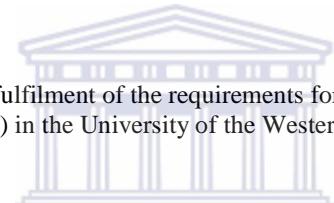


**UNIVERSITY OF THE WESTERN CAPE
Community and Health Sciences**

An evaluation of group differences and items bias, across rural isiXhosa learners and urban isiXhosa learners, of the isiXhosa version of the *Woodcock Muñoz Language Survey* (WMLS).

UNATHI LUCIA SILO



A minithesis submitted in partial fulfilment of the requirements for the degree of MA
(Research Psychology) in the University of the Western Cape

Supervisor: Professor E. Koch
WESTERN CAPE

January 2010

KEYWORDS

Differential item functioning (DIF), scalar equivalence, psychological testing, cross-cultural assessment, language in education, language policies, language proficiency, isiXhosa dialects, test translation, test adaptation.

ABSTRACT

In many countries defined by multilingualism, language has been identified as a great influence during psychological and educational testing. In South Africa (SA), factors such as changes in policies and social inequalities also influence testing. Literature supports the translation and adaptation of tests used in such contexts in order to avoid bias caused by language. Different language versions of tests then need to be evaluated for equivalence, to ensure that scores across the different language versions have the same meaning. Differences in dialects may also impact on the results of such tests. Results of an isiXhosa version of the *Woodcock Muñoz Language Survey* (WMLS), which is a test used to measure isiXhosa learners' language proficiency, show significant mean score differences on the test scores across rural and urban first-language speakers of isiXhosa. These results have indicated a possible problem regarding rural and urban dialects during testing. This thesis evaluates the item bias of the subtests in this version of the WMLS across rural and urban isiXhosa learners. This was accomplished by evaluating the reliability and item characteristics for group differences, and by evaluating differential item functioning across these two groups on the subtests of the WMLS. The sample in this thesis comprised of 260 isiXhosa learners from the Eastern Cape Province in grade 6 and grade 7, both males and females. This sample was collected in two phases: (1) secondary data from 49 rural and 133 urban isiXhosa learners was included in the sample; (2) adding to the secondary data, a primary data collection from 78 rural isiXhosa learners was made to equalise the two sample groups. All ethical considerations were included in this thesis. The results were surprising and unexpected. Two of the subtests in the WMLS showed evidence of scalar equivalence as only a few items were identified as problematic. However, two of the subtests demonstrated more problematic items. These results mean that two subtests of the WMLS that demonstrated evidence of scalar equivalence can be used to measure the construct of language proficiency, while the other two sub-tests that showed problematic items need to be further investigated, as the responses given by learners on these items seem to be determined by their group membership and not by their ability.

DECLARATION

I declare that *An evaluation of group differences and items bias across rural and urban isiXhosa learners, of the isiXhosa version of the Woodcock Muñoz Language Survey (WMLS)*, is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Full name.....

Date.....

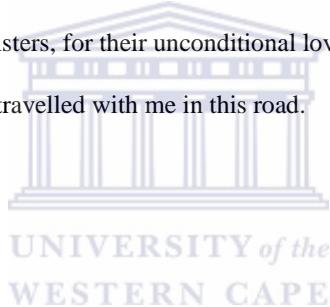
Signed.....



ACKNOWLEDGEMENTS

Sincere gratitude and appreciation to the following people, for their guidance, inspiration, direction and management, all of which enabled me to fulfil the requirements for the degree, MA (Research Psychology).

- Almighty God, for giving me strength during difficult times in my studies.
- My supervisor, Professor Koch, for guidance and for the opportunity to be involved in the ABLE project.
- NRF for financial support.
- My mother, father and two sisters, for their unconditional love and support.
- Friends and colleagues who travelled with me in this road.



CONTENTS	PAGE NO
ABSTRACT	ii
DECLARATION	iii
ACKNOLEDGEMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	ix
LIST OF FOMULAS	x
CHAPTER 1 INTRODUCTION	1
1.1. Background of the study	1
1.2. Background of language in the educational context of South Africa	2
1.3. Testing in the educational context of South Africa	3
1.4. Policies and institutions guarding against test bias	4
1.5. Assessing in cross-cultural contexts	4
1.6. Progress in the implementation of policies for testing practitioners	5
1.7. Dynamics of rural and urban isiXhosa dialects in language testing	5
1.8. Problem statement	6
1.9. Aim and specific objectives	7
1.9.1. Aim	7
1.9.2. Specific objectives	7
CHAPTER 2 LITERATURE REVIEW	9
2.1. Introduction	9
2.2. Background of psychological and educational testing, and language as an issue in testing in diverse societies.	10
2.2.1. Background of psychological testing	10
2.2.2. Language as an issue in cross-cultural testing	12
2.3. Standards and guidelines to address the issue of psychological and educational tests.	14
2.4. Psychological and educational tests in South Africa's multicultural society.	16

2.5.	Status of indigenous languages in South Africa, and as the context of assessment.	18
2.6.	IsiXhosa language varieties and their relevance in testing.	19
2.7.	Theoretical framework of equivalence and bias	20
2.8.	Research on equivalence and bias in tests in South Africa	23
CHAPTER 3 METHOD		26
3.1.	Introduction	26
3.2.	Research design	26
3.3.	Participants	26
3.3.1.	Description of participants	28
3.4.	Data collection procedures	29
3.5.	Instrument	32
3.6.	Ethical considerations	33
3.7.	Data analysis	33
3.7.1.	Specific objective one	33
3.7.2.	Specific objective two	34
3.7.3.	Specific objective three	34
3.7.4.	Specific objective four	35
CHAPTER 4 RESULTS		37
4.1.	Introduction	37
4.2.	Group differences on total mean scores across rural and urban isiXhosa learners.	37
4.3.	Group differences on indexes of reliability across the four subtests of the WMLS, between rural and urban isiXhosa learners.	39
4.4.	Group differences on item characteristics in the subtests of the WMLS, across rural and urban isiXhosa learners.	41
4.5.	Item bias/DIF	51
4.5.1.	Picture Vocabulary	51

4.5.2. Verbal Analogies	53
4.5.3. Letter-Word-Identification	55
4.5.4. Dictation	59
4.5.5. Summary of DIF results	63
CHAPTER 5 DISCUSSION AND CONCLUSION	66
5.1. Introduction	66
5.2. Summary and discussions	66
5.3. Limitations of the study	71
5.4. Conclusion	72
5.5. Recommendations	72
REFERENCES	73
APPENDICES	83
APPENDIX A (Description of instrument WMLS)	83
APPENDIX B (Information Sheet: English and IsiXhosa version)	84
APPENDIX C (Consent Form: English and IsiXhosa version)	90
APPENDIX D (Assent Form: English and IsiXhosa version)	92
APPENDIX E (Ethics Statement)	94
APPENDIX F (Language editing verification)	95

LIST OF TABLES

- Table 1: Number and percentages of participants
- Table 2: Description of both rural and urban isiXhosa learners in grades
- Table 3: Descriptive of sample by gender (rural and urban)
- Table 4: Mean Scores and Standard Deviations for rural and urban isiXhosa learners on each of the subscales of the WMLS
- Table 5: Hotelling's T² results: differences between two groups on all the sub-tests
- Table 6: Presents the indexes of Reliability (Cronbach's Alpha) of the sub-tests of the WMLS across rural and urban isiXhosa learners
- Table 7: Test of Equality Reliability for the WMLS subtests
- Table 8: Mean item difficulty values and mean standard deviations across rural and urban isiXhosa learners on the four subtests of the WMLS
- Table 9: Mean item discrimination on the four subtests of the WMLS across rural and urban isiXhosa learners
- Table 10: Item difficulty and item discrimination analysis per group: Picture Vocabulary
- Table 11: Item difficulty and item discrimination analysis per group: Verbal Analogies
- Table 12: Item difficulty and item discrimination analysis per group: Letter-Word-Identification
- Table 13: Item difficulty and item discrimination per group: Dictation
- Table 14: Stepwise Logistic regression DIF procedure: Picture Vocabulary
- Table 15: Direction of DIF: Picture Vocabulary
- Table 16: Stepwise Logistic regression DIF procedure: Verbal Analogies subtest
- Table 17: Direction for DIF: Verbal Analogies subtest
- Table 18: Stepwise Logistic regression DIF procedure: Letter-Word-Identification
- Table 19: Direction for DIF: Letter-Word-Recognition subtest
- Table 20: Stepwise Logistic regression DIF procedure: Dictation subtest

Table 21: Direction of DIF: Dictation subtest

Table 22: Summary results of the Logistic Regression method



LIST OF FIGURES

- Figure 1: Presents grade differences between rural and urban isiXhosa learners
- Figure 2: Presents gender differences between rural and urban isiXhosa learners
- Figure 3: Mean scores on the four sub-tests across rural and urban isiXhosa learners



LIST OF FORMULAS

Formula 1: DIF Model



CHAPTER ONE

INTRODUCTION

1.1. BACKGROUND OF STUDY

Cultural diversity has brought about the necessity to adapt, translate and develop tests which are suitable for use in different cultural contexts. Assuring the appropriateness of these tests for different cultural groups is of great importance, as this assists in making valid inferences related to individual differences. This thesis forms part of a project known as the Additive Bilingual Education Project (ABLE) (Koch, 2005), which is aimed at promoting additive bilingual education through the medium of both English and isiXhosa. This implies that learners who are isiXhosa mother-tongue speakers are taught in their mother tongue during the foundation phase of their education, while being taught English as a subject. They continue with isiXhosa as part-medium of instruction until grade 6, while English is introduced gradually as a medium of instruction, until all instruction takes place in English from grade 7. The idea of mother-tongue-based bilingual education for the early school years has been promoted by, amongst others, the former Minister of Education, Naledi Pandor (Webb, 2001). Literature has also shown that the use of mother-tongue education during the early developmental years promotes children's cognitive advancement and facilitates the learning of additional languages (Classen, Krynauw, Peterson, & Mathe, 2001b).

The *Woodcock Muñoz Language Survey* (WMLS) was selected for the ABLE project as one of the tests that are used to evaluate this project by assessing learners' academic language proficiency in English and isiXhosa. The original English form of this test was imported from America, and for this project, the English version was adapted by the researchers (involved in the ABLE) for use in the South African context and was then translated into isiXhosa, following International Test Commission (ITC) guidelines (Hambleton, 1994). Detailed information about the adaptation process of this test in both versions is indicated in the following paper "The case of bilingual tests: a study of test adaptation and analysis" by (Koch, 2009). A second project was implemented from this project,

to assess the psychometric properties of both (English and Xhosa) versions of this test. After the adaptation of these two versions, a further step was taken to evaluate the equivalence of the English and the isiXhosa version. Based on these results, further adaptation of the isiXhosa version is being conducted and the equivalence of both these versions is still being investigated, an investigation that this thesis forms part of.

The specific focus of this thesis is evaluating the equivalence of the isiXhosa version of this test across rural and urban isiXhosa learners, using procedures explained in Chapter 3. The investigations in this thesis follow previous research conducted on this test in the second project mentioned above, which indicated mean differences between rural and urban isiXhosa learners on this instrument (Ntantiso, 2009). This chapter will introduce the dynamics of testing in cross-cultural contexts such as South Africa, focusing mainly on language in the educational context, test bias, and equivalence across groups. The rationale of this thesis and the outline of the other chapters will also be included in this chapter.

