 2015). An important issue was whether the questions stayed at the same level of difficulty across the three languages as translations could affect the extent to which information in a SL text remains explicit or implicit in the TL text. It was found that the level of difficulty of the processes of comprehension remained the same across the different languages (*cf.* Tables 7.7, 7.18 and 7.39). This finding strengthens the linguistic and metric equivalence of the texts across the different languages.

In additional to the above findings, this study also checked the number of items across each text. *The Pearl* contained 15 items. Of the 15 items, 7 were MC type questions while 8 were CR items which required the learners to provide a written response. Three of these items displayed DIF amongst the English, Afrikaans and isiZulu languages. Two of the DIF items were MC. *Flowers on the Roof* comprised 13 items, of which seven were MC questions and six were CR items. As this text was included in both PIRLS Literacy 2016 and PIRLS 2016, Grade 4 and 5 learners, respectively, completed the items. In terms of Grade 4, five items displayed DIF and three were CR items. The Grade 5 contained seven items that displayed DIF, of which five were CR items. *Macy and the Red Hen* included 16 items, half (8) of the items were MC questions while the remaining half (8) were CR items. Seven of these items functioned differently across the three languages of which four were CR type items. The DIF items were dealt with in sub-questions 1 and 2 (*cf.* Chapters Five and Six).

From the above findings, it is apparent that the narrative events of each text, across the three languages, were the same. This finding is also true for the different *Processes of Comprehension*. Although the translation of some items could have been refined or translated differently, they should not pose a barrier to the learners – see the discussion below regarding the translation of the texts and items. Lastly, the number of items as well as the type of items remained the same across each language. As such, none of the languages enjoyed favouritism above the other.

- *Micro Level Findings of the Narrative Texts*

The *micro* level looked at the vocabulary used in each text, average word length as well as the words per sentence. In terms of narrative texts, it also included rhetorical devices used during

the text. The learners' responses were also investigated as part of the micro level of equivalence.

As *The Pearl*, *Flowers on the Roof* and *Macy and the Red Hen* are narrative texts, each text made use of a rich vocabulary. Note that the IEA, as part of the passage mapping, included key words or vocabulary. However, as mentioned in *Chapters Seven* and *Eight*, the criteria for selecting the key words or vocabulary was not clear and did not always include words from the title, for example '*pearl*' was not considered a key word for *The Pearl* text. Even though the IEA provided key words, this study further looked at the frequency level (low, mid or high) of the vocabulary used for each text in English[90]. If 95% of the words used in the text fall within the high (K1 to K3) frequency levels, the text is considered to be easy (Laufer, 2013; Laufer & Ravenhorst–Kalovski, 2010) and the words should be known by Grade 4 learners.

Since most (94%) of the words from *The Pearl* fall within the K1 to K3 frequency range, the text is considered to be balanced and South African learners should be able to read the text with relative ease. Interestingly, the word '*pearl*' is considered a mid frequency word. Both Afrikaans and isiZulu panellists agreed that it is not a commonly used word but as it was described several times with illustrations, learners should have sufficient exposure to understand the word. The *Flowers on the Roof* text contained mostly high frequency words, with 95% of the vocabulary ranging between K1 and K2 frequency levels. The text did include one low frequency word namely '*cluck*', which was used only once in the text. This text had the highest percentage of high frequency words of the three narrative texts. *Macy and the Red Hen* included a richer vocabulary compared to the other two texts as it only reaches the 95% threshold at K5 level. This finding indicates that the text included more words from the mid frequency range. For instance, '*squawked*' is a K9 frequency word. This text could be considered appropriate for adept readers and less appropriate for beginner readers based on the higher percentage of mid frequency words. Based on the outcome, two of the three texts in English confirm previous research done on vocabulary and its association with reading comprehension (*cf.* Laufer, 1989; Nguyen & Nation, 2011; Schmitt et al., 2017). Overall, these findings support and strengthen the linguistic equivalence of the PIRLS narrative texts.

[90] Currently, there is no software available to conduct the same analyses in Afrikaans or isiZulu.

In terms of the average word and sentence length per language, as expected the average isiZulu word was longer than both English and Afrikaans words – the latter two languages' average word length was similar across each text. The average sentence length also differed across languages with isiZulu having shorter sentences but longer word units. Again, the English and Afrikaans average sentence length was similar across each text. The differences in average word and sentence length between isiZulu and the other two languages are not surprising as the orthography of isiZulu is different. Furthermore, there is a remarkably consistent pattern across all three narrative texts.

This phase of the study also looked at the learners' answers to constructed response (CR) type items, which were flagged during the Rasch analysis during Phase One. These items comprised both lower and higher order comprehension skills. What is unexpected was the high number of literal questions that displayed differential functioning across the languages (*cf*. Chapters Five and Six). These items are expected to be easy as they are cognitively less demanding than higher order comprehension items and answers to these questions are explicitly stated in texts. Learners only had to copy the word(s) from the text and provide these as their answers or they had to make a straightforward inference from the text. The panellists, overall, did not find the translated versions of these items to be problematic. While some panellists did mention that some items that could have been phrased differently or included in a different format, the existing translated forms comprised explicit information from which answers to the less demanding questions could be derived.

For instance, *Flowers on the Roof* item 8 was an item that indicated DIF, with an isiZulu panellist indicating that the first part of question could have been confusing for the learners. In the ST, the question asked: '*Find the part of the story next to this picture of Granny Gunn*'. The isiZulu version asked '*Thola ingxenye yalolu daba ngalo mfanekiso kaGogo u-Gunn*' (Get the one part of the serious matter using this illustration of grandmother Gunn). However, a better translation would have been '*Thola ingxenye yale ndaba ehanbelana nalo mfanekiso kaGogo u-Gunn*' (Get the part of the story which goes together with the illustration of grandmother Gunn).

This is only an example from one panellist, as the majority of panellists did not find serious translation errors across the texts or items. Thus, the translations might have some bearing on some aspects of the South African learners' poor performance on the PIRLS assessments but not consistently so, there are other reasons for poor reading comprehension. For example,

learners struggled to provide written responses to CR type items – not all CR items tested higher order comprehension. *Macy and the Red Hen* item 6, which is a literal CR item that displayed DIF asked learners to provide two things Macy did that did not work (to get the red hen into the cage). The text explicitly indicates four different actions that Macy took but which did not work. Learners seemed to have struggled with CR type items in general and not necessarily with higher order skills CR items.

- *The Quality of the Afrikaans and isiZulu Translations of Each Text*

As part of the *macro*, *meso* and *micro* levels of equivalence, this question also touched on the translations of each text. As such, each text was scrutinised in terms of the quality of the translation from the ST to TTs.

Overall, the translations for each text were considered to be adequate although the texts included some inappropriateness of style such as ungrammatical language, inappropriate wording or direct translations. The majority of the texts were simply written and mostly easy to understand.

The panellists felt that the texts should not pose as a barrier for any of the language sub-group as each text's theme was something to which South African learners could relate. For instance, an explicit theme of *The Pearl* was for one character to follow his dream (of becoming a pearl merchant). Even though each text included some inaccuracies, one reviewer felt that the *Flowers on the Roof* text was poorly translated into isiZulu as the translator sometimes used free translations where either a different word was used to English, or an isiXhosa word was used instead of an existing isiZulu one (even though the two languages share a large number of cognate words). In terms of the Afrikaans version of *Flowers on the Roof*, it was felt that the translator used 'old fashioned writing' that would not always be easily understandable to the average Afrikaans speaking learner.

Furthermore, each text was evaluated for construct relevance. Although each text contained some infelicities, where the panellists thought that the texts could have been translated differently, each text was considered relevant and the reviewers agreed that learners should be able to cope with the texts. The texts include illustrations which also assist comprehension as they portray major events across the story. Some of the concepts may be abstract, but a main goal of narrative texts is to expose the reader to a richer vocabulary and stimulating features that transport the reader from their known world to other ways being or doing things.

9.5.3 Research Sub-Question 5

The last sub-question is similar to sub-question 4; however, sub-question 5 focuses on the informational texts. Sub-question 5 asks the following:

5. *How does textual coherence, as measured by macro, meso and micro levels of equivalence, differ amongst the source and translated PIRLS informational texts?*

This sub-question deals specifically with the PIRLS Literacy 2016 and PIRLS 2016 informational (expository) texts, namely *African Rhinos and the Oxpecker Birds* and *The Green Sea Turtle's Journey of a Lifetime*. As with sub-question 4, the aforementioned texts were analysed according to *macro*, *meso* and *micro* levels of equivalence (*cf*. Section 7.2) and followed the same structure in reporting.

- *Macro Level Findings of the Informational Texts*

In terms of the macro level of equivalence of the two informational texts, this study considered the purpose of each text as well as the organisational elements used during each text. Both texts were expository in nature and were about animals.

The PIRLS Literacy 2016 text, namely *African Rhinos and the Oxpecker Birds* focused on the mutually beneficial relationship between the African rhinos and the oxpecker birds. The text made use of description as well as compare and contrast to convey the purpose of the text and to introduce the reader to the rhinos and oxpeckers, what they look like, how much each weigh and more importantly the special relationship between these two animals. The PIRLS 2016 text, namely *The Green Sea Turtle's Journey of a Lifetime* entailed the migratory patterns of a green sea turtle throughout her life, that is, the life cycle of the green sea turtle. This text made use of chronological (sequence) and spatial organisational elements to orientate the reader and to inform the reader of each part of the turtle's life cycle, including her size, what she ate, possible dangers she might face as well as returning to her birth place. As both of these texts are informational, use is made of headings and illustrations (photographs and maps) to guide the reader through the text. The information within each text across English, Afrikaans and isiZulu did not change, meaning that the purpose of each text remained the same during translation.

In addition, both texts were scrutinised in terms of the layout. Because one text was developed for PIRLS Literacy and the other for PIRLS, as such these texts have different layouts.

The *African Rhinos and the Oxpecker Birds* text comprised six pages of text and seven pages with question items. The text contained several headings that only dealt with specific information. A photograph accompanied each part of the text. Learners had to read a portion of the text and answer the questions on the opposite page before continuing with the next part of the text. The layout of this text is the same as the narrative text, *The Pearl*. The layout of the text across the English, Afrikaans and isiZulu versions was the same, which means that learners are able to find the same information pertaining to a question at the same place across the languages.

The Green Sea Turtle's Journey of a Lifetime included five pages of text and seven pages of questions. The layout of this text was similar to that of the narrative text *Macy and the Red Hen*, where the questions only followed after the text and that the text and questions were provided in two separate booklets. This text made use of specific headings which encapsulate the main idea of the paragraph(s) and included several illustrations, including photographs, a map and a diagram. As with the previous informational text, the information under each heading was the same across the different languages. The learners could find the answers pertaining to questions at the same location across the three languages. Moreover, the layout was specified by the IEA and as such indicates equivalence across the ST and TT in terms of layout and structure.

Similar to the narrative texts, both the informational texts were longer in the English and Afrikaans versions as the isiZulu versions contained fewer words. The isiZulu sentences were on average shorter but contained on average longer words. The text length in English and Afrikaans was similar. As mentioned in sub-question 4, this finding is not surprising due to the longer word units in isiZulu orthography.

The *African Rhinos and the Oxpecker Birds* and *The Green Sea Turtle's Journey of a Lifetime*, based on the above discussions, provide evidence that both informational texts were the same across English, Afrikaans and isiZulu. Furthermore, the macro level of equivalence adds to the face validity of the PIRLS informational texts.

- *Meso Level Findings of the Informational Texts*

The *meso* level findings of the informational texts are slightly different compared to the findings of the narrative texts. As the narrative texts included main or major events that occurred during the text, the informational texts made use of organisational elements that

described each part of the text with each heading and subsequent paragraph(s) focusing on describing or comparing each of the animals. The information separated by the different headings across the two texts were the same for the English, Afrikaans and isiZulu versions of the texts.

In order to assist in determining the *meso* level of equivalence of the informational texts, the PIRLS *Processes of Comprehension* were investigated. Each question item was developed to test one of four skills as indicated by the PIRLS 2016 Framework (*cf*. Mullis & Martin, 2015). This process is the same for both the literary (narrative) and informational texts. Similar to the narrative texts, it was important to determine whether the questions remained at the same difficulty level across the ST and TT as the translation of the ST could affect whether the explicit and implicit content remains the same across the TTs (*cf*. Tables 8.4 and 8.19). Based on the translations of the items, it was found that the difficulty of the items did not change, thus each question across the languages tested the same *Processes of Comprehension*.

By looking further into each texts' items, it was found that *African Rhinos and the Oxpecker Birds* consisted of 17 items, 10 of the items were CR items while the remaining seven items were MC. Nine items displayed DIF across the English, Afrikaans and isiZulu languages. Four of these items were CR and all four tested the literal skill to focus and retrieve explicitly stated information from the text. *The Green Sea Turtle's Journey of a Lifetime* included 16 items, of these nine were CR items and seven were MC. Three items displayed DIF between the three languages, of which two were CR items, with one CR item testing lower order comprehension skills and the other testing higher order skills. The items that functioned differently across the three languages were discussed in *Chapters Five* and *Six*. In terms of the first text, the finding was unexpected as learners struggled with a basic, literal type item that required less cognitive processing.

Based on the above discussions, it is evident that the ST and TTs followed the same structure; each text made use of specific organisational elements and these were the same in the translated versions. Each text included a specific set of items that covered the Processes of Comprehension. None of the processes changed during the translation, meaning that the difficulty, in terms of the *Processes of Comprehension*, remained the same across the three languages. All of the items were included in the translated versions.

- *Micro Level Findings of the Informational Texts*

This sub-section provides the findings of the *micro* level equivalence of the two informational texts. Similar to the narrative texts, the vocabulary used in each text was explored as well as the average word and sentence length. In addition to these, the learner responses to the DIF items were also investigated.

Both *African Rhinos and the Oxpecker Birds* and *The Green Sea Turtle's Journey of a Lifetime* were informational (expository) texts and included a wide range of vocabulary to convey the purpose of each text. As part of the passage mapping developed by the IEA, some key words were included. However, as with the narrative texts, the criteria for selecting these words were not apparent as neither included the animals in each of the texts, such as '*rhino*', '*oxpecker bird*' and '*turtle*', as mentioned in *Chapters Seven* and *Eight*. In order to gain a holistic view of the vocabulary used in each informational text, the frequency level of each text was calculated. The number of low, mid and high frequency words contained in each text is an indicator of the difficulty level of the text (Laufer, 2013; Laufer & Ravenhorst–Kalovski, 2010). Ideally, as mentioned in Section 9.5.2, if 95% of the vocabulary used in the text ranges between K1 to K3 levels, the text is considered easy and appropriate for beginner readers.

The *African Rhinos and Oxpecker Birds* text only reached the 95% threshold at the K8 level whereas the narrative (literary) texts reached this at about K3. As such, the text included a higher percentage of mid frequency range vocabulary. Thus, this finding indicates that a richer, technical vocabulary was used in the expository text and which required the reader to engage in more abstract language to make connections throughout the text (Pentimonti et al., 2010). Words such as '*rhinoceros*' and '*oxpecker bird*' in Afrikaans '*renoster*' and '*renostervoël*' and in isiZulu '*obhejane*' and '*ihlalankomo*' are not high frequency words – this matches the findings of the English text that indicated that these words are at mid-frequency level (*cf.* Table 8.3). Even though all South African learners do not know these words, the photographs and descriptions of each animal should provide enough information for the learners to expand on their knowledge.

The Green Sea Turtle's Journey of a Lifetime text reached the 95% threshold at the K5 level, whereas *African Rhinos and the Oxpecker Birds* only reached it at the K8 level. This finding indicates that this text includes fewer low frequency words compared to the PIRLS Literacy 2016 text. About 90% of the words used in *The Green Sea Turtle's Journey of a Lifetime* are

high frequency words. Similar to the previous text, this text introduced the readers to the animal by frequently in using the animal noun '*turtle*' across the text and by providing photographs that clearly depict the turtle.

In addition to the above, this study also looked at the average word and sentence length of each text across the three languages. Similar to the narrative texts, the average isiZulu word was longer in comparison to English and Afrikaans words, while the English and Afrikaans average word length was similar. In terms of the average sentence length, the English and Afrikaans sentences were similar and longer than the average isiZulu sentence. Again, this finding is not surprising due to isiZulu orthography.

As part of the micro level of equivalence, this study examined the learner response to CR items which were identified during the Rasch analysis in Phase One. In terms of the lower and higher order comprehension skills tested during each of the informational texts, all four CR type items of *African Rhinos and the Oxpecker Birds* were basic, literal questions requiring learners to access and retrieve explicitly stated information in the text (*cf.* Chapter Five). This finding is similar to the narrative texts that also indicated a high number of DIF items that tested lower order comprehension skills. In terms of *The Green Sea Turtle's Journey of a Lifetime*, only two CR items displayed DIF, one of which tested lower order comprehension while the other tested higher order comprehension skills. The majority of the panellists found that the items were adequately translated, although one panellist in particular explained that there were infelicities such as clumsy phrasing, incorrect word order and direct translation (in both text and items) that might have hampered the learners' understanding.

For example, item 11 of *The Green Sea Turtle's Journey of a Lifetime* asks the learners to complete a table – but that: '*Three have been done for you*'. In the isiZulu version the aforementioned sentence is confusing as it says '*Sewenzeliwe vele okwesithathu*' (already done three times). The sentence could rather have read '*Sewufakelwe izimpendulo ezintathu*' (there have been three answers filled in).

The above is only one example offered by the panellists; however, this item was equally difficult for English, Afrikaans and isiZulu learners and it may have been due to the format of the question, that of completing a table. Reading and completing tables is part of the Grade 5 South African curriculum (*cf.* DBE, 2011) and learners should be used to reading and comprehending tables. Even so, most of the learners struggled to complete the table and

provided irrelevant answers to each space of the table (see example of learner responses in Chapter Eight). Across both texts, learners struggled to provide written responses but these are not necessarily related to more difficult items as the majority of the DIF CR items were basic, literal items. As such, it may be that learners do not properly read the question, are simply not used to providing written responses, and are merely used to copying chunks of text by way of answering questions.

- *The Quality of the Afrikaans and isiZulu Translations of Each Text*

During the analysis of the two informational texts, the panellists also evaluated the translations of these texts. For both texts, the panellists disagreed on the complexity of the text. In terms of the first text (African rhinos) half the panellists indicated that the text was simply written with others arguing that it was mostly simply written with some evidence of complex language structures in place. One panellist explained that the register of the ST and TT is similar but the structure of isiZulu makes it more complex than English. In terms of the second text (green sea turtle), the majority of panellists indicated that the text was mostly simply written, while one explained that the text was complex and difficult to understand due to translation inaccuracies and unnecessary complex sentences in the isiZulu version. The Afrikaans translations of both texts were considered adequate as the purpose of each text was clear, the writing was not complex, and although minor infelicities were found, and sometimes direct translation occurred, they were unlikely to pose barriers for the learners.

Even though both texts included some inaccuracies similar to the narrative texts, the majority of the panellists agreed that the English, Afrikaans and isiZulu learners should have an equal understanding of the two texts, especially taking into consideration the structure of the text as well as the informative photographs and other visuals (for example, maps and diagrams) included in each text. Each text had a specific purpose in mind and informed learners of the animals and their behaviours. Reading texts enables the reader to access different aspects of the world, such as green sea turtles without seeing them in real life. As the informational texts made use of scaffolding by means of headings and illustrations, the learners should be able to cope with the texts.

9.6 CONCEPTUAL FRAMEWORK REFLECTIONS

The study's conceptual framework (*cf*. Figure 3.1) was informed by Peña's (2007) four considerations when conducting cross-cultural research. The framework guided the study's

data collection and analysis. In order to focus on the equivalence of the PIRLS assessments, a pragmatic approach was employed to enable the researcher to make use of both mono-methods to gain a better understanding of the research problem (*cf*. Doyle et al., 2016; Johnson & Onwuegbuzie, 2004; Morgan, 2014). The framework was structured in two phases, namely the first quantitative phase followed by the second qualitative phase, which guided the investigation during the two phases. As such, it was possible to investigate the equivalence of the PIRLS Literacy 2016 and PIRLS 2016 released texts and items across English, Afrikaans and isiZulu.

Phase One included the quantitative section of the study and as such relied on the primary study's data whereby an overall score was provided followed by descriptive statistics and Rasch analysis. The findings from the first phase initiated further in-depth investigation during Phase Two of this study which entailed exploring the items which displayed differential functioning during the first phase as well as looking at the five released texts used during the PIRLS assessments. With this in mind, I considered the linguistic, functional, cultural and metric equivalence (Peña, 2007) of these texts and items.

Linguistic equivalence includes whether linguistic meaning of the ST and TT are the same (Chesterman, 2016) by usually employing the back-translation technique (Beck et al., 2003). The translator translates the ST into the TT, which is followed by a back-translation by a different translator (Behr, 2017). During the PIRLS assessments, the texts and items were developed in US English and each participating country had to translate, and where necessary, make adaptations to the texts. For the South African studies, the texts and items were changed to UK English and adaptations, such as units of measurement, were made. The texts and items were then translated into the different test languages. This study found that overall, the translations of the released texts were adequate. However, one text in particular, namely *Flowers on the Roof*, presented several translational issues in both Afrikaans and isiZulu. All of the texts and items seemed to pose some translation challenges. As is the nature of written language, all three languages, across the texts, made use of unfamiliar words which are not part of everyday speech, but as both the narrative and informational texts made use of richer vocabulary, it is expected that learners would come across mid or low frequency words. In very few instances, the isiZulu texts made use of loan words (borrowed from either English or Afrikaans). The use of loan words is not a translational issue *per se*, but the panellists did state that if the isiZulu word was available, it should have been included rather than using loan

words. Within this study, linguistic equivalence was linked with the meso and micro levels of equivalence as each text's items were evaluated in terms of the *Processes of Comprehension* as well as vocabulary and literary devices and whether these changed across the languages.

Functional equivalence entails the "naturalness" (Bermann & Porter, 2014, p. 72) of the translation(s). The goal of functional equivalence is to ensure that the ST and TT elicit the same behaviour (Greenfield et al., 2006). Therefore, functional equivalence connected with macro level as it entailed looking at whether the ST and TT were similar. For the PIRLS assessments, this meant that the English version and the translated versions should, from a measurement perspective, behave similarly, an issue that speaks to the content validity of instruments. In terms of the Rasch analysis, this study found that the released texts all displayed differential item functioning amongst the three languages. Even though some of the items showed differential functioning, it is important to note that there was not a clear pattern as the items did not universally favour one language nor did the texts discriminate universally against a particular language.

Next, cultural equivalence is considered. Cultural equivalence involves the meaning of the text and whether the construct is understood similarly in different cultures (Chan & So, 2017). As PIRLS is an ILSA, which is conducted across different linguistic and culture groups, it is crucial that the meaning of the texts and items be understood for each participating group. This study asked panellists to review the texts and items in terms of cultural equivalence. The literary and informational texts should be equally interpreted and understood by the different linguistic and cultural groups. Moreover, cultural equivalence is linked to the micro level of equivalence as it dealt with vocabulary and literary devices and whether these changed across the ST and TTs.

Metric equivalence focuses on the item difficulty of assessments (Peña, 2007). When a test is developed and translated, it should bear in mind the level of difficulty of the items; that is, whether the difficulty changes for the translated versions. As such, metric equivalence links to the meso level of equivalence as it focused on the item difficulty of each text. Metric equivalence of the PIRLS assessments included investigating whether the items changed in their difficulty across English, Afrikaans and isiZulu. Both the literary and informational texts had a good spread of items in terms of the four *Processes of Comprehension*. Moreover, the difficulty of the items remained the same across the languages. However, some items of each text did display differential functioning, as discussed under functional equivalence.

Figure 9.1 depicts the conceptual framework used during this study. The framework of this study provided guidance to the research study and enabled the researcher to work systematically through the different research questions. Phase Two of this study also included the macro, meso and micro levels of equivalence. Each of these equivalences focused on different aspects. For instance, macro level examined the overview of the texts by focusing on text characteristics and the layout of each text. The meso level looked at the reading comprehension processes as identified by the IEA as well as the item difficulty. The micro level focused on the vocabulary used in the text, average word and sentence length as well as the learners' responses on CR items.

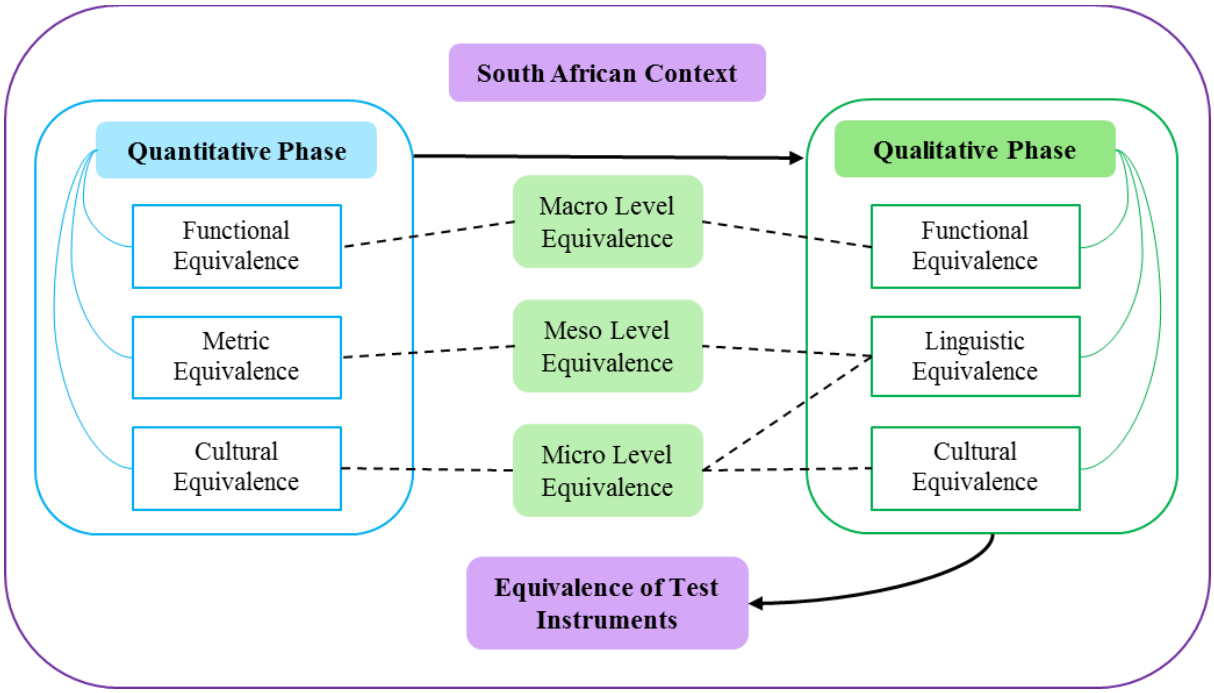

Figure 9.1: Translation Equivalence for Cross-cultural Research (informed by Peña, 2007)

This study provides new insights into the South African PIRLS Literacy 2016 and PIRLS 2016 assessments in terms of learner achievement, differential functioning of items, translation matters, and most importantly, whether these texts were equivalent across languages. Overall, the conceptual framework used during this study provided a suitable lens through which different levels of equivalence were explored in order to determine whether the PIRLS assessments were similar across different languages within the South African context.

As a final reflection based on this study's conceptual framework, which focused on translation equivalence, this study placed emphasis on a previously neglected area of study. Most studies

usually focus on systems theory, school improvement and school effectiveness models to explain the association between contextual variables and learner performance on an assessment such as PIRLS. This study, however, focused on measurement invariance and foregrounds factors such as the different types of equivalence of assessment instruments across different languages.

9.7 MAIN CONCLUSIONS

This section provides the main conclusions drawn, based on the findings that emerged from addressing the main research question of this study: ***To what extent are the PIRLS 2016 limited released texts in English, Afrikaans and isiZulu, in Grade 4 and Grade 5 equivalent?***

- **Main conclusion 1:** ***There is no measured evidence of consistent differential functioning of items for any particular language.***

One important aspect of ILSAs is to ensure that the test instruments are equal across the different languages and cultures (*cf*. ITC, 2017; Peña, 2007). This study examined the PIRLS Literacy 2016 and PIRLS 2016 released texts across English, Afrikaans and isiZulu to determine whether the items of each text functioned the same across the ST and TTs. The PIRLS Literacy 2016 and PIRLS 2016 released texts include: *The Pearl*; *African Rhinos and the Oxpecker Birds*; *Flowers on the Roof, Macy and the Red Hen;* and *The Green Sea Turtle's Journey of a Lifetime.*

Based on the descriptive analysis, it is evident that neither the South African Grade 4 learners nor the Grade 5 learners performed well on these texts. In order to further explore the descriptive results, Rasch analysis was utilised. The Rasch analysis indicated items that functioned differently across the English, Afrikaans and isiZulu learners (*cf*. Chapters Five and Six). One possible reason could be that the texts and items were not equivalent across the languages, and were thus cross-culturally unfair assessments. However, the findings of the Rasch analysis showed that the isiZulu learners found some items easy while Afrikaans and English learners did not (*cf*. Figure 5.8), or where English learners found the items easy but neither Afrikaans or isiZulu learners did (*cf*. Figure 5.12). There were also items where the Afrikaans learners found the item less difficult compared to the English or isiZulu learners (*cf*. Figure 6.10). There were also items that displayed extreme inconsistency across the three languages, which did not follow the model curve (*cf*. Figure 5.18). If the PIRLS Literacy 2016

and PIRLS 2016 assessments had been unfair in one of the three languages, the above kind of variation would not have occurred.

The null hypothesis which assumed that the English, Afrikaans and isiZulu texts and items are equal was rejected. Some of the texts included a high number of items that functioned differently among the three languages. The text with the highest number of DIF items is *African Rhinos and the Oxpecker Birds*, while two texts, namely *The Pearl* and *The Green Sea Turtle's Journey of a Lifetime*, had the lowest number of DIF items. However, as discussed in the previous paragraph, upon further examination of the items that displayed DIF, and specifically the ICC graphs, the Rasch analysis revealed no clear pattern as there was no consistent or universal discrimination against one language.