1.2. BACKGROUND OF LANGUAGE IN THE EDUCATIONAL CONTEXT OF SOUTH AFRICA

South Africa is a socially diverse society, and like other diverse societies around the world, its psychological testing procedures cannot be separated from the political, economic and social history of the country (Classen, 1997). An indication of this can be found in the past language policies and use of psychological tests in the country, which were directed by political aspirations, and have vastly influenced the use of language in the present-day education system and the performance of learners in psychological tests.

The education system of South Africa has undergone various policy changes over the years. One example of these changes is the introduction of the 1953 Bantu Education Act (Mungazi & Walker, 1997). This Act enforced English, and later Afrikaans, as the medium of educating Black learners, and was instrumental in changing the social identity and self-worth of these learners (Van Zyl, 1996).

However, before this Act mother-tongue education was also implemented as a language of teaching and learning in these schools for the first eight years of schooling (Van Zyl, 1996). Heugh (2002) notes that when the Bantu Education Act was introduced, a significant improvement in the matriculation results of learners who were speakers of African languages was identified. Heugh (2002) attributes this increase to the intensive mother-tongue education that was implemented before the Bantu Education Act, even though the conditions of learning were unfavourable. There were very few articles which demonstrated the success of mother-tongue education after the Bantu Education Act, however, because of the educational inequalities between Black and White schools which prevailed during this period (Heugh, 2002). After the 1976 uprising of schoolchildren against the use of Afrikaans as a medium of instruction in schools, the number of years that had to be spent learning through the medium of the African languages was reduced to only the first three years of schooling, on the insistence of the African people, mostly in reaction to the Bantu Education policies of the apartheid regime. Even though the success of the past implementation of mother-tongue education was not recorded, many studies have proven that early learning in one's mother-tongue increases academic language proficiency in this language and also makes it easier to learn an additional language (Webb, 2001; Manyike, 2007). The issue of past and present language policies in education has also significantly influenced the processes of testing learners in the educational context. Owing to the current emphasis on ethical practices in psychological testing, it has been stressed that there is a need to evaluate fairness in tests used in contexts which are marked by inequalities (Van de Vijver & Rothmann, 2004).

1.3. TESTING IN THE EDUCATIONAL CONTEXT OF SOUTH AFRICA

At present, the effects of the past inequalities are still experienced in schools, mainly by learners who are mother-tongue speakers of African languages (Mantsha, 2002). In most schools, the language of learning and teaching is still English or Afrikaans, regardless of the dominant language of the learners in a particular school (Mantsha, 2002). According to Mantsha (2002), this issue of

language in education also affects the assessment processes for learners who are mother-tongue speakers of African languages. In other words, learners have difficulty in processing the content of the assessment tools when these are presented in English, as this is not their mother tongue. Mantsha (2002) has noticed that these learners also have difficulty in understanding the content of assessment tools which are presented in their mother tongue, as they are not academically fluent in these languages. These are issues that need to be addressed in cross-cultural testing.

In addition, changes in the social environment, such as globalisation and urbanisation, need to be taken into account during assessments conducted in cross-cultural contexts (Heugh, Diedericks, Prinsloo, Herbst, & Winnaar, 2007). Such social changes have been found to result in a shift in language, and the development of new dialects. Therefore, before making inferences based on assessments in these contexts, assessment tools should be thoroughly evaluated for bias, and test-users need to be certain that the assessment tools used in such contexts are equivalent across the diverse spectrum of test-takers.

UNIVERSITY of the FREE STATE

1.4. POLICIES AND INSTITUTIONS GUARDING AGAINST TEST BIAS

The new Democratic Government has set means to restructure and transform the divided, fragmented, discriminatory and authoritarian education system into a more democratic, open, flexible and inclusive system (Sayed, 1998; Mauer, 2000). The Constitution and the Bill of Rights are evidence of these attempts, as both these legal documents aim at eliminating any discrimination by race, class, sex, language, etc. in South Africa (Department of Labour, 2009). Various other institutions, such as the Health Professions Council of South Africa (HPCSA) and Human Rights policies, are in place to guard against any unethical use of assessment tools or tools measuring human abilities, such as language tests. A similar pattern has been followed with the APA Standards and the guidelines for ethical testing in the United States of America (USA), which will be discussed in detail in Chapter 2. Ensuring the fairness of these assessment tools is vitally

important when using tools in a multicultural or multilingual society, especially if the results are to be compared across these cultures.

1.5. ASSESSING IN CROSS-CULTURAL CONTEXTS

Van de Vijver (cited in Koch, 2005) states that there has been an increase in assessing and testing across cultural and language groups around the world. He adds that this increase has been followed by a great demand for enhancing the testing of diverse groups (Koch, 2005). Mauer (2000) has suggested that assessment tools used across groups should encompass the basic psychometric properties for all groups, that is, reliability and validity, to ensure that all bias and discrimination in such tests is eliminated. These properties will be discussed in more detail in Chapter 2. In addition, policies are in place to guard against any unethical use of psychological tests.



1.6. PROGRESS IN THE IMPLEMENTATION OF TESTING POLICIES FOR TESTING PRACTITIONERS

Foxcroft and Roodt (2001) have established that even though there are institutions guarding against unethical practices in assessment, many of the assessment tools that are currently used (whether imported from abroad or local) are not been cross-culturally validated. The responsibility of ensuring that these assessment tools are bias-free and are equivalent across cultures has been placed on test users, publishers and test constructors (Taylor, 1987). Van de Vijver and Rothmann (2004) find that the problem in establishing and ensuring equity in cross-cultural assessment has not been adequately solved in the modern South Africa. Moreover, language is identified as being the core source of bias in assessment tools (Foxcroft & Aston, 2006).

1.7. DYNAMICS OF RURAL AND URBAN ISIXHOSA DIALECTS IN LANGUAGE TESTING

The focus of this thesis is on the isiXhosa version of the WMLS, which measures language proficiency. There are eleven official languages, nine of which are indigenous South African languages; isiXhosa is one of them. This language

group has the second largest number of speakers in South Africa. Calteaux (1996) has shown that there are eight identified dialects in isiXhosa, but only two of these dialects form part of the standard isiXhosa dialect which is used in official sectors such as the educational sector. According to Alexander (2005), it is power within a particular dialect which gives it the authority to become a standard language. Kerswell (2006) indicates that migration is a contributing factor to the development of new dialects within the same language group.

Harrison, Malake, and Amoateng, cited in Kerswell (2006), classify the development of dialects in a language into three categories: (1) geographical location in which it is spoken, (2) use by particular ethnic/ social class, and (3) prestige within the community, which is reflected in standard or non-standard dialects. These three categories can be also identified in the differences between urban and rural isiXhosa speakers, which are the result of migration (Kerswell, 2006). According to Kerswell (2006), this additional dialect in isiXhosa in urban areas is mainly found in townships, and is commonly used by young people leaving the area. Dialects share more features than differences, but there are identified differences between dialects, including phonology, lexicon, and syntax/morphology. These differences may give rise to potential bias in tests assessing language proficiency, as the aspects in which the dialects differ are measured in tests measuring language proficiency. Hence the evaluation of fairness in these tests, for use in different dialects of a language, is a priority.

1.8. PROBLEM STATEMENT

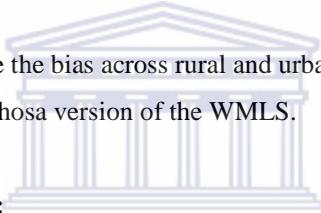
The main concern of this thesis is the equivalence of the isiXhosa version of the WMLS across urban and rural isiXhosa speaking groups, to explore the reasons for, and factors contributing to, the mean group differences referred to under 1.1. The differences amongst the groups may be linked to the changing political and educational systems of South Africa and their impact on language shifts and changes, which will be discussed in Chapter 2. These issues are of great concern in the field of educational testing in South Africa, as they affect the response patterns of learners during testing.

With these concerns, this thesis evaluates the sub-tests of the isiXhosa version of *Woodcock Muñoz Language Survey* (WMLS), looking specifically for bias across isiXhosa learners from the rural and urban Eastern Cape areas. The evaluation of this test contributes to the development of an assessment tool which is bias-free and fair towards the sample groups for which this tool was translated. The significance of this study is to contribute to the development of validity and reliability of the isiXhosa version of the WMLS, as a possible measure of language proficiency that could be used in the South African educational context, with special consideration of the urban dialects.

1.9. AIM AND OBJECTIVES

1.9.1. Aim:

The aim of this study is to evaluate the bias across rural and urban isiXhosa learners on the subtests of the isiXhosa version of the WMLS.



1.9.2. The specific objectives are:

1. To evaluate the mean differences between rural and urban isiXhosa learners on the subscales of the isiXhosa version of the WMLS
2. To evaluate group differences in terms of the reliability of the test between rural and urban isiXhosa learners on the subscales of the isiXhosa version of the WMLS
3. To evaluate group differences in terms of the mean item characteristics of the test across rural and urban isiXhosa learners on the subscales of the isiXhosa version of the WMLS
4. To evaluate the item bias or Differential Item Functioning (DIF) across rural and urban isiXhosa learners on the subscales of the isiXhosa version of the WMLS.

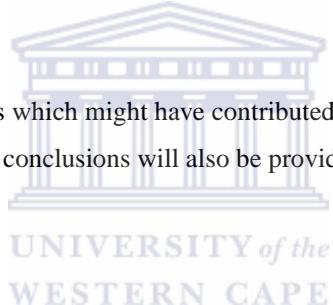
The hypotheses will be stated under the data analysis section.

Chapter 2 of this thesis will review past and present literature on the international development of testing, focusing on issues leading to the development of APA Standards, historical development of testing in South Africa, and issues of bias and equivalence in tests used in a multicultural context. In addition, the theory of equivalence will be discussed in this chapter, focusing mainly on scalar equivalence.

In Chapter 3, a discussion on the research methods used in this study is presented. This includes the research design, sampling techniques, description of the instrument, methods of analysis used, and ethical considerations of the study.

In Chapter 4, the results of the empirical investigations are presented, together with interpretation.

Chapter 5 is a discussion on issues which might have contributed to the results of this study. Recommendations and conclusions will also be provided in this chapter.

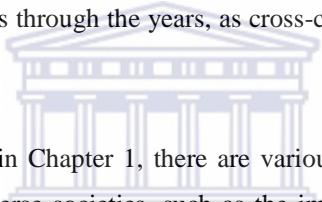


CHAPTER TWO

TESTING IN MULTILINGUAL SOCIETIES: ISSUES, THEORY AND RESEARCH

2.1. INTRODUCTION

The universal use of psychological tests in diverse cultural contexts has brought about great debates in this field. One of these is the debate on culture-fair tests, where one group of psychologists argue that differences in test scores of different cultural groups are the result of true differences, while the other group argues that the sources of test score differences amongst different cultural groups needs to be investigated as this may be the result of test bias in the psychological test used (Van de Vijver, 1998). The subject of test bias has drawn a great deal of attention from researchers and psychologists through the years, as cross-cultural testing has advanced in this field.