- **Main conclusion 2:** ***PIRLS 2016 and PIRLS Literacy 2016 released texts are similar to the texts prescribed for use by Grade 4 and 5 in schools in South Africa: Evidence of Face Validity.***

In terms of face validity and macro equivalence of the PIRLS released texts, it was found that there was consistency across the texts for each language. The overview of each text remained the same across each language, for the narrative texts, it included text characteristics such as genre, theme, and characters. For informational texts, it included the purpose and organisational elements. The majority of panellists agreed that the texts are similar to the texts that South African Grade 4 and 5 learners may encounter. A further investigation of each of the PIRLS released texts showed that the layout of the texts across the different languages was the same. The information pertaining to a question was located in the same paragraph on the same page irrespective of the language.

Furthermore, each of the visuals, whether pictures, photographs or maps were all located at precisely the same place for each language sub-group. The photographs[91] included in the informational texts were of high quality and clearly showed the different animals. However, some of the panellists argued that a few of the illustrations in *The Pearl* and *Macy and the Red Hen* could have been drawn with more precision; for example, in *The Pearl* the merchant,

[91] Note that the PIRLS Literacy 2016 released texts (including photographs, maps and diagrams) were only available in black and white while PIRLS 2016 released texts were available in colour.

Reuben, holds a pearl in his hand but the pearl is unusually large. In *Macy and the Red Hen*, the leftover scraps that Macy feeds the hens could have been more clearly illustrated.

- **Main conclusion 3:** ***The Processes of Comprehension remained the same across the languages: Evidence of meso level equivalence.***

Part of the *meso* level of equivalence required examining whether the PIRLS *Processes of Comprehension* changed during translation, in other words if these processes were different in English, Afrikaans and isiZulu. The four processes included focus on and retrieve explicitly stated information, make straightforward inferences, interpret and integrate ideas and information and evaluate and critique content and textual elements (Mullis & Martin, 2015). These processes are also linked to the perceived level of difficulty ranging from lower order comprehension skills to higher order comprehension skills. The first two processes are grouped into *Retrieving Explicitly Stated Information* and while the last two comprise *Inferencing, Interpreting and Evaluating* (Mullis et al., 2017).

During the investigation of possible changes in these processes during translation, it was found that none of the processes had changed (*cf*. Tables 7.7, 7.18, 7.39, 8.4 and 8.19). Therefore, the difficulty also remained the same as it tested the same reading comprehension skill across the three languages. Consistency in the processes provides partial evidence of equivalence across the languages.

- **Main conclusion 4:** ***The PIRLS Literacy 2016 and PIRLS 2016 released texts made use of a rich range of vocabulary: Evidence of micro level equivalence.***

During Phase Two, the vocabulary used during both PIRLS Literacy 2016 and PIRLS 2016 released texts was examined. This included the frequency of key words, as identified by the IEA, used in each text as well as the word and sentence length for each text. In addition, word profiles were included for each text; however, this was only established for the English texts as there is no software currently available to create word profiles of the Afrikaans and isiZulu texts.

The PIRLS texts, in terms of the frequency of key words, word and sentence length, showed remarkable consistency across the languages. Although, due to the isiZulu orthography, the average word length and sentence length differed, which is consistent with the language's orthography.

In terms of the word profile, it was found that the most of the English texts reached the text-

lexis threshold of 95% required for reading comprehension (*cf.* Laufer, 1989; Laufer & Ravenhorst–Kalovski, 2010; Nguyen & Nation, 2011; Schmitt et al., 2017) and as such confirms to the research done on vocabulary and reading comprehension. Even though there is no software available to conduct the word profile of the Afrikaans and isiZulu texts, these findings could possibly be extrapolated to the Afrikaans and isiZulu texts as the panellists highlighted the same words which were considered to be either mid or low frequency words.

For instance, some texts such *African Rhinos and the Oxpecker Birds*, *Macy and the Red Hen* and *The Green Sea Turtle's Journey of a Lifetime* included a higher percentage of mid-frequency words. These included words between the K3 and K9 frequency range (Schmitt & Schmitt, 2014), that is, between the 3 000 - 9 000 words as identified by the BNC-COCA corpus. For example, *pearl*(*s*), *oyster*(*s*), *rhinoceros*, *bloated*, *hissing*, *squint*, *cluck*, *fluffed*, *squawked*, *algae*, *herons* and *slurps* (*cf.* Table 7.4, 7.16, 7.37, 8.2 and 8.17). Moreover, these mid to low frequency words are similar across Afrikaans and isiZulu which suggests that the learners' poor reading comprehension could be partially due to the vocabulary used in each text and reflects the differences between spoken and written registers of language. Some of the vocabulary, especially those in the mid and low frequency word range, used in the texts do not typically occur in everyday spoken language. As such, if learners are not exposed to written texts at home or at school, they will not acquire these types of words.

Even though this study did not focus *per se* on the vocabulary used or on learners' vocabulary problems, it is included as part of the reflection of qualitative evidence generated during Phase Two of the study.

- **Main conclusion 5:** ***Despite evidence of meso level of equivalence, learners across languages struggled to provide written responses.***

In order to study the learners' constructed responses on the DIF items further, the CEA data manager provided a selection of achievement booklets covering each of the five released texts. These learners' written responses were captured in Chapters Seven and Eight in this study. The learner responses provided a better understanding of poor reading comprehension results during the PIRLS assessments.

The PIRLS Literacy 2016 and PIRLS 2016 assessments required learners to provide written responses to some items to a maximum of 3 points. The overall results of both assessments provided some initial evidence that learners struggled with providing written responses (*cf.*

Figures 5.1 and 6.1). Although most of the English, Afrikaans and isiZulu learners attempted to answer each item, some items included very high incorrect percentages (*cf*. Tables 7.8, 7.19, 7.27, 7.40, 8.5 and 8.20).

The PIRLS constructed response items tested both lower and higher order skills linked to the *Processes of Comprehension*, which include focus on and retrieve explicitly stated information, make straightforward inferences, interpret and integrate ideas and information and evaluate and critique content and textual elements (Mullis & Martin, 2015). Thus, the learners were required to provide a range of written responses ranging from copying an answer from the text to providing the author's perspective. Previous PIRLS reports (Howie et al., 2008; 2012) also indicated that learners struggled with constructed response items.

Overall, when the item required learners to seek the answer in the text and copy it as their answer, some learners were not able to do so and sometimes copied irrelevant answers. These learners may not have read or fully understood the question. The learners also struggled to provide longer responses requiring full sentences or complex sentences. This finding may also reflect on classroom practices, specifically whether teachers encourage learners to find text-based evidence for literal questions. Some learners, especially the Grade 4 learners, provided nonsensical answers where they randomly wrote letters. This scribbling could be due to attempts to hide their lack of writing or reading skills by appearing to be answering the item. Most learners were not able to answer more challenging CR questions, but neither could they answer easy questions; for example, questions that required them to do some perspective taking or completing a table. In other words, skills that are most associated with the ability to interpret and integrate findings or to evaluate textual elements are still mostly lacking.

- **Main conclusion 6:** ***Despite evidence of metric equivalence, learners across languages struggled with literal type questions.***

As described in Chapter Two, the PIRLS Literacy 2016 and PIRLS 2016 tested learners' reading comprehension skills according to the *Processes of Comprehension*. These processes include focus on and retrieve explicitly stated information, make straightforward inferences, interpret and integrate ideas and information and evaluate and critique content and textual elements (Mullis & Martin, 2015). The PIRLS Literacy 2016 items comprised more literal, lower order type items, whereas the PIRLS 2016 items added fewer literal items and more higher order skill items.

This study indicated which items functioned differently across English, Afrikaans and isiZulu languages and these items were further examined to check whether the more lower order or higher order items indicated DIF. Note that PIRLS achievement scores are divided into two subscales namely *Retrieving Explicitly Stated Information* and *Inferencing, Interpreting and Evaluating*. The former includes the two lower order comprehension skills namely focus on and retrieve explicitly stated information and make straightforward inferences, while the latter includes higher order reading comprehension skills which includes interpret and integrate ideas and information as well as evaluate and critique content and textual elements (for a detailed explanation see Mullis et al., 2017).

Literal items are text based and straightforward and are part of the definitions of lower order skills (Bloom, 1956; Krathwohl, 2002). When completing these items, learners are not required to move beyond the literal interpretation of the text. Therefore, these questions are expected to be easy since the information is explicitly stated in the text. In contrast, higher order items require learners to demonstrate an understanding of the text by integrating ideas from across the text, to include their own interpretations, and to construct meaning of what they are reading (McLeod Palane, 2017). Unexpectedly, South African learners performed poorly on lower order questions. One possible explanation could be that the learners may have some form of reading difficulty or disability (Mahdavi & Tensfeldt, 2013), or that learners have not yet mastered decoding skills, which is essential to reading comprehension (Pretorius, 2014; Spaull et al., 2020). Another possible reason for learners struggling with these basic questions could be due to external factors, which could include the teaching methods employed by teachers as well as lack of parental involvement in their children's reading and literacy development (Mohammed & Amponsah, 2018; Pretorius, 2014; Pretorius & Klapwijk, 2016; Roux; 2014).

In terms of the PIRLS Literacy texts, all 17 items across the released texts indicated DIF, of which 12 items tested lower order comprehension skills. Of the 12 items, half tested learners' ability to seek the answer from the text and provide it as their answer, while the other half required learners to make straightforward inferences. Nine of these items were multiple choice (MC) type items and the remaining eight were constructed response (CR) items. As such, the type of item (either CR or MC) may only play a small role in the possible reason for learners struggling with lower order comprehension skills.

The PIRLS 2016 DIF results also indicated 17 items that functioned differently across the English, Afrikaans and isiZulu languages. Nine of the DIF items were literal items; five items

required learners to retrieve explicitly stated information in the text while four items required leaners to make straightforward inferences. In terms of the two question formats, MC and CR, the majority (11) of the 17 DIF items took the form of CR while only six items were MC. This finding differs with the Grade 4 results and points to possible instructional issues. Teachers may not know how to teach learners to read with comprehension and find answers in the text or move beyond the literal interpretation of the text. Higher order questions require learners to answer questions that require higher order skills, such as to generate their own answers by relying on the ideas or information contained in the text as well as constructing meaning from what they have read.

9.8 METHODOLOGICAL REFLECTIONS

Mixed methods research includes both quantitative and qualitative methods in order to investigate a research problem from more than one dimension (Plano Clark & Ivankova, 2016; Tashakkori & Teddlie, 2003); as such, it is considered a more robust analysis (Green & Caracelli, 1997). However, as this study utilised a mixed methods design, it is important to reflect on both the quantitative as well as the qualitative methods of this study.

As the first quantitative phase of this study entailed a secondary analysis of the PIRLS Literacy 2016 and PIRLS 2016 South African data, it is necessary to discuss the perceived limitations thereof. Secondary analysis refers to the re-analysis of data which was collected by a different researcher for another purpose (Johnston, 2014). Even though secondary analysis enables a secondary researcher to analyse existing data to answer a new research question (Cheng & Phillips, 2014), the purpose of the original study may be different. Moreover, the scope and depth of the existing data may not be adequate for the secondary researcher. However, as the aim of this study was to determine whether the PIRLS assessments are equivalent across three different languages, it could be argued that this research study extends the validity of the original assessments within South Africa and may have addressed some of these concerns. The availability and quality of the original study's data could also be a cause for concern (Vartanian, 2011). In terms of availability, the PIRLS datasets are available for scholars, students and other interested parties; however, these persons are required to apply for permission in order to use the data and permission is given by the IEA and, where necessary, the participating countries' NRCs. Furthermore, the PIRLS assessments underwent several quality assurance measures, as discussed in Chapter Two. Another possible limitation could be that the secondary researcher is far removed from the original study (Zimmerman, 2010). Yet as I was involved during both

PIRLS Literacy 2016 and PIRLS 2016, during instrument development and refinement, translations, data interpreting and report writing as well as receiving scoring training, I may have dealt with this concern.

This study, as part of the secondary data analysis, conducted both descriptive and Rasch analysis. Descriptive research was conducted to describe the data by providing a summary of the data as well as to organise it in a meaningful way (Pietersen & Maree, 2016). After the initial analyses were conducted, Rasch analysis was employed.

The Rasch Model is centred on a person or persons' probability of success on an item (Linacre 2013). It uses a scale where the persons', or in this case the learners' ability and the item difficulty scores are compared. A major advantage of using Rasch Measurement Theory is that it can examine the possible measurement invariance of an assessment instrument (*cf.* Andrich 2011; Linacre 2016). In addition, by utilising Rasch Measurement, I was able to test the PIRLS Literacy 2016 and PIRLS 2016 released texts (contained in achievement booklets) to determine whether the items behaved the same across different languages.

The second phase of the research involved a qualitative analysis of the PIRLS Literacy 2016 and PIRLS 2016 released texts and associated items. The qualitative phase was informed by the findings of the first phase of the study.

During the second phase, various data collection and analysis techniques were applied to best analyse the PIRLS texts and items. By conducting a sequential explanatory mixed methods design, I was able to further analyse problematic items with the assistance of a panel of experts. Content analysis was conducted on the five released texts and included scrutinising the translation of the texts. As part of content analysis, the panellists were requested to complete an open-ended questionnaire for each text that included items about the translation of the texts, construct, content and face validity of the texts as well as the incidence of low frequency words. In addition, workshops were held where the panellists discussed each of the above. After these discussions and exploring the panellists' answers in the questionnaires, a decision was made to further analyse the learners' responses by inspecting a random selection of PIRLS Literacy 2016 and PIRLS 2016 booklets (provided by the CEA). The learners' responses were captured and provided in Chapters Seven and Eight. As such, the qualitative phase provided a deeper understanding of the South African PIRLS Literacy 2016 and PIRLS 2016 results.

Nevertheless, there is room for improvement. This study only scratched the surface of the possible reasons why the South African learners have performed poorly on the PIRLS assessments over the various cycles. It would have been more beneficial if a larger panel of Afrikaans and isiZulu experts, from different provinces, took part in this study in order to provide a more holistic view of each language's texts and items. Had the study included experts from different provinces, valuable insights might have emerged on the dialect issue from an ILSA perspective. Furthermore, this study only focused on English, Afrikaans and isiZulu which, while it was ambitious, only provided a glimpse of the equivalences amongst South African languages; however, the inclusion of the other African languages could have provided a more comprehensive understanding of the PIRLS 2016 results.

9.9 STRENGTHS AND LIMITATIONS OF THE STUDY

This study was conceptualised to answer questions surrounding the PIRLS South African results, specifically item bias in the national versions of the PIRLS assessments. In order to address issues related to possible bias, I systematically examined the PIRLS Literacy 2016 and PIRLS 2016 results by utilising a mixed methods approach so as to look at the results from different perspectives. This study design is a departure from traditional secondary analysis designs of ILSA data as it made use of both quantitative and qualitative methods (*cf.* Section 4.2.2). During Phase One, I explored the quantitative data and based on the results, continued with a qualitative probe (Phase Two) of the released texts as well as examining a selection of learners' responses in both PIRLS Literacy and PIRLS booklets.

During the quantitative phase, I made use of the PIRLS Literacy 2016 and PIRLS 2016 datasets provided by the CEA at the University of Pretoria. As I only investigated English, Afrikaans and isiZulu, the data manager removed the other languages from the PIRLS Literacy 2016 dataset. Note that PIRLS 2016 only tested Grade 5 learners in English, Afrikaans and isiZulu. Both final datasets used in this study had a large sample size.

The first phase included calculating percentages and means as well as conducting a Rasch analysis. The percentages and means were conducted by using the IDB-analyzer developed by the IEA. The Rasch analysis was conducted using the statistical programme RUMM2030.

A possible limitation of the quantitative phase is that even though the sample sizes were large, the focus was only on English, Afrikaans and isiZulu languages, which are not generalisable to the entire Grade 4 or 5 South African population.

In terms of the second, qualitative phase, I gathered a panel of experts who reviewed and provided feedback on the released texts as well as items which had displayed DIF during the first phase of the study. Although the panel of experts who assisted in this study were language specialists, future studies should include experienced Intermediate Phase teachers. During the second phase, workshops were held with the panel of experts to discuss each text. These experts were also asked to complete open-ended questionnaires for each of the released texts. As the study continued, it was necessary to examine some of the learners' responses in the PIRLS Literacy 2016 and PIRLS 2016 booklets.

A limitation of the qualitative phase include that I am only an Afrikaans-English bilingual and not multilingual speaker and so had to find isiZulu experts to assist with the exploration of the isiZulu texts, items and learner responses. Another limitation of this phase included the number of learner booklets examined, only ten answers for each problematic item across the five texts were included. It would prove insightful if a larger sample of booklets were examined across each language and text.

9.10 POLICY, PRACTICE AND RESEARCH RECOMMENDATIONS

Based on the findings and conclusions of this study, the following section presents recommendations relating to policy (9.10.1), test development (9.10.2), practice (9.10.3) and further research (9.10.4).

9.10.1 Policy Recommendations

It is widely acknowledged that theories of translation, including adaptations, have made advancements over the years (*cf.* Dore, 2016; Gentzler, 2001; Rios & Sireci, 2014; van de Vijver & Leung 1997). Some of the translation changes could partially be attributed to ILSAs (ITC, 2017) as they are usually cross-cultural studies researching a specific phenomenon. However, when ILSAs are conducted in countries where multiple languages are present and which may have dialects, it is important to carefully consider the translation and adaptations of the instruments. The insights gained from this research indicated that developing equivalent translated versions of the original instrument, especially in the case of informational texts, might be difficult as the corpus of some languages may not yet have the relevant words or technical terms and as a result, make use of the nearest synonym or loan words. Even though this study did not specifically investigate dialects, or the possible effect dialects might have on

learner achievement, it was nevertheless worthy to make this suggestion as it is pertinent in countries which have dialects (*cf*. Stubbe, 2011).

In terms of policy, the South African Department of Basic Education should consider developing more informational texts across the African languages to ensure that learners also receive exposure to different types of texts besides narrative texts. In addition, the current text length recommendations in CAPS should be revisited, as it should be more in line with international expectations of the length of texts that Grade 4 and 5 learners are expected to read. Moreover, language bodies such as the Pan South African Language Board (PanSALB) should be included when cross-lingual and cross-cultural instruments are developed. It should be noted that nationally, the South African Government has developed a draft policy that includes the incremental introduction of African languages (IIAL) in schools (DBE, 2013). The IIAL draft policy was developed to promote social cohesion within the South African society as well as the develop African languages which were previously marginalised during apartheid (DBE, 2013). This is a commendable policy initiative from government.

Furthermore, policy should recommend that when local or international assessment instruments are developed, language bodies such as PanSLAB should provide inputs as each language and cultural group should be treated fairly. In terms of PIRLS, the local NRC consults various organisational bodies when developing the translated versions of the instruments. However, these consultations should include a representative of each language tested.

9.10.2 Test Development Recommendations

It would be useful if guidelines could be developed for ILSAs, or even national assessments and evaluations, to take into consideration possible limitations of the participating language groups' corpus. Usually ILSAs develop their testing instruments in English which is then translated into the TT. As such these instruments are designed from a western perspective. It may, however, prove difficult to develop guidelines which are more flexible as it will leave room for the different participating countries to make interpretations of the text beyond that which the author meant. Furthermore, as ILSAs make use of texts usually already published, it may impose on the copyright of the authors, as the text cannot be changed beyond what the authors originally write.

Another aspect to consider during translations and adaptations of instruments is to employ appropriate translators who have relevant experience and qualifications (Chan & So, 2017).

Moreover, it would be advantageous to only ask translators who understand the finer nuances of the linguistic discourse which need to be translated. For example, in literacy assessments the translators should have experience in the SL and TT as well as knowledge of children's literature.

In South Africa, which is still in its emerging stage of developing its own systemic evaluations for internal, national testing purposes, it would be prudent to develop standard practical guidelines for ensuring translation equivalence across the different languages, especially in the Foundation Phase as the majority of learners have an African language as their home language and LoLT at school. When national systemic tests or other evaluations are planned, the qualification and accreditation of the translators should be checked. The translators should also be checked if they have experience with translating literary and informational texts within the educational sector.

9.10.3 Pedagogic Practice Recommendations

Some practice recommendations for the classroom environment should also be considered. In terms of teacher pre-service training as well as in-service workshops, emphasis should be placed on how to teach reading comprehension and ensuring the teachers teaching language know how to introduce and teach the language as a second language. Although additional training and workshops are included as part of professional development in South Africa, these activities could possibly assist in the learners' language acquisition, and possibly enhance their reading comprehension skills.

In addition to the enhancement of reading skills, teachers should also place emphasis on learners' writing abilities. The link between reading and writing is important and this link could be strengthened during pre-service as well as in-service training. During the training sessions, teachers could be trained to encourage learners' vocabulary growth, sentence construction, enable learners to self-correct mistakes, assist learners in forming coherent sentences, look at the logical flow of information and to present their writing in a structured format.

In the language classroom, teachers should, where possible, expose learners to a variety of texts and not only focus on narrative texts. In recent years, South African NGOs such as Nal'ibali and African Storybook, have created and provided stories in African languages, which are freely downloaded and used as a possible resource within the classroom. However, many South African schools and teachers may not have a variety of resources to use during lessons, and as

such, teachers should be trained on how to creatively use the resources that are available to them.

Furthermore, it would be valuable to expose teachers to ILSAs to provide them with an idea of the level of expectation of what is required of the learners in order to perform well on such assessments. As part thereof, teachers could be informed about the different cognitive levels tested during an ILSA in order to provide an accurate picture of the learners' attainment. Moreover, if teachers gain some exposure to ILSAs, it could provide them with a wide-ranging view of how different question types are used. It may also be worthwhile to ensure that learners are exposed to testing to familiarise them with different types if items, the time allocation and increase their test-stamina.

9.10.4 Further Research Recommendations

This study utilised mixed methods research to find evidence for the equivalence of the PIRLS Literacy 2016 and PIRLS 2016 released texts across English, Afrikaans and isiZulu. However, as this study only focused on linguistic, functional, cultural and metric equivalences, the following recommendations for further research are made:

- A study that focuses on the learners' responses to constructed response items. This should focus on both literal and inferential items as South African learners during the PIRLS assessments struggled with both these types of items. This could possibly be done by testing a sample of learners, scoring the responses and conducting interviews with teachers about how they teach learners to answer reading comprehension questions.
- A study which further investigates possible linguistic and functional equivalence by exploring the translated PIRLS assessments in terms of dialects. By doing so, it may provide valuable insights into the extent to which dialects may play a role during formal testing.
- A study is recommended to investigate the translations of the eight remaining African languages. It is possible that translation challenges occurred for each of these languages and could partially explain the poor results of the reading comprehension test. This could possibly be done by examining the PIRLS texts at the macro, meso and micro levels of equivalence which would include panellists for each language.

- A study investigating learners' vocabulary in each of the official languages at the end of the Foundation Phase could be considered. The PIRLS assessments made use of a rich vocabulary with low, middle and high frequency words. Even though some words according to the BNC-COCA corpus are considered high frequency and should be commonly known to English learners, the same may not be applicable to other languages.
- Even though issues of decoding are not within the PIRLS ambit, the South African learners' decoding abilities should be investigated, especially in light of their poor performance with literal questions. If learners cannot decode with appropriate accuracy and ease, they are unlikely to understand much of what they have read. Thus, the learners' decoding skills might, to some extent, account for their poor performance during the PIRLS assessments.

Even though this study provided an in-depth look at the PIRLS assessments, it only focused on three of the 11 official languages and did not consider contextual factors. As such, it may be worthwhile to conduct other multivariate analysis of the PIRLS assessments by looking at contextual factors.

9.11 CONTRIBUTIONS OF THE STUDY

This study took the form of a mixed methods research design in order to utilise both quantitative and qualitative approaches to best address the research problem (*cf.* Plano Clark & Ivankova, 2016). In addition, by using both quantitative and qualitative methods it added more depth and breadth to the findings as the quantitative findings (Phase One) informed the qualitative findings (Phase Two), meaning that the second phase of the study extended the first phase of the study (Creswell, 2013).

By using a sequential explanatory mixed methods design (Plano Clark & Ivankova, 2016), this study firstly determined whether measurement invariance existed among the English, Afrikaans and isiZulu learners during the PIRLS Literacy 2016 and PIRLS 2016 assessments. Secondly, based on the findings of Phase One, it was then possible to further investigate the texts and more specifically, the items which displayed differential functioning between the languages. Thus, this study was able to provide richer research findings due to the design of the study, which set out to determine whether the released texts were equivalent across different languages.

Through the pragmatic lens, I was able to apply and extend on Peña's (2007) considerations for cross-cultural research. As ILSAs involve testing participants across different languages as well as different cultural groups, it is vital to take cognisance of the differences between these languages and cultures by examining the linguistic and cultural equivalences of ILSA instruments. In addition to linguistically and culturally equivalent instruments, other considerations such as functional and metric equivalence should be considered as they can highlight possible content and construct validity issues.

Based on the above statements, it is important for ILSAs to ensure that the instruments, whether they be standardised tests or questionnaires, be equivalent and not be more advantageous to one group than another. The PIRLS assessments have taken great strides to ensure good quality cross-cultural assessments; however, by taking into consideration the four equivalences, it could assist further in safeguarding against possible threats to validity.

In addition, ILSA results, such as the PIRLS results, should not only be discussed at the result releases in each participating country, such as South Africa, but rather be carefully scrutinised for their contribution, meaning and better understanding of the local context to inform policy debates, policy reform as well as for evidence-based decision making. For instance, if the South African education sphere takes into consideration the value, as mentioned above, of ILSAs, the country can move beyond looking at league tables (van Staden et al., 2019) and focus on drafting and implementing policies for the betterment of South African education. Then ILSAs will be able to fulfil their mandate as a tool for monitoring the education system (van Staden, *in press*).

9.12 CLOSING THOUGHTS

As indicated at the beginning of this study, the South African Grade 4 and 5 learners performed poorly on the PIRLS assessments. As such, it was important to investigate possible reasons for learners' poor performance in the reading comprehension assessments. In order to partially assist in answering this research problem, this study aimed to determine whether the PIRLS assessments are equivalent across three languages, namely English, Afrikaans and isiZulu.

When conducting ILSAs, the research organisation should keep in mind several aspects to ensure that their assessment instruments are equivalent across different linguistic and cultural groups. These include linguistic, functional, cultural and metric equivalences as well as cross cutting macro, meso and micro levels of equivalence. In so doing, it could guard against

possible threats to the validity of the assessment instruments. These four equivalences form part of the overarching 'translation equivalence' concept. There are many different ways of conducting translations, but adherence to the four equivalences could assist in producing an equivalent instrument that is similarly understood by the source and target reader. In conclusion, "Translating literary works … requires much more than exchanging one word for another, more than trading sentences between languages. It takes an understanding of the author's message to capture the vision, the essence, of the writing" (McLeod, 2016, para. 1).

-⌘-

REFERENCES

Abdi, H. (2010). Guttman scaling in N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 559-560). SAGE Publications. https://dx-doi-org.uplib.idm.oclc.org/10.4135/9781412961288.n173

Ackermann, C. (2017). *The effect of parental involvement on Grade 4 learner literacy: Evidence from prePIRLS 2011* [Master's Thesis, University of Pretoria]. UP Space. https://repository.up.ac.za/handle/2263/65433

Afflerbach, P., & Cho, B. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. Israel & G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 69-90). Routledge.

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1061-1070.

American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing* (2014 ed.). American Psychological Association

Anastasi, A. (1988). Psychological testing (6th ed.). Macmillan.

Anderson, R., & Pearson, P. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. Pearson (Ed.), *Handbook of reading research* (pp. 255-291). Longman.

Andrich, D. (2011). *Rasch models for measurement*. SAGE Publications.

Andrich, D., & Marais, I. (2019). *A course in Rasch measurement theory: Measuring in the educational, social and health sciences*. Springer.

Andrich, D., Sheridan, B. E., & Luo, G. (2012). *RUMM2030: Rasch unidimensional models for measurement*. RUMM Laboratory.

Archibald, M. M., Radil, A. I., Zhang, X., & Hanson, W. E. (2015). Current mixed methods practices in qualitative research: A content analysis of leading journals. *International Journal of Qualitative Methods*, *14*(2), 5-33.

Arfé, B., Mason, L., & Fajardo, I. (2018). Simplifying informational text structure for struggling readers. *Read Writ*, *31*, 2191-2210. https://doi.org/10.1007/s11145-017-9785-6

Arffman, I. (2013). Problems and issues in translating international educational achievement tests. *Educational Measurement: Issues and Practice*, *32*(2), 2-14.

Babbie, E. (2007). *The practice of social research* (11th ed.). Thomson Wadsworth.

Baker, L. & Beall, L. (2009). Metacognitive processes and reading comprehension. In S. Israel & G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 373-388). Routledge.

Bamgbose, A. (2011). A*frican languages today: The challenge and prospects for empowerment under globalisation*. In E.G. Bokamba, R.K. Shosted & B.T Ayalew (Eds.), Selected Proceedings of the 40th Annual Conference on African Linguistics: African Languages and Linguistics Today (pp. 1-14). Cascadilla Proceedings Project.

Banda, F. (2009). Critical perspectives on language planning and policy in Africa: Accounting for the notion of multilingualism. *Stellenbosch Papers in Linguistics PLUS*, *3*, 1-11.

Beck, C. T., Bernal, H., & Froman, R. D. (2003). Methods to document semantic equivalence of a translated scale. *Research in Nursing & Health*, *26*(1), 64-73. https://doi.org/10.1002/nur.10066

Behr, D. (2017). Assessing the use of back translation: the shortcomings of back translation as a quality testing method. *International Journal of Social Research Methodology*, *20*(6), 573–584. https://doi-org.uplib.idm.oclc.org/10.1080/13645579.2016.1252188

Bergman, M. M. (2008). *Advances in mixed methods research*. SAGE Publications. https://dx-doi-org.uplib.idm.oclc.org/10.4135/9780857024329

Bermann, S., & Porter, C. (Eds.). (2014). *A companion to translation studies: Companion to translation studies*. John Wiley & Sons, Incorporated. Retrieved from https://ebookcentral-proquest-com.uplib.idm.oclc.org

Beukes, A. M. (2009). Language policy incongruity and African languages in postaparheid South Africa. *Language Matters*, *40*(1), 35-55. https://doi.org/10.1080/10228190903055550

Blaikie, N. (2004). Ontology, ontological. In M. S. Lewis-Beck, A. Bryman & T. F. Liao (Eds.), *The SAGE encyclopedia of social science research methods* (Vol. 1, pp. 767-767). SAGE Publications. https://dx-doi-org.uplib.idm.oclc.org/10.4135/9781412950589.n663

Blaikie, N. (2011). Epistemology. In M. S. Lewis-Beck, A. Bryman & T. F. Liao (Eds.), *The SAGE encyclopedia of social science research methods* (p. 310). SAGE Publications. https://dx-doi-org.uplib.idm.oclc.org/10.4135/9781412950589.n281

Bloom, B. S. (1956). *Taxonomy of educational objectives: Handbook the cognitive domain*. David McKay.