As mentioned in the discussions in Chapter 1, there are various issues involved when testing is carried out in diverse societies, such as the impact of education policies, language in education, and test bias. These issues anticipate various discussions on the equivalence of tests in such contexts. This chapter will focus on reviewing literature dealing with these issues of testing in diverse societies, focusing more on the issues of test bias. The evaluation of test bias in language proficiency tests across rural and urban isiXhosa learners will also be discussed in greater detail in this chapter. The chapter will be divided into seven sections:

- the background of psychology and educational testing and language as an issue in testing in diverse societies;
- standards and guidelines to address the issues of language in psychological and educational testing;
- psychological and educational testing in South Africa's multicultural society;
- the status of indigenous languages in South Africa as a context for assessment;
- isiXhosa language varieties and their relevance in testing;

- theoretical framework of equivalence and bias; and
- research on equivalence and bias in tests in South Africa.

The sections will link in a way that demonstrates the past, present and future plans in the field of cross-cultural testing, and the issue of language as a main cause of test bias.

2.2. BACKGROUND OF PSYCHOLOGICAL AND EDUCATIONAL TESTING, AND LANGUAGE AS AN ISSUE IN TESTING IN DIVERSE SOCIETIES

2.2.1. Background of psychological testing

Before discussing the background of psychological and educational testing, it is important to understand what the concept of testing entails. According to Kaplan and Saccuzzo (1997), “testing” refers to all the possible uses, and applications, of the underlying constructs of psychological and educational tests. In addition, these authors view these tests as measuring individual differences in ability and personality, assuming that the differences shown on the test may reflect actual differences among individuals (Kaplan & Saccuzzo, 1997).

The discussion on the issue of human differences in this field has long existed. This can be traced to the work of Wilhelm Wundt, who was a German psychologist and was an influential figure in the development and use of tests (Geisinger, 2000). Wundt disregarded the existence of group differences and believed that human beings (across groups) are all the same (Geisinger, 2000). In fact, Geisinger (2000) asserts that Wundt believed that whenever different groups of individuals behaved differently when exposed to the same independent variable, this was an error. Geisinger (2000) compares this perspective to currently existing perspectives of testing, noting that these are two very different views, as the current perspectives on human differences indicate that differences among groups of people may be the result of true differences.

Owing to these identified differences, many researchers emphasise the need for tests which have been specifically designed for diverse societies, such as monolingual and multilingual tests which are bias-free (Van de Vijver & Tanzer, 1997). However, tests produced for cross-cultural contexts have only been found in the early 20th Century (Carter et al., 2004). The fundamental belief of testing in diverse contexts is that culture permeates all aspects of behaviour and thus cannot be seen as separate from measuring human abilities. Greenfield (cited in Carter et al., 2004) indicates that psychological tests reflect the values, knowledge and communication strategies of their origins, which makes it important to evaluate these tests in each context where they will be used Greenfield (cited in Carter et al., 2004).

When comparing the number of published tests for the majority United Kingdom (UK) population with that of the of minority cultural groups, Carter et al. (2004) found that there were fewer available tests constructed for minority cultural groups. This comparison can also apply to other countries around the world, where the UK is the worldwide supplier of most psychological tests. Therefore, in many countries, psychologists and researchers are forced to adapt and translate these tests for use in these contexts, in order for them to be culture-fair. Hau and Chang (2001) adds that before using any psychological test in a certain context, researchers and psychologists need to verify that the test is appropriate for use with that population. Hau and Chang (2001) suggests that the way to verify the appropriateness of psychological tests is to show their validity, reliability and appropriate norm groups to which the population is to be compared.

Foxcroft, Paterson, le Roux and Herbst (2004) in their paper on test-use patterns and needs of psychological assessment in SA, asked practitioners whether psychological tests were appropriate for use cross-culturally. Results showed that only 16% of the practitioners indicated that the tests being used in South Africa were appropriate for use in a cross-cultural context such as SA, while 65.8% of this sample indicated that the tests they used were only sometimes appropriate to use in a cross-cultural context. 11% of the practitioners indicated that they did not

feel that any of the tests they used were appropriate to use in a cross-cultural context. In this study it was concluded that 58% of practitioners indicated that more culturally appropriate tests were needed in South Africa (Foxcroft et al., 2004).

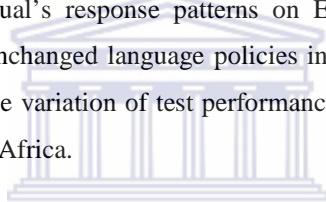
Literature has shown that there are various issues pertaining to testing in a culturally diverse context (Isaac (cited in Carter, et al., 1994), Mahon et al., 2003 (cited in Carter, et al., 1994); Marshall, 2003 (cited in Carter, et al., 1994); & Kunنان, 2000). In the present thesis, only language will be discussed as an issue in cross-cultural testing; the other identified issues are outside the scope of this thesis.

2.2.2. Language as an issue in cross-cultural testing

Cross-cultural testing, according to de Klerk (2008), occurs when psychological tests are used in different cultural settings, in order to compare test-takers from different cultural backgrounds. As stipulated in the previous chapter, this then creates a need to enhance this form of testing and to increase multilingual versions of psychological tests (de Klerk, 2008; Koch 2005). This need for multilingual versions of psychological tests has been mainly created by the influence of language in these tests, and also the variation of language in certain contexts (Foxcroft & Aston, 2006). In other words, de Klerk (2008) identifies language as one of the obvious issues in cross-cultural testing.

Foxcroft and Aston (2006) demonstrate the challenges of testing learners in multilingual contexts, by using the South African context as an example. They explain that South Africa has 11 official languages, 9 of which are African languages, together with Afrikaans and English. They demonstrate that in the educational context, English-speaking learners are educated through the medium of English, while learners who are speakers of African languages are educated in their home language until they reach grade 4, after which they will be taught in English. Afrikaans learners have the opportunity to learn through the medium of this language until undergraduate level at university.

These uses of language in the curriculum have also influenced psychological testing in the educational context. Hence most tests that are in verbal format have been generally available only in Afrikaans or English. Today some tests which are adapted to this context have been translated into indigenous languages as a result of studies which have shown that the average English-language proficiency of grade 12s in South Africa who indicate an indigenous language as their first language, is below the acceptable functional literacy level, based on the results of English Literacy Skills Assessments (ELSA) (Horne, 2001). Horne (2001) indicates that 40-60 % of matriculates who speak an indigenous language as their first language have not reached a grade 12 English functional literacy level by the time they leave school. Thus, a lack of English-language proficiency could have a detrimental effect on an individual's response patterns on English test forms (Schaap & Basson, 2003). The unchanged language policies in education can be seen as a contributing factor to the variation of test performance of learners from different cultural groups in South Africa.



This can also be demonstrated in other countries such as the USA (Stanfield, 2003), where non-native English-speaking students perform differently from students who are native English speakers in English-language proficiency tests used at school. Stanfield (2003) explains this difference to have been caused by factors such as academic background and the level of exposure to English. In the past, non-verbal tests were seen as a solution to the challenge of language in education, but Rosselli and Ardila (2003) suggest that these tests are also susceptible to cultural bias. Bedell, Van Eeden and Van Staden (1999), have recognised test translation as a means to compensate for the problems associated with culture and language in psychological testing. Stansfield (2003) demonstrates this in his paper "Test translation and adaptation in the public education in the US", by stating that any test of ability should be administered in the test-takers most proficient language, unless proficiency in the less proficient language is part of the assessment.

Further solutions to the above mentioned challenges in cross-cultural testing have been suggested, namely developing and norming tests for culturally and linguistically diverse societies (Foxcroft, 1997). However, Schaap and Vermeulen (2008) have noticed three specific problems with test translation in South Africa. Firstly, South Africa has 11 official languages and tests would have to be translated into all 11 official languages, which is costly, while there is a lack of available translators with both language and specialist psychological/human-resource expertise. Secondly, there are not enough test administrators who speak the preferred language of the test-takers, and thirdly there is the problem of different dialects of one language, and the difference in the performance between urban and rural individuals tested in their mother tongue (Bedell et al., 1999).

These challenges of test translation have brought about the need to explore methodological factors related to the responses provided by test-takers in the different language groups, in order to examine the equivalence of these different translated versions of tests. The new Standards for Educational and Psychological Testing provide guidelines regarding the testing of individuals from different linguistic backgrounds (American Psychological Association (APA), 2003).

WESTERN CAPE UNIVERSITY

2.3. STANDARDS AND GUIDELINES TO ADDRESS THE ISSUE OF LANGUAGE IN PSYCHOLOGICAL AND EDUCATIONAL TESTS.

The issue of language in testing has received attention in all the versions of the Standards of Educational and Psychological Testing. These Standards were developed to provide general criteria for evaluating the appropriateness of tests (Camara, 2007). The first version of the Standards of Education and Psychological Testing focused mainly on addressing the issue of proficiency of test-takers in the language in which tests are being administered. In Standard 9.3 of this document, it states that “when testing an examinee’s proficiency in two or more languages for which the test is available, the examinee’s relative language proficiency should be determined.” Linguistic proficiency is generally regarded as the most important single moderator of performance on assessment measures worldwide, mainly in diverse societies (Grieve, 2001; de Klerk, 2008).

However, translated tests are also being used extensively in educational testing for assessing the knowledge and skills of individuals who speak different languages (Allalouf, Hambleton, & Sireci 1999). The American Psychological Association Standards have evolved and have increasingly served as a guiding tool to those who wish to construct new tests or translate tests, and those who need to select wisely from the ones already available (Tyler & Walsh, 1979). The construction of new tests, and the adaptation and translation of tests in different contexts, have led to a stronger emphasis in the latest version of the Standards regarding fair testing in general, and the role of translation in achieving this (American Educational Research, 1992). The Standards also acknowledge the influence of social, political, historical, and economic context on individual behaviour. These Standards have been shown to be a significant influence in both educational and psychological assessments worldwide (American Educational Research, 2003).

The Standards are not the only international document which serves as a controlling body for the use of psychological tests. The American Psychological Association (APA) has adopted the Guidelines Code for Fair Testing which covers the same grounds as the Standards, with the emphasis being on fair testing in all groups, regarding their cultural background. The application of these ethical practices can be illustrated in the education system of many districts in the United State of America, where many schools conduct reading comprehension tests on school children in their native language (Sireci, & Allalouf, 2003). The same principle is being followed in Newark, NJ, where reading tests are administered in English, French, Portuguese and Spanish, so as to provide fair opportunities to all test-takers, regardless of their language (Uiterwijk & Vallen, 2006).

Another set of guidelines are the International Testing Commission's guidelines for the translation and adaptation of tests (International Testing Commission, 2001). Like the Standards, these guidelines also acknowledge good test use and encourage best practice in assessments. In addition, the International Testing Commission's guidelines promote and put great emphasis on good practice in test

adaptation and translation (International Testing Commission, 2001). The International Testing Commission (2001) guidelines view test translation and adaptations as an important step towards assuring uniformity in the quality of tests used in different cultures and languages.

These ethical documents differ from laws and regulations, in that laws and regulations are only designed to protect the public from specific abuse, while these ethical standards and codes attempt to establish a higher normative standard for a broader range of professional activities and behaviours (International Testing Commission, 2001).

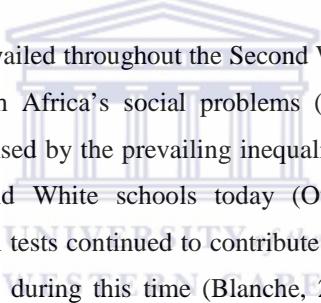
In South Africa laws and ethical standards that require psychologists and practitioners not to use tests that discriminate against any language or cultural group, are the Employment Equity Act 55, of 1998 (Government Gazette, 2009), and the PSYSA's Code of Conduct, (2004).

2.4. PSYCHOLOGICAL AND EDUCATIONAL TESTING IN SOUTH AFRICA'S MULTICULTURAL SOCIETY

The development of modern psychological and educational testing in South Africa has been identified to follow a similar trend to the United States and Europe (Foxcroft & Roodt, 2001). Like many other countries around the world, the context of testing was characterised by racial inequality (Foxcroft & Roodt, 2001). This can be demonstrated by the extensive use of English tests in America, disregarding test-takers' proficiency in the English language (Walsh & Beltz, 2000). This characteristic of inequality amongst the South African population can be traced back to when South Africa was a British Colony. The origin of psychological assessment has thus been identified to stem from this colonial heritage in South Africa (Classen, 1997).

Meiring (2007) states that psychological tests in South Africa can be traced back to as early as 1920-1960. These tests were enthusiastically imported and adapted by psychologists to this context from overseas, while they were specially designed

for use with “White” test-takers but were directly applied to the whole South African population (including Black test-takers). In the process psychologists disregarded the nature of the broader composition of the South African population (Foxcroft & Roodt, 2001). An example of this phenomenon was the development of the Fick Scale, a South African version of the Stanford-Binet, an intelligence test which was designed for testing White schoolchildren and later on was used on a large sample of Black schoolchildren (Fick, 1929). These results revealed that Black schoolchildren had a lower mean score than White schoolchildren (Fick, 1929). This study by Fick (1929) contributed greatly to the inferior teaching in Black schools. This was used to demonstrate the “inferiority” of Black children, instead of examining the effects of poverty and education on their performance on these instruments (Meiring, 2007).



This unethical form of testing prevailed throughout the Second World War, which contributed significantly to South Africa’s social problems (Blanche, cited in Owen, 1998). This can be recognised by the prevailing inequalities in the quality of education between Black and White schools today (Owen, 1998). The inappropriate use of psychological tests continued to contribute to the broadening of the gap between racial groups during this time (Blanche, 2004). Today, the main argument against using psychological tests in South Africa is that these tests are a Western invention, culturally bound, biased, and thus inappropriate to indigenous groups (Owen, 1998).

According to Van de Vijver and Rothmann (2004), the post-apartheid government has recognised the need to control these psychological measurements by currently promoting, through legislation, acknowledgement of the diversity of the South African population. For example, the Employment Equity Act 55 of 1998 specifies that psychological tests should be forbidden unless these tests have been shown to be valid and reliable and also if they are fairly applied, so that there are no biases against any group. However, Bedell et al., (1999) have argued that even though there is recognition of this diversity by government, results in psychological tests are not always equally accurate for all subgroups, mainly in a

heterogeneous society such as South Africa. Van de Vijver and Rothmann (2004) indicate the importance of bringing current practices in line with the legislation. A number of researchers in this field (Foxcroft, 2002a; HPCSA 2002; Kanjee, 2007) have pointed out that there have been a minimal number of empirical studies that have been conducted to assess bias in psychological tests in South Africa. The recognition of the need to examine the bias and equivalence of psychological tests in multicultural contexts is thus vital, as most of these tests are still imported, adapted to the South African context, and then translated into South Africa's indigenous languages, with very little research being done on these phenomena (Van de Vijver & Rothmann, 2004). The diversity in the South African context has brought about the need to thoroughly evaluate psychological and educational tests, looking specifically at the languages in which these tests are constructed, as language became the main cause of bias (Foxcroft, 2002a). Owen (1991) also mentions that most of the bias in psychological tests is brought about by language used in these tests.

2.5. STATUS OF INDIGENOUS LANGUAGES IN SOUTH AFRICA, AND AS THE CONTEXT OF ASSESSMENT

In a study which evaluated the measurements used in a cross-cultural environment, McGorry (2000) has also shown that it is highly important to utilise culturally and linguistically appropriate measurements. In South Africa, the acknowledgment of all languages has taken an extended period until the current democratic era, where there are 11 official languages approved by the Constitution, and everyone has been given the authority to express themselves in the language of their choice (Alexander, 2005). Bekker (2005) maintains that even though this transformation has taken place, indigenous languages are still slow in developing in terms of their status. Foxcroft (1997) has shown that the majority of psychological tests in South African are in English, and she acknowledges the need to develop strategies which will address these language issues within the context of assessment in this country.

2.6. ISIXHOSA LANGUAGE VARIETIES AND THEIR RELEVANCE IN TESTING

IsiXhosa is one of South Africa's indigenous languages and Statistics South Africa (2007) has shown that this language is spoken by 18% of South Africa's population, with most speakers being situated in the Eastern and Western Cape. During the colonial times, most of the indigenous people were divided into homelands according to their languages, with the assumption that these were linguistically and culturally homogeneous societies (Bekker, 2005). History shows that the isiXhosa-speaking individuals were divided into the Transkei and Ciskei, today known as the Eastern Cape (Bekker, 2005). At the moment, the Transkei remains predominantly rural, whereas the Ciskei has developed into a more urban area (Bekker, 2005).

The socio-historical position of this language has led to the development of a variety of dialects which are the result of the geographical location of isiXhosa speakers (Bekker, 2005). At present, migration, urbanisation and globalisation have been found to have an influential role in the development of new urban isiXhosa dialects that people in urban areas use extensively in their daily communication, and are even being used by learners in schools (Finlayson & Slabbert, 2003). In contrast, in rural areas of the Eastern Cape, even though they are influenced by modernity, there are still strong practices maintaining the Xhosa culture, including its languages. However, Bekker (2005) contrasts these practices with the fact that while in the Eastern Cape 83% of the population are speakers of isiXhosa; the language of instruction in schools is not isiXhosa. Strong academic proficiency in isiXhosa can thus not be assumed, even for rural isiXhosa learners. This is the situation with most indigenous languages in South Africa.

In addition, it has now been established that standard South African languages differ substantially from the language varieties that are spoken in multilingual urban areas in our provinces (Calteaux, 1996). Calteaux (1996) further indicates that the standard varieties are linguistically closest to the rural varieties than those non-standard varieties used in urban areas. Assessing the language proficiency of

learners who are speakers of isiXhosa would not be fairly done if the assessment tool that was used favoured one group of learners who were speakers of a certain isiXhosa dialect, and if these issues were not taken into account in testing in general.

As indicated in the previous chapter, one example of a test assessing language proficiency of learners is the WMLS. During an initial evaluation of the isiXhosa version of the WMLS, results revealed significant differences in the mean scores between rural and urban isiXhosa learners, with the rural isiXhosa learners scoring higher than those learners in the urban areas (Ntantiso, 2009). These differences may be the result of item bias on this test, and there is thus a need to assess the item bias of this test across rural and urban groups. If items are biased against members of a group, it means that they have the ability to respond to the item correctly, but for reasons such as item format or phrasing of the item, they do not have the same opportunity as others to do well (Bedell et al., 1999). Similar issues may affect assessment in predominantly monolingual societies in which there are different dialects; therefore the test developer and users must also be careful that their tests do not unfairly penalise speakers of non-standard dialects (Payne and Taylor 2002). Speakers of non-standard dialects have been reported to be more sensitive to the context of assessment, sometimes reducing their verbal output when relating to a standard dialect speaker, or to “hypercorrect” their speech (i.e. inappropriately apply standard dialect features) (Seymour and Miller-Jones, cited in Payne & Taylor, 2002).

2.7. THEORETICAL FRAMEWORK OF EQUIVALENCE AND BIAS

According to Van de Vijver (1998), the evaluation of measurement tools imported from a different context and then adapted to a context with a diverse cultural heritage has a variety of implications, as shown in the previous sections of this chapter. The adaptation of psychological tests is regarded as an effective way of being able to compare candidates who operate in different languages and countries, and these tests are useful when they demonstrate equivalence across different languages and countries (Robi, Sireci, & Hambleton, 2003). However,

simply translating and validating psychological and educational tests for use across groups is not enough, as further empirical steps are necessary to ensure the equivalence of such tests (Robin et al., 2003). Gierl (2000) has demonstrated this by showing that poor translation of tests can change the validity of one set of test scores, and this can therefore influence the comparability, meaning, and interpretability, if the construct measured by the two forms is not equivalent.

From a methodological perspective, Van de Vijver and Rothmann (2004) reveal bias and equivalence to be the most salient issues cross-cultural testing. These authors also state that the concepts of bias and equivalence are often treated as antonyms, with “equivalence” referring to the comparison of test scores that are obtained from different cultural or language groups, and more specifically, the extent to which scores have the same meaning across groups (Van de Vijver & Rothmann, 2004). “Bias” refers to the occurrence of score differences in the indicators of a particular construct that do not correspond with the difference in the underlying trait or ability (Van de Vijver & Tanzer, 1997). Van de Vijver and Rothmann (2004) identify three different types of equivalence, namely construct or structural equivalence, measurement unit equivalence, and scalar equivalence. Van de Vijver and Rothmann (2004) describe the different levels of equivalence as follows:

- structural equivalence occurs when the same constructs are measured
- measurement unit equivalence occurs when the measurement units are identical in both of the groups, but the origins of the scale differ
- scalar equivalence is important when the measurement instrument is used for comparison across cultural groups, and the instrument has the same origin and measurement unit in all cultures. This is the highest level of equivalence, and can assist in concluding whether average scores obtained from two cultures are comparable.

The above-mentioned levels of equivalence demonstrate the basis of cross-cultural comparisons, and as such qualify the interpretation of cultural differences (Van de Vijver & Rothmann, 2004). Furthermore, Van de Vijver and Rothmann (2004)

have shown that the equivalence of measures used for cross-cultural comparisons should be empirically established rather than presumed. According to Van de Vijver and Rothmann (2004), scalar equivalence is established when the instrument has been proven to be unbiased.

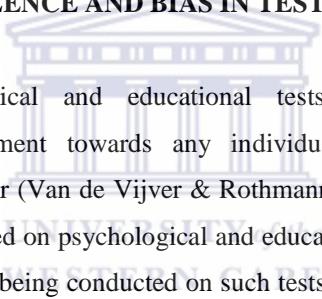
As mentioned before, “test bias” refers to the presence of a “nuisance factor” which impacts on the meaning of scores obtained with instruments. Van de Vijver and Leung (1997) have identified three types of bias, namely construct bias, method bias, and item bias or differential item functioning/DIF. Construct bias occurs when the construct measured is not identical across groups. Method bias is caused by methodological factors when doing comparative studies such as sample bias, instrument bias and administration bias.

Item bias or differential item functioning (DIF) refers to the anomalies at item level (Vijver & Leung, 1997). In other words, DIF occurs when test-takers from different subgroups who are equally skilled in the construct being measured do not stand the same chance of answering an item correctly (Uiterwijk & Vallen, 2006). The main aim of this form of bias analysis is to compare different groups on their performance on an item (Muraki, 1999). When differences between groups are not identified during DIF analysis, this could be an indication of no differences in the response patterns across groups. Identified item response differences are classified into two different kinds of DIF, namely uniform and non-uniform DIF (Zumbo, 1999). Uniform DIF exists when there is no interaction between ability level and group membership. That is, the probability of answering the item correctly is greater for one group than the other, uniformly over all levels of ability. Non-uniform DIF exists when there is interaction between ability level and group membership; that is, the difference in the probability of a correct answer for the two groups is not the same at all ability levels (Zumbo, 1999).