Boer, D., & Fischer, R. (2010). Towards a holistic model of functions of music listening across cultures: A culturally decentred qualitative approach. *Psychology of Music, 40*(2) 179-200.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.

Bonnet, G. (2002). Reflections in a critical eye: On the pitfalls of international assessment: Knowledge and skills for life: first results from PISA 2000. *Assessment in Education, 9*, 387-400.

Boone, W. J., Staver, J. R. & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer

Bornstein, R. F. (2004). Face Validity. In M. S. Lewis-Beck, A. Bryman & T. F. Liao (Eds.), *The SAGE encyclopedia of social science research methods* (Vol. 1, p. 368). SAGE Publications. http://dx.doi.org/10.4135/9781412950589.n323

Borsboom, D., & Mellenbergh, G. J. (2007). Test validity in cognitive assessment. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 85-115). Cambridge University Press.

Bos, L. T., de Koning, B. B., van Wesel, V., Boonstra, A. M., & van der Schoot, M. (2015). What can measures of text comprehension tell us about creative text production?. *Read Writ, 28*, 829-849.

Boucher, H. C. (2010). Understanding Western-East Asian Differences and Similarities in Self-Enhancement. *Social and Personality Psychology Compass*, *4*(5), 304-317, https://doi.org/10.1111/j.1751-9004.2010.00266.x

Britt, M., Goldman, S., & Rouet, J. (Eds.). (2012). *Reading: From words to multiple texts*. Routledge.

Brook Napier, D. (2011). Critical issues in language and education planning in twenty first century in South Africa. *US-China Education Review, 1*, 58-76.

Bryman, A. (1988). *Quantity and Quality in Social Research*. Routledge.

Bryman, A. (2001). *Social Research Methods*. Oxford University Press.

Campbell, J. R., Kelly, D. L., Mullis, I. V. S., Martin, M. O., & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001* (2nd ed.). Boston College.

Cekiso, M., Meyiwa, T. & Mashige, M. (2019). Foundation Phase teachers' experiences with instruction in the mother tongue in the Eastern Cape. *South African Journal of Childhood Education*, *9*(1), a658. https://doi.org/10.4102/sajce.v9i1.658

Chan, D. N. S., & So, W. K. W. (2017). Translation and validation of translation in cross-cultural research: strategies used in a study of cervical cancer screening among ethnic minorities. *International Journal of Nursing Practice*, *23*(6), e12581. https://doi.org/10.1111/ijn.12581

Cheng, H. G., & Phillips, M. R. (2014). Secondary analysis of existing data: opportunities and implementation. *Shanghai Archives of Psychiatry*, *26*(6), 371-375.

Chesterman, A. (2016). *Memes of Translation: The Spread of Ideas in Translation Theory. Revised Edition*. John Benjamins Publishing Company. Retrieved from http://web.a.ebscohost.com.uplib.idm.oclc.org/ehost/ebookviewer/ebook/bmxlYmtfXzExNzIxODBfX0FO0?sid=acd6aca5-6bc3-48ef-8300-5e6123c082ee@sdc-v-sessmgr01&vid=0&format=EB&rid=1

Chisholm, L. (2005). The politics of curriculum review and revision in South Africa in regional context. *Compare: A Journal of Comparative and International Education*, *35*(1), 79-100.

Christie, P. (1996) The right to learn: The struggle for education in South Africa (2nd ed.). SACHED; Ravan.

Chunhua, M. A. (2014). Translation of children's literature from the perspective of functional equivalence. *Studies in Literature and Language*, *9*(3), 151-155.

Cizek, G. J. (2012). Defining and distinguishing validity: Interpretations of score meaning and justification of test use. *Psychological Methods*, *17*(1), 31-43.

Clark, R. & Maynard, M. (2012). Research methodology: using online technology for secondary analysis of survey research data – "act globally, think locally". In J. Goodwin, (Ed.), *SAGE secondary data analysis* (pp. v2-37-v2-54). SAGE Publications. https://dx.doi.org/10.1177/089443939801600108

Coetzee-van Rooy, S. (2018) Dominant language constellations in multilingual repertoires: Implications for Language-in-Education Policy and practices in South Africa. *Language Matters*, *49*(3), 19-46. https://doi.org/10.1080/10228195.2018.1493613

Cohen, J., McCabe, L., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, practice and teacher education. *Teachers College Record, 111*(1), 180-213.

Cohen, L., Manion, L., & Morrison, K. (2008). *Research methods in education* (6th ed.). Routledge.

Colina, S., Marrone, N., Ingram, M., & Sánchez, D. (2017). Translation quality assessment in health research: A functionalist alternative to back-translation. *Evaluation & the Health Professions*, *40*(3), 267–293. https://doi.org/10.1177/0163278716648191

Combrinck, C. (2019). *Rasch unidimensional measurement models: Analysing data in RUMM 2030 hand-out*. Centre for Evaluation and Assessment.

Combrinck, C. M. (2018). *The use of Rasch Measurement Theory to address measurement and analysis challenges in social science research* [Doctoral Thesis, University of Pretoria]. UP Space. https://repository.up.ac.za/handle/2263/67982

Combrinck, C., Roux, K., Tshele, M., Mokoena, G., & McLeod Palane, N. (2017). Research Design and Methods in PIRLS 2016. In S. Howie, C. Combrinck, M. Tshele, K. Roux, N. McLeod Palane & G. Mokoena. *PIRLS 2016 Progress in International Reading Literacy Study 2016 Grade 5 Benchmark Participation: South African Children's Reading Literacy Achievement* (pp. 29-47). Centre for Evaluation and Assessment.

Creamer, E. G. (2018). *An introduction to fully integrated mixed methods research*. SAGE Publications.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed method approaches* (2nd ed.). SAGE Publications.

Creswell, J. W. (2013). *Steps in conducting a scholarly mixed methods study*. DBER Speaker series. University of Nebraska Discipline-Based Education Research Group (Online) Retrieved from: http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1047&context=dberspeakers Accessed 27th December 2019.

Creswell, J. W., & Plano Clark, V. L. (2008). *The mixed methods reader*. SAGE Publications.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). SAGE Publications.

Creswell, J. W., Shope, R., Plano Clark, V. L., & Green, D. O. (2006). How interpretive qualitative research extends mixed methods research. *Research in the Schools*, *13*(1), 1-11.

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. SAGE Publications.

Cumming, A. H., & Berwick, R. (1996). *Validation in language testing*. Multilingual Matters Ltd.

Cummins, J. (1979). Cognitive/Academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, *19*, 121-129.

Cummins, J. (2000). *Language, power and pedagogy: bilingual children in the crossfire*. Multilingual Matters Ltd.

Cummins, J. (2001) Instructional conditions for trilingual development. *International Journal of Bilingual Education and Bilingualism*, *4*(1), 61-75.

Cummins, J. (2008). BICS and CALP: empirical and theoretical status of the distinction. In B.H. Street & N.H. Hornberger (Eds.), *Encyclopaedia of language and education* (2nd ed.), Volume: 2 Literacy (pp.71-83). Springer Science; Media LLC.

Cummins, J. (2014). Bilingualism language proficiency and metalinguistic development. In P. Homel, M. Palij & D. Aaronson. *Childhood bilingualism* (pp. 71-88). Psychology Press.

Dass, S., & Rinquest, A. (2017). School fees. In F. Veriava, A. Thom & T. F. Hodgson (Eds.), *Basic education rights handbook: Education rights in South Africa*. SECTION27. http://section27.org.za/wp-content/uploads/2017/02/Chapter-7.pdf

Davies, A., & Elder, C. (2005). Validity and validation in language testing. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 795-813). Lawrence Erlbaum Associates.

De Beaugrande, R. A. (1995). Text linguistics. In J. Verschueren, J. O., Östman & J. Blommaert (Eds.), *Handbook of Pragmatics* (pp. 536-544). John Benjamins.

De Beaugrande, R., & Dressler, W. (1981). *Introduction to text linguistics*. Longman Inc.

de Kadt, E. (2005). English, language shift and identities: a comparison between 'Zulu-dominant' and 'multicultural' students on a South African university campus. *Southern African Linguistics and Applied Language Studies*, *23*(1), 19-37. https://doi.org/10.2989/16073610509486372

Dent, G. R., & Nyembezi, C. L. S. (2003). *Scholar's Zulu dictionary*. Shuter & Shooter.

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. Praeger.

Department of Basic Education (DBE). (2009). *Report of the task team for the review of the implementation of the National Curriculum Statement*. Government Printer.

Department of Basic Education (DBE). (2011). *Curriculum and Assessment Policy Statement Grades 4-6: English Home Language*. Government Printing Works.

Department of Basic Education (DBE). (2012). *National protocol for assessment in grades R-12*. Government Printing Works.

Department of Basic Education (DBE). (2013). *The incremental introduction of African Languages at Schools* (Draft Policy). https://www.education.gov.za/Portals/0/Documents/Reports/IIAL%20Policy%20September%202013.pdf?ver=2014-04-09-162048-000

Department of Education (DoE). (1997). *Language in education policy*. Department of Education. https://www.gov.za/sites/default/files/gcis_document/201409/languageeducationpolicy19971.pdf

Department of Education (DoE). (2002) *Revised national curriculum statement for schools: grades R–9*. Overview. Government Printer.

Department of Education (DoE). (2003). Revised national curriculum statement grades R-9 (Schools). Government Printer.

Department of Education (DoE). (2009). *Report of the task team for the review of the implementation of the national curriculum statement*. Government Printer.

Dore, M. (Ed.) (2016). *Achieving consilience: Translation theories and practice*. Cambridge Scholars Publishing.

Downe-Wamboldt, B. (1992). Content analysis: Method, applications, and issues. *Health Care for Women International*, *13*, 313-321.

Doyle, L., Brady, A.-M., & Byrne, G. (2016). An overview of mixed methods research - revisited. *Journal of Research in Nursing*, *21*(8), 623-635. https://doi.org/10.1177/1744987116674257

du Plessis, E. (2013). *Introduction to CAPS*. Curriculum and Instructional Studies, UNISA.

du Plessis, G. A., Saccaggi, C. F., & du Bruin, G. P. (2015). Cross-cultural differences in the character strength of citizenship in South Africa. *Psychology in Society*, *48*, 22-44. http://dx.doi.org/10.17159/2309-8708/2015/n48a2

Duke, N., & Carlisle, J. (2011). The development of comprehension. In M. L. Kamil, P. D. Pearson, E. B. Moje & P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4, pp. 199–228). Routledge.

Dymock, S. (2005). Teaching expository text structure awareness. *The Reading Teacher*, *59*(2), 177-182.

Ebbs, D., & Wry, E. (2017). Translation and layout verification for PIRLS 2016. In M. O. Martin, I. V. S. Mullis & M. Hooper (Eds.), *Methods and procedures* (pp. 7.1-7.15). Boston College.

Education Testing Service (ETS). (2014). *Standards for Quality and Fairness*. Educational Testing Service.

Eitel, A., & Scheiter, K. (2015). Picture or text first? Explaining sequence effects when learning with pictures and text. *Educational Psychology Review*, *27*(1), 153-180. https://doi.org/10.1007/s10648-014-9264-4

Eleyan, D., Othman, A., & Eleyan, A. (2020). Enhancing software comments readability using Flesch reading ease score. *Information*, *11*(9), 430-430. https://doi.org/10.3390/info11090430

Elley, W. B. (1992). *How in the world do students read? IEA study of reading literacy*. IEA.

Engelhard, G. (2013). *Invariant Measurement: Using Rasch Models in the Social, Behavioral, and Health Sciences*. Routledge.

Erkut, S., Alarcón, O., Coll, C. G., Tropp, L. R., & García, H. A. V. (1999). The dual-focus approach to creating bilingual measures. *Journal of Cross-Cultural Psychology*, *30*, 206-218.

Evans, R. (2017). Koeksister op 'n royal Albert piering: 'n Kritiese nabetragting van 'n nie-moedertaalonderwys-ervaring. *Tydskrif vir Geesteswetenskappe*, *57*(3), 693-705. http://dx.doi.org/10.17159/2224-7912/2017/v57n3a2

Evans, R., & Cleghorn, A. (2014). Parental perceptions: a case study of school choice amidst language waves. *South African Journal of Education*, *34*(2), Art. # 819, 19 pages.

Feilzer, Y. M. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, *4*(1), 6-16. https://doi.org/10.1177/1558689809349691

Fengling, L. (2017). A comparative study of Nida and Newmark's translation theories. *International Journal of Liberal Arts and Social Science*, *5*(8), 31-39.

Ferreira-Meyers, K. A. F., & Horne, F. (2017). Multilingualism and the language curriculum in South Africa: Contextualising French within the local language ecology. *Stellenbosch Papers in Linguistics Plus*, *51*, 23-40. http://dx.doi.org/10.5842/51-0-696

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). SAGE Publications.

Fischer, J., Praetorius, A. K., & Klieme, E. (2018). The impact of linguistic similarity on cross-cultural comparability of students' perceptions of teaching quality. *Educational*

Assessment, Evaluation and Accountability, *31*, 201-220. https://doi.org/10.1007/s11092-019-09295-7

Fisher, W. (1992). Reliability Statistics. *Rasch Measurement Transactions*, (*6*)3, 238.

Fives, A., Russel, D., Kearns, N., Lyons, R., Eaton, P., Canavan, J., Devaney. C., & O'Brien, A. (2014). The association between academic self-beliefs and reading achievement among children at risk of reading failure. *Journal of Research in Reading*, *37*(2), 215-232.

Foley, A. (2002). South African education and the dilemmas of multilingualism. *English Academy Review*, *19*, 50-63.

Forrin, N. D., Risko, E. F., & Smilek, D. (2019). On the relation between reading difficulty and mind-wandering: A section-length account. Psychological research, 83(3), 485-497. https://doi.org/10.1007/s00426-017-0936-9

Foy, P. (2017). *PIRLS 2016 user guide for the international database*. TIMSS & PIRLS International Study Center. https://timssandpirls. bc.edu/pirls2016/international-database/downloads/P16UserGuide.pdf

Gaber, J. (2010). Face validity. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 471-474). SAGE Publications.

Gathercole, V. C. M. (Ed.) (2013). *Solutions for the assessment of bilinguals*. Multilingual Matters.

Gauton, R., & Schryver, G. (2004). *Translating technical texts into Zulu with the aid of multilingual and/or parallel corpora*. Unisa Press.

Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative of assessment instruments. *Psychological Assessment*, *6*(4), 53-38.

Gentzler, E. (2001). *Contemporary translation theories: Revised second edition*. Multilingual Matters.

Ghaith, G. M., & Harkouss, S. A. (2003). Role of text structure awareness in the recall of expository discourse. *Foreign Language Annals*, *36*, 86-96. http://dx.doi.org/10.1111/j.1944-9720.2003.tb01935.x

Gierl, M. J., & Khaliq, S. N. (2001). Identifying sources of differential item and bundle functioning on translated achievement tests: A confirmatory analysis. *Journal of Educational Measurement, 38*(2), 164-187.

Goertz, G., & Mahoney, J. (2012). Concepts and measurement: Ontology and epistemology. *Social Science Information*, *51*(2), 205-216. https://doi.org/10.1177/0539018412437108

Gordon, S. L., & Harvey, J. (2019) Choice of language in education: Do we know what South Africans want?. *Language and Education*, *33*(3), 226-243. https://doi.org/10.1080/09500782.2018.1488865

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text. In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82-98). Guilford Publications.

Graven, M. H. (2014). Poverty, inequality and mathematics performance: The case of South Africa's post-apartheid context. *ZDM–International Journal on Mathematics Education*, *46*(7), 1039-1049. https://doi.org/10.1007/s11858-013-0566-7

Greene, J. C., & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. *New Directions for Evaluation*, *74*, 5-17. https://doi.org/10.1002/ev.1068

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274.

Greenfield, P. M., Trumbull, E., Keller, H., Rothstein-Fisch, C., Suzuki, L., & Quiroz, B. (2006). Cultural conceptions of learning and development. In P. A. Alexander, P. R. Pintrich & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 675-694). Lawrence Erlbaum Publishers.

Hall, R. (2013). Mixed methods: in search of a paradigm. In T. Lê & Q. Lê (Eds.), Conducting Research in a Changing and Challenging World (p.71-78). Nova Science Publishers.

Hall, S. (1994). *Using Picture Storybooks to Teach Literary Devices: Recommended Books for Children and Young Adults*. Vol 2. Oryx Press.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Halliday, M. A. K., & Hasan, R. (1995). *Cohesion in English* (14th ed.). Longman

Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Lawrence Erlbaum Publishers.

Hammersley, M. (2011). Pragmatism. In M. S. Lewis-Beck, A. Bryman & T. Futing Liao (Eds.), *SAGE encylopedia of social science research methods* (pp. 848-849). SAGE Publications. https://dx.doi.org/10.4135/9781412950589.n739

Harkness, J. (2003). Questionnaire translation. In J. Harkness, F. J. R. van de Vijver & P. Ph. Mohler (Eds.), *Cross-Cultural Survey Methods* (pp. 35-56). Wiley.

Hasbrouck, J., & Tindal, G. (2017). *An update to compiled ORF norms* (Technical Report No. 1702). Behavioral Research and Teaching, University of Oregon.

Hathcoat, J. D., & Nicholas, M. C. (2014). Epistemology in D. Coghlan, & M. Brydon-Miller, *The SAGE encyclopedia of action research* (pp. 303-306). SAGE Publications.

Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. (2016). The effects of text structure instruction on expository reading comprehension: A meta-analysis. *Journal of Educational Psychology, 108*(5), 609-629. http://dx.doi.org/10.1037/edu0000082

Hebert, M., Bohaty, J. J., Nelson, J. R., & Roehling, J. V. (2018). Writing informational text using provided information and text structures: an intervention for upper elementary struggling writers. *Read Writ, 31,* 2165-2190. https://doi.org/10.1007/s11145-018-9841-x

Hemphill, L., & Tivnan, T. (2008). The importance of early vocabulary for literacy achievement in high-poverty schools. *Journal of Education for Students Placed at Risk, 13*, 426-451. http://dx.doi.org/10.1080/10824660802427710

Henning, E., Mamiane, A., & Pheme, M. (2001). Entering the academy as "the other" – about writing competence and the bridge to the discourse community. *South African Journal of Higher Education, 151,* 109-128.

Heugh, K. (2013). Multilingual education policy in South Africa constrained by theoretical and historical disconnections. *Annual Review of Applied Linguistics, 33,* 215-237. https://doi.org/10.1017/S0267190513000135

Hilton, M. (2006). Measuring standards in primary English: Issues of validity and accountability with respect to PIRLS and national curriculum test scores. *British Educational Research Journal, 32*(6), 817-837. https://doi.org/10.1080/01411920600989453

Hlatshwayo, S. A. (2000). *Education and independence: Education in South Africa, 1658-1988*. Greenwood Press.

Hooper, M., Mullis, I. V. S., & Martin, M. O. (2015). PIRLS 2016 context questionnaire framework. In I. V. S. Mullis & M. O. Martin (Eds.), *PIRLS 2016 assessment framework* (2nd ed) (pp.31-51). TIMSS & PIRLS International Study Center.

Howie, S. J., Combrinck, C., & Roux, K. (2017). South Africa encyclopedia. In I. V. S. Mullis, M. O. Martin, S. Goh & C. Prendergast (Eds.), *PIRLS 2016 encyclopedia: Education policy and curriculum in reading* (pp.1-15). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/pirls2016/international-results/encyclopedia/countries/south-africa/

Howie, S. J., Combrinck, C., Roux, K., Tshele, M., Mokoena, G. M., & McLeod Palane, N. (2017a). *PIRLS Literacy 2016 Progress in International Reading Literacy Study 2016: South African children's reading literacy achievement*. Centre for Evaluation and Assessment.

Howie, S. J., Combrinck, C., Tshele, M., Roux, K., McLeod Palane, N. & Mokoena, G. M. (2017b). *PIRLS 2016 Progress in International Reading Literacy Study 2016 grade 5 benchmark participation: South African children's reading literacy achievement*. Centre for Evaluation and Assessment.

Howie, S. J., van Staden, S., Tshele, M., Dowse, C., & Zimmerman, L. (2012). *PIRLS 2011: South African children's reading literacy achievement*. Centre for Evaluation and Assessment.

Howie, S. J., Venter, E., van Staden, S., Zimmerman, L., Long, C., du Toit, C., Scherman, V., & Archer, E. (2008). PIRLS 2006 summary report: South African children's reading literacy achievement. Centre for Evaluation and Assessment.

Hsieh, H., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288.

International Association for the Evaluation of Educational Achievement (IEA). (2011). *IEA studies*. Retrieved 29 July 2017 from http://www.iea.nl/studies.html

International Association for the Evaluation of Educational Achievement (IEA). (2018). *Brief history of the IEA*. Retrieved 4 April 2018 from http://www.iea.nl/brief-history-iea

International Association for the Evaluation of Educational Achievement (IEA). (n.d.-a). *The pearl passage map*. TIMSS & PIRLS International Study Center.

International Association for the Evaluation of Educational Achievement (IEA). (n.d.-b). *Flowers on the roof passage map*. TIMSS & PIRLS International Study Center.

International Association for the Evaluation of Educational Achievement (IEA). (n.d.-c). *Macy and the red hen passage map*. TIMSS & PIRLS International Study Center.

International Association for the Evaluation of Educational Achievement (IEA). (n.d.-d). *African rhinos and oxpecker birds passage map*. TIMSS & PIRLS International Study Center.

International Association for the Evaluation of Educational Achievement (IEA). (n.d.-e). *The green sea turtle's journey of a lifetime passage map*. TIMSS & PIRLS International Study Center.

International Test Commission (ITC). (2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). Available from www.InTestCom.org

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods*, *18*(1), 3-20. https://doi.org/10.1177/1525822X05282260

Jakobson, R. (1959). On linguistics aspects of translation. In R. A. Brower (Ed.), *On translation* (pp.113-118). Harvard University Press.

Jansen, J. D. (1998). Curriculum reform in South Africa: a critical analysis of outcomes-based education. *Cambridge Journal of Education*, *28*(3), 321-331. https://doi.org/10.1080/0305764980280305

Jansen, J. D. (2002). Political symbolism as policy craft: Explaining non-reform in South African education after apartheid. *Journal of Education Policy*, *17*(2), 199-215. https://doi.org/10.1080/02680930110116534

Jindal, P., & MacDermid, J. C. (2017). Assessing reading levels of health information: Uses and limitations of Flesch formula. *Education for Health*, *30*(1), 84-88. https://doi.org/10.4103/1357-6283.210517

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: a research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26. https://doi.org/10.3102/0013189X033007014

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112-133. https://doi.org/10.1177/1558689806298224

Johnston, M. P. (2014). Secondary data analysis: a method of which the time has come. *Qualitative and Quantitative Methods in Libraries* (QQML), *3*, 619-626.

Joubert, H. J., & Bray, E. (2007). Public school governance in South Africa. Interuniversity Centre for Education Law and Education Policy.

Joubert, R., & Prinsloo, S. (2001). *Education law: A practical guide for educators*. Van Schaik Publishers.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1-73. https://doi.org/10.1111/jedm.12000

Kanjee, A., & Sayed, Y. (2013). Assessment policy in post-apartheid South Africa: challenges for improving education quality and learning. *Assessment in Education: Principles, Policy & Practice*, *20*(4), 442-469. https://doi.org/10.1080/0969594X.2013.838541

Keet, C. M., & Khumalo, L. (2017a). Toward a knowledge-to-text controlled natural language of isiZulu. *Lang Resources & Evaluation*, *51*, 131-157. http://dx.doi.org/10.1007/s10579-016-9340-0

Keet, C. M., & Khumalo, L. (2017b). Grammar rules for the isiZulu complex verb. *Southern African Linguistics and Applied Language Studies*, *35*(2), 183-200. https://doi.org/10.2989/16073614.2017.1358097

Kiecolt, K. J., & Nathan, L. E. (1985). *Quantitative Applications in the Social Sciences: Secondary analysis of survey data*. SAGE Publications. https://dx.doi.org/10.4135/9781412985796

Kim, M., Han, H.-R., & Phillips, L. (2003). Metric equivalence assessment in cross-cultural research: Using an example of the Center for Epidemiological Studies - depression scale. *Journal of Nursing Measurement*, *11*, 5-18.

Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In R. B. Ruddell & U. J. Norman (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1270-1327). International Reading Association.

Kintsch, W. (2013). Revisiting the construction-integration model of text comprehension and its implications for Instruction. In D. Alvermann, N. Unrau & R. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 807-841). International Reading Association. https://doi.org/10.4324/9781315110592

Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. Paris & S. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 71-92). Lawrence Erlbaum Associates.

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *Blackwell handbooks of developmental psychology. The science of reading: A handbook* (p. 209-226). Blackwell Publishing.

Koch, E. (2009). The case for bilingual language tests: a study of test adaptation and analysis. *Southern African Linguistics and Applied Language Studies*, *27*(3), 301-317. https://doi.org/10.2989/SALALS.2009.27.3.8.942

Korstjens, I., & Moser, A. (2018) Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice*, *24*(1), 120-124. https://doi.org/10.1080/13814788.2017.1375092

Krathwohl, D. R. (2002) A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212-218.

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology* (3rd ed.). SAGE Publications.

Kros, C. (2010). *The seeds of separate development: The origins of Bantu Education*. Unisa Press.

Kuhn, T. S. (1962). *The structure of scientific revolutions* (2nd ed.). The University of Chicago Press, Ltd.

Labuschagne, M. J. (2014). *Validation of the Performance of Tshivenda Learners in PIRLS 2006* [Master's Dissertation, University of Pretoria]. UP Space. http://hdl.handle.net/2263/46165

Land, S. (2015). Reading and the orthography of isiZulu. *South African Journal of African Languages*, *35*(2), 163–175. http://dx.doi.org/10.1080/02572117.2015.111300

LaRoche, S., Joncas, M., & Foy, P. (2017). Sample design in PIRLS 2016. In M. O. Martin, I. V. S. Mullis & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016* (pp. 3.1-3.34). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-3.html

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Multilingual Matters.

Laufer, B. (2013). Lexical thresholds for reading comprehension: What they are and how they can be used for teaching purposes. *TESOL Quarterly*, *47*(4), 867-872. https://doi.org/10.1002/tesq.140

Laufer, B., & Ravenhorst–Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, *22*, 15-30.

Leu, D., Mcverry, J.G., O'Byrne, W.I., Zawilinski, L., Castek, J., & Hartman, D.K. (2009). The new literacies of online reading comprehension and the irony of No Child Left Behind. In L. M. Morrow, R. Rueda & D. Lapp (Eds.), *Handbook of research on literacy and diversity* (pp. 173-194). The Guilford Press.

Lexico. (2020a). Ellipsis. In *Lexico.com*. Retrieved April 12, 2020, from https://www.lexico.com/en/definition/ellipsis

Lexico. (2020b). Trek. In *Lexico.com*. Retrieved June 13, 2020, from https://www.lexico.com/en/definition/trek

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean. *Rasch Measurement Transactions*, *16*(2), 878.

Linacre, J. M. (2012). *Winsteps Rasch Tutorial 3*. Retrieved from https://www.winsteps.com/a/winsteps-tutorial-3.pdf

Linacre, J. M. (2013) Disconnected Subsets, Guttman Patterns and Data Connectivity. *Rasch Measurement Transactions*, *27*(2), 1415-1417.

Linacre, J. M. (2016). *Winsteps® Rasch measurement computer program User's Guide*. Winsteps.com.

Linacre, J. M. (n.d.). *Summaries of persons and items*. Retrieved from https://www.winsteps.com/winman/table3_1.htm

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. SAGE Publications.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contra-dictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The handbook of qualitative research* (2nd ed., pp. 163-188). SAGE Publications.

Literary Devices Editors. (2019). *Suspense*. Retrieved December 17, 2019, from https://literarydevices.net/suspense/

Long, J. (2013). Translation definitions in different paradigms. *Canadian Social Science*, *9*(4), 107-115. http://dx.doi.org/10.3968/j.css.1923669720130904.2703

Louwerse, M. M., & Graesser, A. C. (2007). Coherence in discourse. In P. Strazny (Ed.), *Encyclopedia of linguistics* (pp. 216-218). Fitzroy Dearborn

Madiba, M., & Mabiletja, M. (2008). An evaluation of the implementation of the new Language-in-Education Policy (LiEP) in selected secondary schools of Limpopo Province. *Language Matters*, *39*(2), 204-229. https://doi.org/10.1080/10228190802579601

Mahdavi, J. N., & Tensfeldt, L. (2013). Untangling reading comprehension strategy instruction: Assisting struggling readers in the primary grades. *Preventing School Failure: Alternative Education for Children and Youth*, *57*(2), 77-92. https://doi.org/10.1080/1045988X.2012.668576

Makalela, L. (2017). Language Policy and Education in Southern Africa. In T. L. McCarty, & S. May (Eds.). *Language Policy and Political Issues in Education* (3rd ed, pp. 1-11).

Springer. Retrieved from https://link.springer.com/content/pdf/10.1007%2F978-3-319-02320-5_38-1.pdf

Makalela, L., & Fakude, P. (2014). 'Barking' at texts in Sepedi oral reading fluency: Implications for edumetric interventions in African languages. *South African Journal of African Languages*, 34, 71-78. https://doi.org/10.1080/02572117.2014.896538

Malan, J. J. (2012). 'n Oorweging van die sosiopolitieke kragte wat inwerk op die betekenis en toepassing van die diskresionêre taalklousule van die Suid-Afrikaanse grondwet. *LitNet Akademies*, *9*(1), 55-79.