Zumbo (1999) has shown that all these forms of bias can result in systematic error that distorts the inferences made during selections and classification.

Hence de Klerk (2008) stresses that both bias and equivalence are important concepts in cross-cultural assessments, in order to have valid comparisons across cultural groups, or for the valid use of tests across cultural or language groups. In other words, in cross-cultural testing the equivalence of translated tests across different cultural groups is optimal when bias is minimal (de Klerk, 2008). If test versions show construct bias, there will be inequivalence in the psychological concept on which the test is based, and no comparison can be made between the two groups (de Klerk, 2008). Method and item bias may not influence construct equivalence, but they definitely influence the scalar equivalence of the test (de Klerk, 2008).

2.8. RESEARCH ON EQUIVALENCE AND BIAS IN TESTS IN SOUTH AFRICA

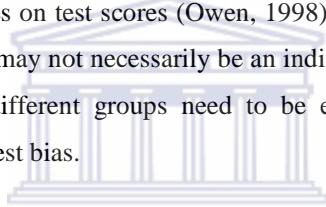


In many countries, psychological and educational tests which display discrimination and unfair treatment towards any individual or group are prohibited, as was indicated earlier (Van de Vijver & Rothmann, 2004). In South Africa, the same emphasis is placed on psychological and educational testing, and extensive research is increasingly being conducted on such tests. An example is a study conducted by Foxcroft and Aston (2006) focusing on examining language bias in the South African adaptation of the WAIS-III. A standardised sample including English first- and second-language speakers who were largely educated in English or Afrikaans was used for this study. Results revealed that the equivalence of the WAIS- III across diverse language groups has not been clearly established, and that some bias might exist for English second-language test-takers, mainly those who are Black or Afrikaans-speakers.

A variety of tests measuring constructs such as personality, intelligence and aptitude, have been found to have an element that measures a test-taker's language proficiency in the language of the test used (Owen, 1998). Another study is that of Koch (2007a; 2007b) on an English reading test used for admission into

university, in which item bias and construct bias were found across Afrikaans-, English- and African-speaking students.

In addition to the bias in English tests used across different language groups, in many contexts, tests that measure language skills are also being translated into several different languages, so that parallel versions exist for use in a multilingual context (Sireci & Allalouf, 2003). Differences in translated versions have been found to be a common challenge contributing to test bias of adapted tests in cross-cultural societies. Establishing the equivalence of different language versions is thus essential for comparisons (Foxcroft & Aston, 2006). It is thus the responsibility of the researcher and the psychologist to evaluate the causes of means-score differences between different cultural groups on a psychological test, to avoid making invalid inferences on test scores (Owen, 1998). De Klerk (2008) adds that mean-score differences may not necessarily be an indication of test bias, but these differences amongst different groups need to be explored so as to eliminate any possible causes of test bias.



In this thesis to evaluate the bias and equivalence of the WMLS, the first step was to examine group differences based on their mean scores. After differences were detected, further analyses were conducted to examine the bias of the items in this instrument across rural and urban isiXhosa learners. This was done to establish the extent to which the instrument can be validly used to assess the language proficiency of both rural and urban isiXhosa learners. It is essential to make sure that all test-takers from the different language or dialect groups with the same abilities have the same understanding of the construct (Owen, 1998). Test equivalence needs to be ensured for all groups (whether of languages or dialects) to which the test is administered (Schaap & Basson, 2003). This would ensure fairness during testing, and also when interpreting the performance of the test-takers.

The purpose of this thesis is therefore to demonstrate the scalar equivalence of the WMLS as a tool measuring language proficiency, in order to compare rural and

urban isiXhosa-speaking learners in their language proficiency. This will be done by using empirical evidence of this assessment tool. Chapter 3 will discuss the methods which were used to achieve the general and specific aims of this thesis.



CHAPTER THREE

METHOD

3.1. INTRODUCTION

This chapter will reflect on the processes that were undertaken in conducting this study. The chapter will include the research design that was followed, procedures, participants, instruments, analysis and ethical considerations which were made during this process.

3.2. RESEARCH DESIGN

A quantitative research approach was applied in this study. The approach coincides with the aims and objectives of the study. An essential quality of this approach is that it views social research through an objective lens, so as to avoid any bias during the process of acquiring data and analysing this data (Creswell, 2003). For the purpose of this study, a differential research design was used to evaluate the item bias of the isiXhosa version of the WMLS across urban and rural isiXhosa learners to measure their language proficiency. In differential research it is important to identify the variable that defines each group, so that the test used can accommodate these differences even though they are measured on the same construct (Graziano & Raulin, 2000). In this study the variable that defines each group is the context in which these participants are located, namely a rural or an urban South African context. Graziano and Raulin (2000) identify an important use of differential research design, which is to compare existing groups on theoretically relevant variables, and also to test experimental procedures if they are thought to be unethical.

3.3. PARTICIPANTS

Both secondary data and primary data collection also formed part of this study. The secondary data consisted of 133 urban (Port Elizabeth in the Eastern Cape) isiXhosa learners and 55 rural (King Williams Town in the Eastern Cape) isiXhosa learners who spoke a more standard dialect of isiXhosa. In the original study the groups were of unequal sizes, as comparison of these groups was not the

focus of the study at the time. For the purpose of the data analysis in the present study, equal group sizes were ensured so as to avoid sampling error. Messick (1989) classifies “sampling error” as one of the threats when drawing research conclusions. Messick (1989) further notes that when this type of error occurs, analysts may identify differences between groups even though this difference may have occurred as a result of unequal or improper sampling. In this study, additional data was collected in order to avoid sampling error. Hence, additional data from 78 learners in the deep rural Eastern Cape was collected, and the conditions of testing and context of testing were similar to the previous data that was collected from the deep rural area.

In differential research, participants are assigned to groups on the basis of the same pre-existing variable, and the groups will often differ on several variables other than the dependent variable (Graziano & Raulin, 2000). The sampling strategy that was used in the original study was a purposive convenience sampling technique; in this sampling technique one can select participants dependent on the purpose of the study, but it is non-probability sampling. The same sampling procedure was followed in the added data of this study. There was no generalisability in this study because the sample group was limited in terms of their geographical location. Learners from both of these groups in this data used English as a medium of instruction in school (both groups are from ex-DET school settings), but the school context was dominated by isiXhosa speakers in both school settings. The learners in the original study consisted of males and females in grade 6 and 7, ranging in age from 11-14 years, and both males and females from the same grades will be used in the added sample. The age group selected in this study falls in the age range in which this test can be administered, which is 3-99 years (Woodcock & Sandoval, 2005). Owing to time limitations, and because of the exploratory nature of this research, the 11-14-year age range was selected.

3.3.1. DESCRIPTION OF PARTICIPANTS

This section will describe the nature of the sample used in this thesis, based on their grade and gender. This sample was selected from a population of learners from schools in the Eastern Cape from both rural and urban areas.

Table 1
Number and percentages of participants

IsiXhosa learners	n	%
Rural	127	49
Urban	133	51
Total	260	100

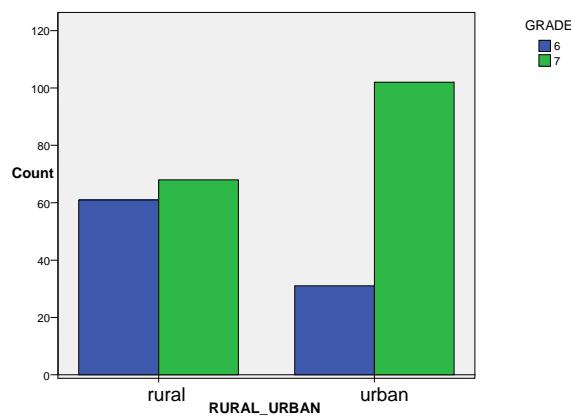
Table 1 demonstrates the distribution of the sample by groups out of 260 learners who took part in this study, and indicating that the size of the groups was fairly similar, with slightly more urban isiXhosa learners than rural ones.

Table 2
Description of both rural and urban isiXhosa learners in grades

IsiXhosa learners	GRADE	
	6	7
Rural	60	67
Urban	31	102
Total	91	169

Table 2 indicates that there were more grade 7 isiXhosa learners than grade 6 isiXhosa learners in the urban areas. Below is a bar graphic representation of the difference by grade in this sample.

Figure 1
Presents grade differences between rural and urban isiXhosa learners



The above figure shows that there were more grade 7 isiXhosa learners in both groups. However, the proportion of grade 7 learners is higher for the urban isiXhosa learners than for the rural learners. This variation in the proportion of grade 7 to grade 6 learners across these groups may influence the results in this thesis, as the group with more grade 7 learners may perform better owing to their level of academic exposure. This factor will be taken into account in the discussion of the results throughout.

Table 3
Descriptive of sample by gender (rural and urban)

Group	GENDER		Total
	Female	Male	
Rural	63	64	127
Urban	78	55	133
Total	141	119	260

Table 3 indicates that there were more males than females in the sample of rural isiXhosa learners; however this table further indicates that there were more female than male urban isiXhosa learners from this sample.

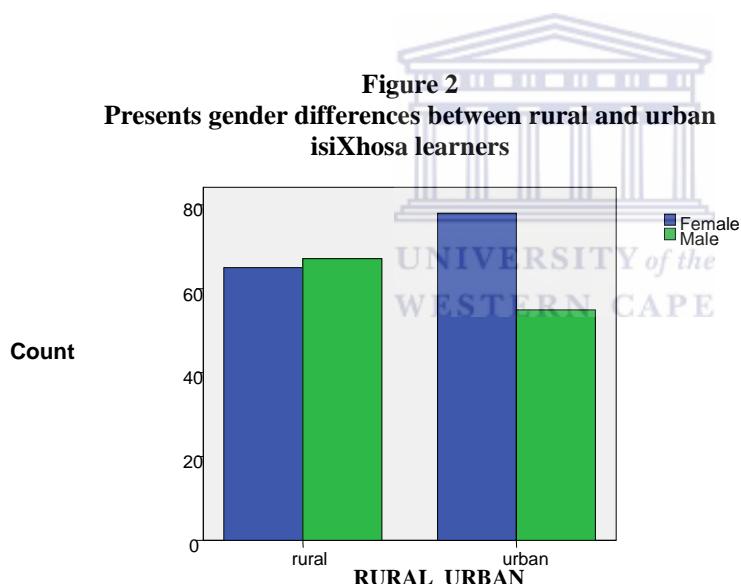


Figure 2 shows that there were slightly more males than females amongst the rural isiXhosa learners, while the opposite applied to the urban isiXhosa learners, where there were more female than male learners.

3.4. DATA COLLECTION PROCEDURES

This study stems from a bigger project, namely, the Additive Bilingual Education (ABLE) project that aims to promote additive bilingual education through the

medium of both English and isiXhosa. The WMLS was selected as an instrument that was used in the evaluation of this project. For the purpose of this project the test needed to be in English as well as isiXhosa. The instrument was thus translated and adapted into isiXhosa. A second project, consisting of several phases, was implemented to assess the psychometric properties of both versions of this test for the South African context. A brief description of the different phases of this project follows.