Manyike, T. V., & Lemmer, E. M. (2014). Research in Language Education in South Africa: Problems & Prospects. *Mediterranean Journal of Social Sciences*, *5*(8), 251-258. https://doi.org/10.5901/mjss.2014.v5n8p251

Markus, K. A., & Smith, K. M. (2010). Content validity. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 239-243). SAGE Publications. https://dx.doi.org/10.4135/9781412961288

Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2017). *Methods and procedures in PIRLS 2016*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/publications/pirls/2016-methods.html

Martinez, L. S. (2017). Validity, face and content. In M. Allen (Ed.), *The SAGE encyclopedia of communication research methods* (pp. 1823-1824). SAGE Publications.

Maruyama, G., & Ryan, C. S. (2014). *Research methods in social relations* (8th ed.). John Wiley & Sons, Inc.

Maskew Miller Longman. (2012). *Survival guide to the FET Curriculum and Assessment Policy Statements (CAPS)*. Platinum.

Masoke-Kadenge, E., & Kadenge, M. (2013). 'Declaration without implementation': An investigation into the progress made and challenges faced in implementing the Wits language policy. *Language Matters*, *44*(3), 33-50. https://doi.org/10.1080/ 10228195.2013.837949

Mason, J. (2006). Mixing methods in a qualitatively driven way. *Qualitative Research*, *6*(1), 9-25.

Masrai, A. (2019). Vocabulary and reading comprehension revisited: Evidence for high-, mid-, and low-frequency vocabulary knowledge. *SAGE Open*, 1-13. http://dx.doi.org/10.1177/2158244019845182

Matthews-López, J. L. (2003). *Best practices and technical issues in cross-lingual, cross-cultural assessments: an evaluation of a test adaptation* [Doctoral Thesis, Ohio

University]. OhioLINK. https://etd.ohiolink.edu/pg_10?::NO:10:P10_ETD_SUBID:58664

Maykut, P., & Morehouse, R. (1994). *Beginning qualitative research: A philosophic and practical guide*. The Falmer Press.

McLeod Palane, N. L. (2017). *The Effect of Language of Instruction and Contextual Factors on Higher-order Reading Comprehension Performance* [Doctoral Thesis, University of Pretoria]. UP Space. https://repository.up.ac.za/handle/2263/65461

McLeod, S. (2016, July 20). *Translator's job is to be invisible*. The Kingston Whig Standard. https://www.thewhig.com/2016/07/20/translators-job-is-to-be-invisible/wcm/7f876daa-7f62-368a-67f1-4b9e292b7b81

Menzel, K., Lapshinova-Koltunski, E., & Kunz, K. (2017). Cohesion and coherence in multilingual contexts. In K. Menzel, E. Lapshinova-Koltunski & K. Kunz (Eds.), *New perspectives on cohesion and coherence*, (pp. 1-11). Language Science Press.

Merriam-Webster. (2020). Homonym. In *Merriam-Webster.com dictionary*. Retrieved March 31, 2020, from https://www.merriam-webster.com/dictionary/homonym

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*(2), 5-11.

Messick, S. (1992). Validity of test interpretation and use. In M.C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed., pp. 1487-1495). MacMillan.

Messick, S. (1995). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, *23*(2), 13-23.

Miao, J. (2000). The limitations of 'equivalent effect'. *Perspectives: Studies in Translatology*, *8*(3), 197-205. https://doi.org/10.1080/0907676X.2000.9961388

Mikhchi, H. H. (2011). Standards of textuality: Rendering English and Persian texts based on a textual model. *Journal of Universal Language*, *12*(1), 47-74.

Miller, R.T. (2018). English orthography and reading. In J. I. Liontas, M. DelliCarpini, & TESOL International Association (Eds.) *The TESOL encyclopedia of English language teaching* (pp. 1-7). Wiley-Blackwell. https://doi.org/10.1002/9781118784235.eelt0461

Millin, T. (2015). Reading to learn: A literature review within a South African context. *Stellenbosch Papers in Linguistics*, *44*, 105-124. http://dx.doi.org/10.5774/44-0-170

Mohammed, I., & Amponsah, O. (2018). Predominant factors contributing to low reading abilities of pupils at Elsie Lund Basic School in the Tamale metropolis, Ghana. *African Educational Research Journal*, *6*(4), 273-278.

Mohlala, M. (1994). Opposition to unequal education: Towards understanding Black people's objection to apartheid education in South Africa. Amherst.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained. Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research, 1*(1), 48-76. https://doi.org/10.1177/2345678906292462

Morgan, D. L. (2014). *Integrating Qualitative and Quantitative Methods: A Pragmatic Approach*. SAGE Publications. https://dx.doi.org/10.4135/9781544304533

Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology, 52*(2), 250-260. https://doi.org/10.1037/0022-0167.52.2.250

Morse, J. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research, 40*(1), 120-123.

Moss, P. A. (2007). Reconstructing validity. *Educational Researcher, 36*(8), 470-476.

Mostert, M. (2007). Face validity. In N. J. Salkind (Ed.), *Encyplopedia of measurement and statistics* (pp. 380-386). SAGE Publications.

Mtsatse, N. (2018). *Exploring differential item functioning on reading achievement between English and isiXhosa language subgroups* [Master's Thesis, University of Pretoria]. UP Space. https://repository.up.ac.za/handle/2263/65447

Mullis, I. V. S, & Martin, M. O. (Eds.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). TIMSS & PIRLS International Study Center.

Mullis, I. V. S., & Prendergast, C. O. (2017). Developing the PIRLS 2016 achievement Items. In M. O. Martin, I. V. S. Mullis & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016*. TIMSS & PIRLS International Study Center.

Mullis, I. V. S., Martin, M. O., & Sainsbury, M. (2015). PIRLS 2016 reading framework. In I. V. S. Mullis & M. O. Martin, *PIRLS 2016 assessment framework* (2nd ed., pp 11-30). TIMSS & PIRLS International Study Center.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 international results in reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/pirls2016/international-results/

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 international report*. TIMSS & PIRLS International Study Center.

Murray, S. (2002). Language issues in South African education: An overview. In R. Mesthrie (Ed.), *Language in South Africa* (pp. 434-448). Cambridge University Press.

Najafi, E., Valizadeh, A., & Darooneh, A. H. (2020): The effect of translation on text coherence: A quantitative study. *Journal of Quantitative Linguistics*, 1-15. https://doi.org/10.1080/09296174.2020.1774297

Netten, A., Droop, M. & Verhoeven, L. (2011). Predictors of reading literacy for first and second language learners. *Read Writ, 24***,** 413–425. https://doi.org/10.1007/s11145-010-9234-2

Newton, P. E., & Shaw, S. D. (2014). *Validity in educational & psychological assessment*. SAGE Publications.

Ngcobo, M. (2013). Loan words classification in isiZulu: The need for a sociolinguistic approach, *Language Matters: Studies in the Languages of Africa*, *44*(1), 21-38. http://dx.doi.org/10.1080/10228195.2012.706317

Nguyen, L. T. C., & Nation, P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. *RELC Journal*, *42*, 86-99. https://doi.org/10.1177/0033688210390264

Ngwenya, T. (2012). The North-West University language policy: A glimmer of hope and flashes of red lights. *Language Matters*, *43*(2), 221-239. https://doi.org/10.1080/10228195.2011.637574

Nida, E. A. (1964). *Towards a Science of Translating*. Brill.

Nida, E. A. (1993). *Language, culture, and translating*. Shanghai Foreign Language Education Press.

Nida, E. A., & Taber, C. R. (1969). *The Theory and Practice of Translation* (4th ed.). Brill.

Ntshangase-Mtolo, P. (2009). *The translatability of English academic discourse into isiZulu with reference to the discourse of mathematics*. [Master's Dissertation, University of KwaZulu-Natal]. UKZN Space. http://ukzn-dspace.ukzn.ac.za/handle/10413/1010

Nwammuo, A. N., & Salawu, A. (2018) Media roles in disseminating strategies for teaching and learning indigenous languages: The case of South Africa's language-in-education policy in post-apartheid era. *Cogent Arts & Humanities*, *5*(1), 1553653. https://doi.org/10.1080/23311983.2018.1553653

O'Cathain, A. (2010). Assessing the quality of mixed methods research: toward a comprehensive framework. In A. Tashakkori, & C. Teddlie. *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 531-556). SAGE Publications. https://dx.doi.org/10.4135/9781506335193

Ogbonnaya, U. I., & Awuah, F. K. (2019). Quintile Ranking of Schools in South Africa and Learners' Achievement in Probability. *Statistics Education Research Journal*, *18*(1), 106-119.

Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, *12*(2), 281-316. Retrieved from https://nsuworks.nova.edu/tqr/vol12/iss2/

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, *13*(1), 48-63.

Onwuegbuzie, A. J., & Leech, N. L. (2004). Enhancing the interpretation of significant findings: The role of mixed methods research. *The Quantitative Report*, *9*(4), 770-792.

Onwuegbuzie, A. J., Johnson, R. B., & Collins, K. M. T. (2009). Call for mixed analysis: A philosophical framework for combining qualitative and quantitative approaches. *International journal of multiple research approaches*, *3*(2), 114-139. https://doi.org/10.5172/mra.3.2.114

Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.). *Handbook of Mixed Methods in Social and Behavioural Research* (pp. 351-383). SAGE publications.

Östlund, U., Kidd, L., Wengström, Y., & Rowa-Dewar, N. (2011). Combining qualitative and quantitative research within mixed method research designs: A methodological review. *International Journal of Nursing Studies*, *48*, 369-383. https://doi.org/10.1016/j.ijnurstu.2010.10.005

Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*, *46*, 1-18.

Payne, G. & Payne, J. (2004). Secondary analysis. In Payne, G., & Payne, J. *Sage key Concepts: Key concepts in social research* (pp. 214-218). SAGE Publications. https://dx.doi.org/10.4135/9781849209397

Payne, G., & Williams, M. (2011). *Teaching quantitative methods: Getting the basics right*. SAGE Publications. https://dx.doi.org/10.4135/9781446268384

Peña, E. D. (2001). Assessment of semantic knowledge: Use of feedback and clinical interviewing. *Seminars in Speech and Language*, *22*, 51-63.

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, *78*(4), 1255-1264. https://www.jstor.org/stable/4620701

Pentimonti, J. M., Zucker, T. A., Justice, L. M., & Kaderavek, J. N. (2010). Informational text use in preschool classroom read-alouds. *The Reading Teacher*, *63*(8), 656–665. http://dx.doi.org/10.1598/RT.63.8.4

Petöfi, J. S. (1973). Towards an empirically motivated grammatical theory of verbal texts. In J. S. Petöfi & H. Rieser (Eds.) *Studies in text grammar. Foundations of language (Supplementary Series)*, V.19 (pp. 205-275). Springer. https://doi.org/10.1007/978-94-010-2636-9_6

Pietersen, J. & Maree, K. (2016). Statistical analysis I: Descriptive statistics. In K. Maree (Ed.), *First steps in research* (2nd ed., pp. 204-217). Van Schaik Publishers.

Pinnock, A. J. E. (2011). *A practical guide to implementing CAPS: A toolkit for teachers, schools managers and education officials to use to assist in managing the implementation of a new curriculum*. NAPTOSA.

Pinto, R. M. (2010). Mixed methods design. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 813-818). SAGE Publications. https://dx.doi.org/10.4135/9781412961288

Plano Clark, V., & Ivankova, N. (2016). *Mixed methods research: A guide to the field*. SAGE Publications. https://dx.doi.org/10.4135/9781483398341

Plüddermann, P. (2015). Unlocking the grid: Language-in-Education Policy realisation in postapartheid South Africa. *Language in Education*, *29*(3), 186-199. https://doi.org/10.1080/09500782.2014.994523

Pretorius, E. J. (2008). What happens to literacy in (print) poor environments? Reading in African Languages and School Language Policies. In M. Lafon & V. Webb (Eds.), *The Standardisation of African Languages: Language Political Realities* (pp. 60-87). Institut Français d'Afrique du Sud Johannesburg.

Pretorius, E. J. (2000). *Inference generation in the reading of expository texts by university students*. [Doctoral Thesis, University of South Africa]. UnisaIR. http://uir.unisa.ac.za/handle/10500/15816

Pretorius, E. J. (2014). Supporting transition or playing catch-up in grade 4? Implications for standards in education and training. *Perspectives in Education*, *32*(1), 51-76.

Pretorius, E. J. (2016, 30 September – 2 October). *Narrowing the gap: Building vocabulary during the foundation phase* [Paper presentation]. RASA 2016 Conference, KwaZulu-Natal, South Africa. https://www.zenexfoundation.org.za/images/Narrowing_the_gap_-_Building_Vocabulary_During_the_Foundation_Phase.pdf

Pretorius, E. J., & Klapwijk, N. M. (2016). Reading comprehension in South African schools: Are teachers getting it, and getting it right? *Per Linguam*, *32*(1), 1-20. http://dx.doi.org/10.5785/32-1-627

Pretorius, E., & Stoffelsma, L. (2017). How is their word knowledge growing? Exploring grade 3 vocabulary in South African township schools. *South African Journal of Childhood Education*, *7*(1), 13 pages. https://doi.org/10.4102/sajce.v7i1.553

Probyn, M. (2009). 'Smuggling the vernacular into the classroom': Conflicts and tensions in classroom codeswitching in township/rural schools in South Africa. *International Journal of Bilingual Education and Bilingualism*, *4*(4), 249-266. https://doi.org/10.1080/13670050802153137

Randall, J. & Engelhard, G. (2010). Using confirmatory factor analysis and the Rasch model to assess measurement invariance in a high stakes reading assessment. *Applied Measurement in Education*, *23*(3), 286-306. https://doi.org/10.1080/08957347.2010.486289

Rashed, M. A. (2013). Culture, salience, and psychiatric diagnosis: Exploring the concept of cultural congruence & its practical application. *Philosophy, ethics, and humanities in medicine: PEHM*, *8*(5), 1-12. https://doi.org/10.1186/1747-5341-8-5

Ray, M. N., & Meyer, B. J. F. (2011). Individual differences in children's knowledge of expository text structures: A review of literature. *International Electronic Journal of Elementary Education*, *4*(1), 67-82.

Reddy, V. (2006). *Mathematics and science achievement at South African schools in TIMSS 2003*. Human Sciences Research Council.

Republic of South Africa (RSA). (1953). *Bantu Education Act, Act No. 47 of 1953*. Government Printing.

Republic of South Africa (RSA). (1996a). *The Bill of Rights of the Constitution of the Republic of South Africa*. Government Gazette (No. 17678).

Republic of South Africa (RSA). (1996b). *The Constitution of the Republic of South Africa*. http://www.justice.gov.za/legislation/constitution/SAConstitution-web-eng.pdf

Republic of South Africa (RSA). (1996c). *National Education Policy Act 27 of 1996* (NEPA). https://www.education.gov.za/LinkClick.aspx?fileticket=l73mPb_ja4c%3D&tabid=419&portalid=0&mid=4023

Republic of South Africa (RSA). (1996d). *South African Schools Act 84 of 1996* (SASA). https://www.gov.za/sites/default/files/gcis_document/201409/act84of1996.pdf

Richards, J. C., & Schmidt, R. (2010). Text. In *Longman dictionary of language teaching and applied linguistics* (4th ed.). Pearson Education Limited.

Rios, J. A., & Sireci, S. G. (2014) Guidelines versus practices in cross-lingual assessment: A disconcerting disconnect. *International Journal of Testing*, *14*(4), 289-312. https://doi.org/10.1080/15305058.2014.924006

Roberts, K. L., Norman, R. R., & Cocco, J. (2015). Relationship between graphical device comprehension and overall text comprehension for third-grade children. *Reading Psychology*, *36*(5), 389-420. https://doi-org.uplib.idm.oclc.org/10.1080/02702711. 2013.865693

Roberts, K. L., Norman, R. R., Duke, N. K., Morsink, P., Martin, N. M., & Knight, J. A. (2013). Diagrams, timelines, and tables - Oh, My! Fostering graphical literacy. *Reading Teacher*, *67*(1), 12-24. https://doi-org.uplib.idm.oclc.org/10.1002/TRTR.1174

Robertson, S.-A., & Graven, M. (2020). A mathematics teacher's response to a dilemma: 'I'm supposed to teach them in English but they don't understand'. *South African Journal of Childhood Education*, *10*(1), a800. https://doi.org/10.4102/sajce.v10i1.800

Roehling, J. V., Hebert, M., Nelson, J. R., & Bohaty, J. J. (2017). Text structure strategies for improving expository reading comprehension. *The Reading Teacher*, *71*(1), 71-82. https://doi.org/10.1002/trtr.1590

Rogler, L. H. (1999). Methodological sources of cultural insensitivity in mental health research. *American Psychologist*, *54*, 424-433.

Rothon, C., Head, J., Klineberg, E., & Stansfeld, S. (2011). Can social support protect bullied adolescents from adverse outcomes? A prospective study on the effects of bullying on the educational achievement and mental health of adolescents at secondary schools in East London. *Journal of Adolescence*, *34*(3), 579–588.

Roux, K. (2014). *Identifying the relationship between the home environment, parental attributes and learner achievement in reading literacy* [Master's Thesis, University of Pretoria]. UP Space. https://repository.up.ac.za/handle/2263/66516

Rueda, R. (2011). Cultural perspectives in reading: Theory and research. In M. Kamil, P.D. Pearson, E.B. Moje & P. Afflerbach (Eds.), *Handbook of reading research IV* (pp. 84-104). Routledge.

Sabatini, J. P., Albro, E., O'Reily, T. (Eds.). (2012). *Measuring up: Advances in how we assess reading ability*. Rowman & Littlefield Publishers, Inc.

Sainsbury, M., & Schagen, I. (2004). Attitudes to reading at ages nine and eleven. *Journal of Research in Reading*, 27: 373-386. https://doi.org/10.1111/j.1467-9817.2004.00240.x

Sandilands, D., Oliveri, M. E., Zumbo, B. D., & Ercikan, K. (2013). Investigating sources of differential item functioning in international large-scale assessments using a confirmatory approach. *International Journal of Testing*, *13*(2), 152-174. https://doi.org/10.1080/15305058.2012.690140

Scheffler, I. (2011). *Four pragmatists: A critical introduction to Peirce, James, Mead and Dewey*. Routledge.

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, *47*(4): 484-503. https://doi.org/10.1017/S0261444812000018

Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2017). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching*, *50*(2), 212-226. http://dx.doi.org/10.1017/S0261444815000075

Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 189-213). SAGE Publications.

Sechrest, L., Fay, T. L., & Hafeez Zaidi, S. M. (1972). Problems of translation in cross-cultural research. *Journal of Cross-Cultural Psychology*, *3*, 41-56.

Shepard, L. A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practices*, *16*(2), 5-8. https://doi.org/10.1111/j.1745-3992.1997.tb00585.x

Sibanda, J., & Baxen, J. (2018). Third-grade English second language teachers' vocabulary development practices. *South African Journal of Childhood Education*, 8(1), a545. https://doi.org/10.4102/sajce.V8i1.545

Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(10), 2621-2624. https://doi.org/10.1073/pnas.1516947113

Smith, E. V. & Smith, R. M. (Eds.) (2004). *Introduction to Rasch measurement: Theory, models and applications*. JAM Press.

Snow, C. E., Burns, S. M., & Griffin, P. (Eds). (1998). *Preventing reading difficulties in young children*. National Academy Press. Online available at https://www.nap.edu/read/6023/chapter/1

South African Government. (2019). *President Cyril Ramaphosa: State of the nation address 2019*. Online available at https://www.gov.za/speeches/2SONA2019.

Spaull, N. & Pretorius, E. (2019). Still falling at the first hurdle: Examining early grade reading in South Africa. In N. Spaull & J. Jansen (Eds), *South African schooling: The enigma of inequality* (pp. 147-168). Springer Nature.

Spaull, N. (2015a). Schooling in South Africa: How Low Quality Education Becomes a Poverty Trap. In A. De Lannoy, S. Swartz, L. Lake, C. & Smith (Eds), *South African Child Gauge 2015* (pp. 34-41). University of Cape Town.

Spaull, N. (2015b). *Education quality in South Africa and sub-Saharan Africa: An economic approach* [Doctoral Thesis, Stellenbosch University]. SUNScholar Research Repository. https://scholar.sun.ac.za/handle/10019.1/96775

Spaull, N. (2016, 30 September - 2 October). *Learning to Read and Reading to Learn*. [Paper presentation]. RASA 2016 Conference, KwaZulu-Natal, South Africa. https://www.zenexfoundation.org.za/images/Laerning_to_read_and_reading_to_learn_-_Nic_Spaull.pdf

Spaull, N., & Kotze, J. (2015). Starting behind and staying behind in South Africa: The case of insurmountable learning deficits in mathematics. *International Journal of Educational Development*, *41*, 13-24. https://doi.org/10.1016/j.ijedudev.2015.01.002

Spaull, N., Pretorius, E., & Mohohlwane, N. (2020). Investigating the comprehension iceberg: Developing empirical benchmarks for early-grade reading in agglutinating African languages. *South African Journal of Childhood Education*, *10*(1), a773. https://doi.org/10.4102/sajce.v10i1.773

Statistics South Africa (StatsSA). (2012). *Census 2011 in Brief*. Retrieved March 19, 2017 from http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf

Stoffelsma, L. (2019). English vocabulary exposure in South African township schools: Pitfalls and opportunities. *Reading & Writing*, *10*(1), a209. https://doi.org/.4102/rw.v10i1.209

Stubbe, T. C. (2011). How do different versions of a test instrument function in a single language? A DIF analysis of the PIRLS 2006 German assessments. *Educational Research and Evaluation*, *17*(6), 465-481. https://doi.org/10.1080/13803611.2011.630560

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Pearson.

Tamayo, J. (1987). Frequency of use as a measure of word difficulty in bilingual vocabulary test construction and translation. *Educational & Psychological Measurement*, *47*, 893-902.

Tashakkori, A., & Teddlie, C. (2010). *Handbook of mixed methods in social and behavioral research* (3rd ed.). SAGE Publications.

Tashakkori, A., & Teddlie, C. (Eds). (2003). *Handbook of mixed methods in social and behavioral research*. SAGE Publications.

Taylor, C. S. (2013). *Validity and validation* (Series in understanding statistics). Oxford University Press.

Taylor, S., & von Fintel, M. (2016). Estimating the impact of language of instruction in South African primary schools: A fixed effects approach. *Economics of Education Review*, *50*, 75-89. https://doi.org/10.1016/j.econedurev.2016.01.003

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research. Integrating quantitative and qualitative approaches in the social and behavioral sciences*. SAGE Publications.

Thomson, D. R., Besner, D., & Smilek, D. (2013). In pursuit of off-task thought: Mind wandering-performance trade-offs while reading aloud and color naming. *Frontiers in Psychology*, 4, A360. https://doi.org/10.3389/fpsyg.2013.00360

Tolley, E. (2016). *Qualitative methods in public health: A field guide for applied research* (2nd ed.). Wiley.

Trochim, W. M. K., & Donnelly, J. P. (2006). *The research methods knowledge base* (3rd ed.). Atomic Dog.

Truddell, B., & Schroeder, L. (2007). Reading methodologies for African languages: Avoiding linguistic and pedagogic imperialism. *Language, Culture and Curriculum*, *20*(3), 165-180. http://dx.doi.org/10.2167/lcc333.0

Turner, N., & Wildsmith-Cromarty, R. (2014) Challenges to the implementation of bilingual/multilingual language policies at tertiary institutions in South Africa (1995-2012). *Language Matters*, *45*(3), 295-312. https://doi.org/0.1080/10228195.2014.961525

United National Educational, Scientific and Cultural Organization (UNESCO). (1953). *The use of vernacular languages in education*. Paris: UNESCO.

United National Educational, Scientific and Cultural Organization (UNESCO). (2012). *International standard classification of education: ISCED 2011*. UNESCO. http://uis.unesco.org/en/topic/international-standard-classification-education-isced

United National Educational, Scientific and Cultural Organization (UNESCO). (2017). *International study concludes reading literacy on the rise*. UNESCO. https://en.unesco.org/news/international-study-concludes-reading-literacy-rise

Usunier, J. C. (2011). Language as a resource to assess cross-cultural equivalence in quantitative management research. *Journal of World Business*, *46*(3), 314-319. https://doi.org/10.1016/j.jwb.2010.07.002

Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme development in qualitative content analysis and thematic analysis. *Journal of Nursing Education and Practice*, *6*(5), 100-110.

van de Vijver, F. J. R. & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. SAGE Publications.

van den Broek, P., & Kendeou, P. (2017). Development of reading comprehension: Change and continuity in the ability to construct coherent representations. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.). *Theories of reading development*: *Studies in written language and literacy* (Vol. 15) (pp. 283-306). John Benjamins Publishing Company. Retrieved from https://ebookcentral-proquest-com.uplib.idm.oclc.org

van der Berg, S. (2008). How effective are poor schools? Poverty and educational outcomes in South Africa. *Studies in Educational Evaluation*, *34*(3), 145-154. https://doi.org/10.1016/j.stueduc.2008.07.005

van der Berg, S., & Gustafsson, M. (2019). Educational outcomes in post-apartheid South Africa: Signs of progress despite great inequality. In N. Spaull, & J. Jansen (Eds). *South African schooling: The enigma of inequality. Policy Implications of Research in Education* (Vol. 10) (pp. 25-45). Springer. https://doi.org/10.1007/978-3-030-18811-5_2

van der Berg, S., Spaull, N., Wills, G., Gustafsson, M., & Kotzé, J. (2016). *Identifying binding constraints in education*. Research on Socio-economic Policy. http://resep.sun.ac.za/wp-content/uploads/2017/10/PSPPD_BICiE-email-01062016.pdf

van der Merwe, A., & Le Roux, M. (2014). Idiosyncratic sound systems of the South African Bantu languages: Research and clinical implications for speech-language pathologists and audiologists. *South African Journal of Communication Disorders*, *61*(1), Art. #86, 8 pages. http://dx.doi.org/10.4102/sajcd.v61i1.86

van der Veer, K., Ommundsen, R., Hak, T., & Larsen, K. S. (2003). Meaning shift of items in different language versions. A cross-national validation study of the illegal aliens scale. *Quality & Quantity International Journal of Methodology*, *37*, 193-206.

van der Walt, J. L., & Steyn, H. S. (2007). Pragmatic validation of a test of academic literacy at tertiary level. *Ensovoort*, *11*(2), 138-153.

van der Walt, J. L., & Steyn, H. S. (2008). The validation of language tests. *Stellenbosch Papers in Linguistics*, *38*, 191-204. https://doi.org/10.5774/38-0-29

van Dijk, T. A., & Kintch, W. (1983). *Strategies of Discourse Comprehension*. Academic Press.

van Dyk, H., & White, C. J. (2019). Theory and practice of the quintile ranking of schools in South Africa: A financial management perspective. *South African Journal of Education*, *39*(1), 1-9. https://doi.org/10.15700/saje.v39ns1a1820

van Rooy, B., & Pretorius, E. J. (2013). Is reading in an agglutinating language different from an analytic language? An analysis of isiZulu and English reading based on eye movements. *Southern African Linguistics & Applied Language Studies*, *31*(3), 281-297. http://dx.doi.org/10.2989/16073614.2013.837603

van Staden, S. (2018). Exploring possible effects of differential item functioning on reading achievement across language subgroups: A South African perspective. In L. D. Hill & F. J. Levine (Eds.), *Global Perspectives on Education* (pp. 9.1-9.22). Routledge. https://doi.org/10.4324/9781351128421

van Staden, S. (in press). Reading comprehension in early childhood: Putting new evidence of trend growth in South Africa into perspective. In K. S. Adeyemo (Ed.). *Education systems for Africa*. Springer.

van Staden, S., Bosker, R., & Bergbauer, A. (2016). Differences in achievement between home language and language of learning in South Africa: Evidence from prePIRLS 2011. *South African Journal of Childhood Education*, *6*(1), 1-10. https://dx.doi.org/10.4102/sajce.v6i1.441

van Staden, S., Combrinck, C., Roux, K., Tshele, M. & Palane, N. M. (2019). Moving beyond league table standings: How measures of opportunity to learn can inform educational quality and policy directives. *South African Journal of Childhood Education*, *9*(1), a712. https://doi.org/10.4102/sajce.v9i1.712

Vartanian, P. (2011). *Secondary Data Analysis*. Oxford.

Venuti, L. (2012). *The translation studies reader*. Routledge.

Vogt, W. P. (2005). *Dictionary of statistics & methodology* (3rd ed.). SAGE Publications. https://doi.org/10.4135/9781412983907

Vottonen, E. (2016). Does theory meet practice? The realisation of global and local strategies in self-reflection analysis. In M. Dore (Ed.), *Achieving Consilience: Translation Theories and Practice* (pp. 206-227). Cambridge Scholars Publishing.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures*. Rand McNally.

Webb, V. (2012). Managing multilingualism in higher education in post-1994 South Africa. *Language Matters*, *43*(2), 202-220. https://doi.org/10.1080/10228195.2011.618508

Weiner, I. B. (Ed.). (2003). *Handbook of psychology: Assessment psychology* (10th ed.). John Wiley & Sons, Inc.

Wharton-McDonald, R. & Swiger, S. (2009). Developing higher order comprehension in the middle grades. In S. Israel & G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 510-530). Routledge.

Whetton, C., Twist, L., & Sainsbury, M. (2007). Measuring standards in primary English: the validity of PIRLS - a response to Mary Hilton. *British Educational Research Journal*, *33*(6), 977-986. https://doi.org/10.1080/01411920701657082

Widdowson, H. (2007). *Discourse Analysis*. Oxford University Press.

Wilkes, A., & Nkosi, N. (2010). *Complete Zulu Beginner to Intermediate Book and Audio Course*. John Murray Press.

Wools, S. (2015). *All About Validity: An evaluation system for the quality of educational assessment*. Ipshamp Drukkers.

World Wildlife Fund. (2020). *Rhino Species*. Retrieved from https://www.worldwildlife.org/species/rhino

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement*, *14*, 97-116. https://www.jstor.org/stable/1434010

Wright, B. D., & Stone, G. E. (2004). *Making measures*. Phaneron Press.

Wright, B. D., & Stone, M. H. (1999). *Measurement essentials* (2nd ed.). Wide Range.

Zimmerman, L. (2010). *The influence of schooling conditions and teaching practices on curriculum implementation for Grade 4 learner reading literacy development*. [Doctoral Thesis, University of Pretoria]. UP Space. https://repository.up.ac.za/handle/2263/24982

Zimmerman, L. (2017). Learning from the best: Reading literacy development practices at a high-performing primary school. *Per Linguam*, *33*(2), 36-50. https://doi.org/10.5785/33-2-740