- First phase: The adaptation of the English and isiXhosa versions of the test.
- Second phase: First round of data collection and the statistical evaluation of the equivalence of the English and isiXhosa versions of the test.
- Third phase: Further adaptation of the isiXhosa version on the basis of the results of the second phase, and further research into the equivalence of the two language versions of the test.
- Fourth phase: An evaluation of the predictive validity of both versions of the WMLS, as well as an evaluation of the content validity of this instrument.

The specific focus of the present thesis is the evaluation of the equivalence in the isiXhosa version of this instrument across rural and urban isiXhosa learners; using procedures as explained under “Participants” (see 3.3 above.) The procedures which were followed for data collection were the same as those of the primary data which was used in this study, in that, before the data collection started, test administrators went through intensive training to learn how to administer the WMLS so as to keep the process of administration standard. For this thesis, the data collection was carried out in a school in the deep rural areas of the Eastern Cape, with permission provided by the Department of Education in the Eastern Cape, the principal, parents, and learners. As per ethical procedures for testing, the data collection took place in September 2008 before the beginning of the end-of-year exams, so as not to disrupt learners during exams. The instrument was individually administered to learners in a controlled environment. The data collection was conducted during school hours, and as a preventative measure not to disrupt the teaching and learning process in the classroom, individual learners

were assessed and they went back to class after the assessment. It took an estimated 40 minutes to administer the instrument to individual learners; the data collection was conducted over a period of 5 days to acquire a sample of 78 rural isiXhosa learners. Three days were spent assessing grade 6 learners, while two days were spent assessing grade 7 learners.

After the data collection the data was stored in a safe place, and later captured and cleaned by the researcher. This data was added to the existing database of the primary data. All the analyses in this thesis were conducted by the researcher on the Statistical Package for Social Sciences (SPSS) version 15.0.

3.5. INSTRUMENT

The secondary data and additional primary data which was used in this study was based on the isiXhosa version of the *Woodcock Muñoz Survey* (WMLS). The WMLS was originally developed to assess cognitive-academic language proficiency in English and Spanish in Argentina, Costa Rica, Mexico, Peru, Puerto Rico, Spain and the United States of America (USA). It consists of four sub-tests, which are: Test 1: Picture Vocabulary; Test 2: Verbal Analogies; Test 3: Letter-Word-Identification; Test 4: Dictation. The four subtests of the WMLS form a cluster which is a combined measure of expressive vocabulary, verbal reasoning, reading identification, and writing skills. Descriptions of this instrument are indicated in Appendix A. The reliability of the original test was calculated for the USA population, using the split-half procedure with the median reliability ranging from 0.80 to 0.93 content. The concurrent and construct validity of the WMLS Normative Update was identified in the USA context. These proved to be satisfactory (Woodcock & Sandoval, 2005). The psychometric characteristics of both SA versions of the test are still being established, and this thesis forms part of this research project.

3.6. ETHICAL CONSIDERATIONS

Permission to conduct the original study with grade 6 and 7 learners was sought from the Department of Education in the Eastern Cape, and the same procedure was followed to access participants for this study. The principals of the schools in which the research was conducted were also asked for permission to conduct the study in their schools, and were informed that their willingness to form part of the project was voluntary. The isiXhosa version of the information sheet (Appendix B), informed consent forms (Appendix C), and the assent form (Appendix D) were available for the original study, and these forms were also included in this study as required by the University of the Western Cape (UWC) Senate Research Committee. Participants were informed about the purpose of the study, and told that participation was voluntary and they could withdraw at any stage of the testing. Confidentiality of the results during the original study was provided, and clear instructions were given before the testing, in a language that was understood by participants. The same procedures were followed for this study. Results of this test would be stored in a safe place in the same way as in the original study. No harm was caused during the participation in the original study and participation in this study also caused no harm to participants. The ethical statement is attached in Appendix E.

3.7. DATA ANALYSIS

3.7.1. Specific objective one: to evaluate mean differences between rural isiXhosa learners and urban isiXhosa learners on the isiXhosa version of the WMLS:

For the above-mentioned specific aim, descriptive statistics, namely means and standard deviations (SD), were derived, after which a Hotelling's T² was conducted. Hotelling's T² assesses the overall mean differences between two groups on a profile of tests scores (Field, 2005), in this case, the sub-tests of the WMLS. Post-hoc T-tests were also conducted to identify where differences were to be found in cases where significant overall F-statistics were found.

Null hypothesis: There are no mean differences between rural and urban isiXhosa learners on the isiXhosa version of the WMLS.

3.7.2. Specific objective two: to evaluate group differences in terms of the reliability of the test between rural and urban isiXhosa learners of the isiXhosa version of the WMLS.

The Cronbach's Alpha of each sub-test was calculated for each group, and the results of the two groups were compared using the following statistics $(1-\alpha_1)/(1-\alpha_2)$, (Van de Vijver & Lueng, 1997). The statistics follow an F- distribution at n_1-1 and $n_2 - 1$ degrees of freedom

Null hypothesis: there will be no difference between the Cronbach Alphas of the two groups on the sub-tests of the WMLS.

3.7.3. Specific objective three: to evaluate group differences in terms of the item characteristics of the test between rural and urban isiXhosa learners on the isiXhosa version of the WMLS.

Mean item difficulties and the mean standard deviation of item difficulties was determined for this specific objective. This computation was conducted to identify how many learners, on average, in each group scored the items correctly; these groups were compared descriptively based on these scores. Mean item discrimination values were calculated and compared descriptively for this specific aim.

Individual items analyses were also conducted to identify possible problematic items that performed differentially across the two groups. An item characteristic analysis was conducted; in which differences and similarities were interpreted, based on item difficulty or easiness across groups, and how items discriminated differently across both groups.

No null hypothesis will be formulated for this aim.

3.7.4. Specific objective four: to evaluate the item bias or differential item functioning, across rural and urban isiXhosa learners of the isiXhosa version of the WMLS.

Logistic regression analysis was used to answer this objective. This analysis technique was used to detect the contribution of ability, group membership and interaction to the probability of a correct response. This statistical technique also calculates correct response to an item by means of the following model in order to identify DIF (Zumbo, 1999):

$$P = (u = 1 | \theta, g) = \frac{e^{\tau_0 + \tau_1 \theta + \tau_2 g + \tau_3 (\theta g)}}{e^{\tau_0 + \tau_1 \theta + \tau_2 g + \tau_3 (\theta g)}} \quad (1)$$

In this model, the parameters τ_0, τ_1, τ_2 , and τ_3 show an intercept and the weights for the ability, group difference, and the ability and group interaction terms, respectively, θ is ability denoted by the total test score, and g is the group membership, in this case coded as 0 for the reference group and 1 for the focal group.

Four steps were undertaken to interpret these results. The analyses were conducted for each item individually per sub-test. The following steps were followed:

- The first step was to evaluate the model fit, by calculating the differences in the Chi-square (DIFF square) between the first and the third step in the model. The significance of the overall model (from step 1 to step 3) was determined, using the critical value of the Chi-Square distribution at 2 degrees freedom. In this study a critical value of 9.55 at alpha = 0.01 was used, to control for increased type 1 error caused by the repeated analyses as well as the sensitivity of item bias analysis for this form of error.
- The R^2 difference values (DIFF R^2) between the first and third steps were calculated. The effect size, using the $R^2 \Delta$, was calculated to identify the

size of the DIF and as a measure to further control for type 1 error (Zumbo, 1999). The effect size was used to categorize the DIF in the following manner:

- a) Negligible DIF: $R^2 \Delta < .035$
- b) Moderate DIF: $.035 < R^2 \Delta \leq .060$
- c) Large DIF: $R^2 \Delta > .060$

Only items which displayed moderate DIF and large DIF were reported in this study.

- Items were classified as either constituting uniform DIF or non-uniform DIF. According to Zumbo (1999), uniform DIF is identified by the $R^2 \Delta$ from step 1 to step 2, while the non-uniform DIF is identified by the $R^2 \Delta$ from step 2 to step 3, as well as by evaluating the beta values in the regression equation for both these forms of DIF. When an item displays uniform DIF, the Beta value for group membership on the item will be significant. A negative Beta value on group membership shows that this item favours the focal group (urban isiXhosa learners), while a positive Beta value on group membership indicates that this item favours the reference group (rural isiXhosa learners). Items displaying non-uniform DIF are identified when the Beta value of the interaction between group membership and ability is significant. An item with a negative Beta value on this interaction shows that this item favours high-ability urban learners and low-ability rural learners. If the Beta value is positive, then high-ability rural and low-ability urban learners are favoured.

Null hypothesis (per subscale and all items): the probability of scoring 1 on the item i will be the function of the intercept and ability only.

The next chapter (5) will give the results of the analyses which were conducted in this study.

CHAPTER FOUR RESULTS

4.1. INTRODUCTION

The overall aim of this study is to evaluate the bias, more specifically item bias, across rural and urban isiXhosa learners on the sub-tests of the isiXhosa version of the WMLS. The evaluation of bias in all psychometric tests is of crucial importance as several inferences are made on such tests (Cole & Moss, 1993). Therefore test users need to verify whether there are any score differences between the groups of test takers, and if there are, it is essential to reveal what the causes of these difference are. A previous study of the WMLS revealed that there were significant mean score differences between rural and urban isiXhosa learners (Ntantiso, 2009). This present study therefore focuses on revealing whether these mean score differences between these groups are the result of item bias (DIF) when we add more rural isiXhosa learners to the original sample in the previous study, or real differences in the constructs of language proficiency as measured in the test. This chapter reflects on the results and discussions of this study, and also reveals steps undertaken to identify the existence of bias on the sub-tests of the WMLS across rural and urban isiXhosa learners. The results will be presented for each specific aim of this study.

4.2. GROUP DIFFERENCES ON TOTAL MEAN SCORES ACROSS RURAL AND URBAN ISIXHOSA LEARNERS

Research Aim 1: to evaluate mean differences between rural isiXhosa learners and urban isiXhosa learners on the isiXhosa version of the WMLS.

Descriptive statistics for the study sample are presented in Table 4.

Table 4
**Mean Scores and Standard Deviations for rural and urban isiXhosa learners
on each of the subscales of the WMLS**

	Picture Vocabulary		Verbal Analogies		Letter-Word-Identification		Dictation	
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
Rural	23.81	3.78	10.06	4.86	50.88	6.29	35.53	6.98
Urban	23.09	4.08	11.62	4.39	49.96	6.95	29.23	7.59

* Std dev = Standard deviation

Overall this table demonstrates that rural isiXhosa learners had slightly higher mean scores on the sub-tests: Picture Vocabulary, Letter-Word- Identification and Dictation than the urban isiXhosa learners, while the urban isiXhosa learners had a higher mean score on the Verbal Analogies sub-test than the rural isiXhosa learners. Figure 3 below is a graphic representation of the overall means score performance across rural and urban isiXhosa learners on the WMLS subtests.

Figure 3
Mean scores on the four sub-tests across rural and urban isiXhosa learners

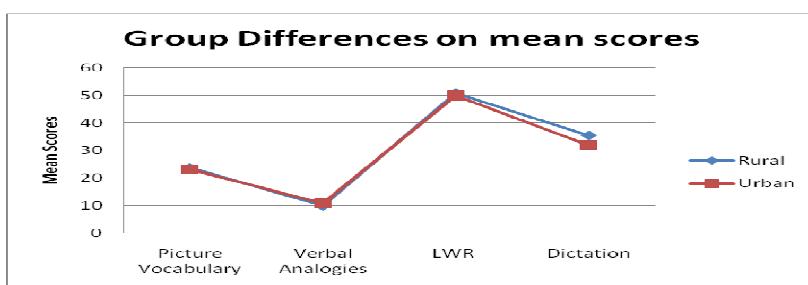


Figure 3 illustrates that the mean scores between rural and urban isiXhosa learners were similar but there seem to be discrepancies in the mean scores of the Verbal Analogies and Dictation sub-tests between the two groups.

Table 5**Hotelling's T² results: differences between two groups on all the sub-tests**

T^2 (casewise MD) = 0.45 F = 28.601 p < .0000				
Subtests	Mean Diff	Post-hoc t-value	Df	P
Picture Vocabulary	33.75	2.17	261	0.14
Verbal Analogies	159.91	7.46	261	0.01
Letter-Word-Identification	68.22	1.54	261	0.21
Dictation	2573.95	48.25	261	0.00

The above results of the Hotelling's T² show that overall there were differences between rural and urban isiXhosa learners on the isiXhosa version of WMLS.

The post-hoc test indicated that these group differences were significant on only two of the sub-tests on the isiXhosa version of the WMLS (at an alpha level of 0.05). These sub-tests were the Verbal Analogies and the Dictation. The mean score differences identified on these subtests need to be further investigated, as it is stated on the APA Standards (2003) that mean score differences between groups need to be investigated to see what contributes to these differences.

The null hypothesis of no differences is thus rejected for the sub-tests Verbal Analogies and Dictation.

Comment [S1]: Spacing changes from here till page 50s – correct. Check page nrs on content page after you have changed

4.3. GROUP DIFFERENCES ON INDEXES OF RELIABILITY ACROSS THE FOUR SUBTESTS OF THE WMLS, BETWEEN RURAL AND URBAN ISIXHOSA LEARNERS

Research Aim 2: To evaluate group differences in terms of the reliability of the test between rural and urban isiXhosa learners in the isiXhosa version of the WMLS.

Table 6
Presents the indexes of Reliability (Cronbach's Alpha) of the sub-tests of the WMLS across rural and urban isiXhosa learners

Rural_Urba n	Cronbach's Alpha			
	Picture Vocabulary	Verbal Analogies	Letter-Word- Identification	Dictation
Rural	0.77	0.85	0.89	0.87
Urban	0.77	0.64	0.90	0.88

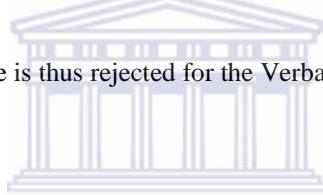
Table 6 illustrates the internal consistency of items for each sub-test of the WMLS for both rural and urban isiXhosa learners. De Vellis (2003) states that an acceptable reliability coefficient is 0.70 and above, and may be used to make decisions about individuals. This table thus shows that the Picture Vocabulary sub-test has an acceptable reliability coefficient for both rural and urban isiXhosa learners. In addition to this, this table shows that the Verbal Analogies sub-test had a questionable internal consistency for one group (urban isiXhosa learners) while for the other group (rural isiXhosa learners) the reliability coefficient was acceptable. Furthermore this table shows that there was a good internal consistency between items on the Letter-Word-Identification sub-test for both rural and urban isiXhosa learners. Lastly the table indicates that there was a good internal consistency of items in the Dictation sub-test for both rural and urban isiXhosa learners. These reliability analyses reveal that there could be a problem with the internal consistency of some of the WMLS sub-tests for the urban isiXhosa learners when comparing rural and urban isiXhosa learners. The following statistics, namely $(1-\alpha_1)/(1-\alpha_2)$, will reveal the equality of reliability for all sub-tests for the two groups. The critical value for the F distribution was 1.3 at an alpha level of 0.01.

Table 7
Test of Equality Reliability for the WMLS subtests

	Picture Vocabulary	Verbal Analogies	Letter-Word-Identification	Dictation
Test of	1.0	2.4	0.90	0.92
p-value	>.01	<.01	>.01	>.01

This table indicates that, for three of the sub-tests (Picture Vocabulary, Letter-Word- Identification and Dictation) there were no significant differences between the reliability of the rural and urban isiXhosa learners. However, for one sub-test of the WMLS (Verbal Analogies) there were significant differences in the reliability of the two groups. These analyses thus point to the direction of equivalence of the Picture Vocabulary, Letter-Word-Recognition and Dictation sub-tests across rural and urban isiXhosa learners, but indicate the possibility of problems in equivalence on the Verbal Analogies sub-test.

The null hypothesis of no difference is thus rejected for the Verbal Analogies sub-test.



4.4. GROUP DIFFERENCES ON ITEM CHARACTERISTICS IN THE SUBTESTS OF THE WMLS ACROSS RURAL AND URBAN ISIXHOSA LEARNERS

Research Aim 3: To evaluate group differences in terms of the item characteristics of the test across rural and urban isiXhosa learners on the isiXhosa version of the WMLS

Tables 8 and 9, below present the descriptive mean item difficulty (p-value) and item-total correlation for each subscale of the WMLS.

Table 8
Mean item difficulty values and mean standard deviations across rural and urban isiXhosa learners on the four subtests of the WMLS.

Rural_Urban Group	Picture Vocabulary		Verbal Analogies		Letter-Word-Identification		Dictation	
	Item Difficulty	Std dev	Item Difficulty	Std dev	Item Difficulty	Std dev	Item Difficulty	Std dev
Rural	0.41	0.17	0.29	0.30	0.89	0.20	0.63	0.30
Urban	0.39	0.19	0.33	0.34	0.88	0.27	0.52	0.34

*Std dev = Standard deviation

The mean item difficulty for the rural and urban isiXhosa learners on three of the four sub-tests (Picture Vocabulary, Verbal Analogies, and Letter-Word-Identification) revealed an approximate similarity. The mean item difficulty scores of most of the sub-tests fall in a range which according to Schultz and Whitney (2005), is the desirable range in item-difficulty index analysis (0.3 and 0.7). The mean item difficulty score on the Verbal Analogies sub-test indicates that the items of this sub-test were slightly more difficult for the rural isiXhosa learners, while the Letter-Word-Identification sub-test's mean item difficulty values showed this sub-test to be easier than the other sub-tests for both rural and urban isiXhosa learners, and this easiness falls out of the accepted range of easiness. Furthermore, the above table shows that the items on Dictation were easier for the rural isiXhosa group than the urban isiXhosa group.

Table 9
Mean item discrimination on the four subtests of the WMLS across rural and urban isiXhosa learners

Rural_Urban Group	Picture Vocabulary	Verbal Analogies	Letter-Word-Identification	Dictation
Rural	0.15	0.39	0.25	0.25
Urban	0.17	0.24	0.35	0.31

The above table shows that the mean values for the item-total correlation of the two groups were approximately similar, with a slightly lower mean item discrimination value for the rural group on the Letter-Word-Identification and Dictation sub-tests, and a slightly lower mean item discrimination value for the

urban group in the Verbal Analogies sub-test. This table also indicates that the Picture Vocabulary subscale was the only which had a poor mean discrimination power (for both groups), as its item-total correlation was below the 0.2 – 0.4. These are acceptable values for an item-total correlation (Foxcroft & Roodt, 2001).

Tables 10 to 13 present the item difficulty and item discrimination analysis for all the individual items of the sub-tests across the two groups.



Table 10
Item difficulty and item discrimination analysis per group: Picture Vocabulary

Items	Item difficulty (p-values)		Item discrimination (item-total correlations)	
	Rural	Urban	Rural	Urban
1	1	0.99	0.00	0.23
2	1	0.97	0.00	0.04
3	0.99	0.95	0.04	0.22
4	1	0.98	0.00	0.14
5	1	0.98	0.00	0.07
6	1	0.98	0.00	0.09
7	1	0.96	0.00	0.06
8	0.98	0.97	0.05	0.08
9	0.98	1	0.07	0.00
10	1	1	0.00	0.00
11	1	1	0.00	0.00
12	1	0.99	0.00	0.04
13	0.98	0.98	0.12	0.11
14	0.53	0.66	0.32	0.34
15	0.99	0.99	0.04	0.04
16	1	0.98	0.00	0.12
17	0.88	0.89	0.11	0.22
18	0.84	0.77	0.15	0.23
19	0.29	0.49	0.27	0.20
20	0.93	0.98	0.04	0.21
21	0.73	0.49	0.22	0.32
22	0.62	0.59	0.16	0.29
23	0.55	0.33	0.17	0.36
24	0.43	0.27	0.34	0.27
25	0.09	0.09	0.28	0.22
26	0.20	0.07	0.62	0.22
27	0.87	0.77	0.21	0.57
28	0.30	0.34	0.41	0.45
29	0.20	0.05	0.26	0.26
30	0.25	0.42	0.58	0.33
31	0.14	0.07	0.22	0.15
32	0.06	0.31	0.53	0.50
33	0.24	0.04	0.30	0.22
34	0.13	0.09	0.26	0.25
35	0.14	0.24	0.62	0.45
36	0.06	0.08	0.36	0.30
37	0.00	0.02	0.00	0.20
38	1	0.02	0.00	0.09
39	0.2	0.01	0.27	0.07
40	0.05	0.03	0.37	0.21
41	0.00	0.00	0.00	0.00
42	0.00	0.01	0.00	0.16
43	0.00	0.01	0.00	0.16
44	0.02	0.01	0.27	0.16
45	0.01	0.02	0.17	0.26
46	0.00	0.02	0.00	0.15
47	0.01	0.02	0.17	0.26
48	0.01	0.00	0.17	0.00
49	0.00	0.01	0.00	0.04
50	0.00	0.00	0.00	0.21
51	0.00	0.00	0.00	0.00
52	0.00	0.00	0.00	0.00
53	0.00	0.00	0.00	0.00
54	0.00	0.00	0.00	0.00
55	0.00	0.00	0.00	0.00
56	0.00	0.01	0.00	0.21
57	0.00	0.01	0.00	0.21
58	0.00	0.00	0.00	0.00

The table shows that at the beginning of the Picture Vocabulary sub-test the items were easy for both groups. It then shows that the degree of easiness in these items was slightly different across the two groups, with the rural isiXhosa learners performing slightly better than the urban learners up to item 21. The performance of both groups decreased slightly from item 21, with the rural learners still performing slightly better than the urban learners. The sub-test items became more difficult after item 29 for both these groups, with a variation in the level of difficulty for each group. Items 3, 17, 18, 20, 22, 27 and 39 showed different item total discrimination across the rural and urban isiXhosa learners, meaning that these items did not distinguish clearly between high- and low-ability learners across rural and urban learners. However, the rest of the items seemed to discriminate well between high- and low-ability learners in both groups.

There seems to be a lot of resemblance in the performance of the two groups on items of this sub-test, and this anticipates the equivalence between the performances of these groups on the isiXhosa version of the WMLS. However, attention needs to be given to the items on which rural isiXhosa learners gave more correct responses than urban ones. The slight differences in these items may be the result of daily exposure of rural isiXhosa learners to the objects in the pictures of this sub-test. In other words, culturally loaded items may be part of the reason why rural isiXhosa learners generally seem to have performed slightly better than urban isiXhosa learners. Where items possibly favoured urban learners (for example items 19 and 32), this may have been the result of the exposure that urban isiXhosa learners have to media such as television and the borrowing of English words into isiXhosa. An example of this may be the picture of the *ihelikhoptha*, (*item 19: Picture Vocabulary*), on which 50% of urban learners scored correctly, compared to the 29% of the rural learners.

Table 11
Item difficulty and item discrimination analysis per group: Verbal Analogies

Items	Item difficulty (p-values)		Item discrimination (item-total correlations)	
	Rural	Urban	Rural	Urban
1	0.76	0.82	0.15	-0.03
2	0.56	0.72	0.32	0.21
3	0.60	0.85	0.38	-0.03
4	0.63	0.84	0.55	0.22
5	0.98	0.97	0.14	-0.05
6	0.67	0.72	0.38	0.05
7	0.70	0.49	0.10	0.08
8	0.74	0.76	0.27	-0.09
9	0.39	0.48	0.45	0.23
10	0.37	0.47	0.57	0.35
11	0.53	0.65	0.49	0.29
12	0.55	0.63	0.50	0.16
13	0.64	0.69	0.41	0.29
14	0.17	0.35	0.33	0.27
15	0.35	0.39	0.49	0.21
16	0.22	0.44	0.50	0.35
17	0.16	0.16	0.32	0.20
18	0.49	0.59	0.44	0.14
19	0.07	0.12	0.37	0.44
20	0.02	0.11	0.36	0.51
21	0.07	0.07	0.36	0.29
22	0.10	0.08	0.45	0.39
23	0.03	0.03	0.38	0.52
24	0.02	0.08	0.39	0.46
25	0.02	0.03	0.43	0.35
26	0.04	0.05	0.40	0.56
27	0.02	0.05	0.39	0.39
28	0.02	0.04	0.36	0.46
29	0.03	0.05	0.30	0.40
30	0.01	0.02	0.44	0.44
31	0.01	0.01	0.44	0.40
32	0.91	0	0.44	0.00
33	0.01	0	0.44	0.00
34	0.01	0.02	0.44	0.20
35	0.01	0	0.44	0.00

The overall item difficulty on the Verbal Analogies sub-test was low for both groups (see Table 8). In this sub-test many items possibly favoured the urban learners, with items 2, 3, 4, 14, 16, and 19 showing large differences between the two groups in terms of item difficulty. However, most of these items were equally discriminating in the two groups. Only item 32 was easier for the rural isiXhosa learners than the urban group. The performance of both groups

decreased from item 19 until the end of this sub-test. The item total correlations across items appears to indicate good discriminatory power for both groups, though items 1 and 5 appear to have poor (with some negative item total correlations) discriminatory power between high and low performers in both groups of learners. Items 3, 6 and 8 displayed differences in the item total correlations across the two groups.

These results appear to be promising for the equivalence of the sub-test across the two groups; however, more attention needs to be given to those items which seem to favour the urban isiXhosa learners, especially in the light of the results on the test of equal reliability. The higher performance of urban isiXhosa learners in this sub-test may also be a result of real developmental differences in the academic verbal reasoning between different grades, as the sample of rural learners was constituted more from grade 6 learners, and the urban isiXhosa group from more grade 7 learners.

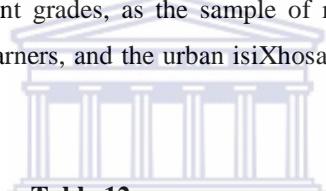


Table 12
Item difficulty and item discrimination analysis per group: Letter-Word-Identification

Items	Item difficulty (p-values)		Item discrimination (item-total correlations)	
	Rural	Urban	Rural	Urban
1	1	1	0	0
2	1	1	0	0
3	1	1	0	0
4	1	1	0	0
5	0	0.89	0	0.26
6	0.99	0.92	-0.07	0.33
7	0.99	0.99	-0.07	0.11
8	0	0.95	0	0.24
9	0	0.98	0	0.33
10	0	0.95	0	0.38
11	0	0.91	0	0.19
12	0	0.91	0	0.29
13	0.98	0.78	-0.09	0.45
14	0	0.97	0	0.31
15	0.99	0.92	-0.06	0.35
16	0	0.98	0	0.29
17	0	0.98	0	0.41
18	0	0.99	0	-0.00
19	0	0.98	0	0.31
20	0.98	0.83	-0.11	0.29
21	0.98	0.76	-0.09	0.43
22	0	0.98	0	0.41
23	0	0.96	0	0.09
24	0	0.99	0	0.47
25	0	0.92	0	0.36

26	0	0.96	0	0.39
27	0.99	0.92	-0.04	0.31
28	0	0.92	0	0.42
29	0	0.93	0	0.40
30	0.95	0.98	0.50	0.38
31	0.94	0.91	0.59	0.29
32	0.94	0.96	0.59	0.31
33	0.96	0.96	0.47	0.45
34	0.90	0.94	0.51	0.26
35	0.62	0.64	0.52	0.41
36	0.87	0.87	0.49	0.35
37	0.90	0.93	0.60	0.37
38	0.67	0.69	0.44	0.53
39	0.89	0.90	0.44	0.46
40	0.75	0.85	0.57	0.48
41	0.69	0.82	0.64	0.51
42	0.86	0.90	0.56	0.52
43	0.87	0.94	0.60	0.49
44	0.68	0.67	0.61	0.49
45	0.87	0.91	0.57	0.44
46	0.83	0.86	0.56	0.49
47	0.86	0.89	0.59	0.46
48	0.81	0.76	0.46	0.36
49	0.56	0.72	0.45	0.45
50	0.52	0.50	0.46	0.39
51	0.81	0.79	0.55	0.41
52	0.70	0.71	0.63	0.41
53	0.70	0.65	0.56	0.47
54	0.65	0.63	0.58	0.43
55	0.75	0.72	0.56	0.51
56	0.76	0.72	0.52	0.49
57	0.74	0.76	0.56	0.52

**UNIVERSITY of the
WESTERN CAPE**

The Letter-Word-Identification sub-test revealed an overall similarity in the performance of both groups in this version of the WMLS, as 50%+ of learners in both groups had correct responses on most of these items, with more rural learners scoring slightly better than urban learners. However, the high performance of these groups decreased slightly in items 49 - 54, as is to be expected.

The slight similarities in the performance of rural and urban learners are promising for the equivalence across the two groups of this sub-test. The problem with this sub-test is its tendency to be quite easy, which may indicate that this sub-test may not reveal the true ability of learners on this aspect of language proficiency. On the other hand, it may be a true reflection of the nature of the orthography of this language, and the fact that children generally do not struggle with word recognition in isiXhosa because of the direct relation between sound and letter combination in the written language. The urban group managed to

answer a number of these items correctly, and from item 30 all items seemed to have good discrimination for this group. Most items in this sub-test seemed to discriminate well for both rural and urban isiXhosa learners. Only items 6 and 20 showed discrepancies across the two groups with regard to item total discrimination.

Again, these results appear to be very promising, with much evidence pointing in the direction of equivalence across the rural and urban isiXhosa dialects on the isiXhosa versions on this subscale. Some attention needs to be given to items which demonstrated possible DIF.

Table 13
Item difficulty and item discrimination per group: Dictation

Items	Item difficulty (p-values)		Item discrimination (item-total correlations)	
	Rural	Urban	Rural	Urban
1	0	0	0	0
2	0	0.99	0.00	0.83
3	0	0	0.00	0.00
4	0	0.99	0.00	0.04
5	0	0	0.00	0.00
6	0	0	0.00	0.00
7	0	0.99	0.00	-0.01
8	0	0.96	0.00	0.21
9	0	0.95	0.00	0.20
10	0.98	0.86	0.09	0.24
11	0.98	0.90	-0.05	0.12
12	0.98	0.83	-0.13	0.15
13	0.98	0.88	-0.13	-0.05
14	0.98	0.87	0.05	-0.07
15	0.98	0.93	0.05	-0.21
16	0.99	0.95	-0.06	0.06
17	0.99	0.92	0.10	0.28
18	0.84	0.59	-0.15	-0.02
19	0.94	0.77	-0.06	0.16
20	0.82	0.50	-0.33	0.15
21	0.85	0.44	-0.17	0.32
22	0.79	0.19	-0.27	0.17
23	0.93	0.75	-0.15	0.14
24	0.84	0.34	-0.20	0.23
25	0.95	0.82	-0.03	0.32
26	0.80	0.26	-0.25	0.15
27	0.95	0.76	-0.01	0.20
28	0.78	0.18	-0.17	0.28
29	0.51	0.55	0.39	0.29
30	0.61	0.78	0.33	0.27
31	0.38	0.14	0.50	0.29
32	0.21	0.19	0.45	0.33
33	0.59	0.64	0.50	0.33
34	0.72	0.73	0.48	0.45
35	0.14	0.08	0.48	0.42
36	0.23	0.25	0.46	0.37

37	0.25	0.21	0.57	0.28
38	0.29	0.15	0.51	0.28
39	0.59	0.62	0.52	0.34
40	0.19	0.16	0.54	0.22
41	0.23	0.15	0.65	0.48
42	0.19	0.08	0.48	0.28
43	0.21	0.06	0.53	0.27
44	0.17	0.14	0.41	0.38
45	0.46	0.42	0.75	0.51
46	0.24	0.12	0.57	0.41
47	0.35	0.18	0.68	0.39
48	0.36	0.27	0.70	0.49
49	0.31	0.26	0.57	0.53
50	0.37	0.32	0.71	0.54
51	0.36	0.26	0.72	0.49
52	0.17	0.12	0.57	0.43
53	0.29	0.21	0.73	0.45
54	0.13	0.11	0.49	0.39
55	0.37	0.24	0.66	0.48
56	0.22	0.15	0.67	0.50

The Dictation sub-test showed high performance for both groups, especially at the beginning where all items seemed to be easy for both groups. However, there seemed to be discrepancies between the two groups on these items, with most rural learners scoring these items more correctly than the urban group, while these items also presented with negative or very low item discrimination in this group. These discrepancies were demonstrated by items 10, 20, 21, 22, 23, 24, 25, 26, 27, and 28, on the item total discrimination. As this sub-test progresses, the difficulty level increases from item 28 to the end, but more rural learners still performed better on these items.

The differences in the performance of rural and urban isiXhosa learners may indicate problems in the equivalence of this sub-test, but it may also be that less emphasis is placed on the teaching of writing conventions and grammar in the urban setting as a result of the low status of isiXhosa in the urban areas as compared to the rural areas. However, the identified problems with this sub-test indicate the necessity for analysis of item bias across these groups. These analyses will be run for all sub-tests, as slight differences in items on all the sub-tests were detected.

4.5. ITEM BIAS /DIF

Research Aim 4: To evaluate the item bias or differential item functioning, across rural isiXhosa and urban isiXhosa learners of the isiXhosa version of the WMLS

4.5.1. PICTURE VOCABULARY

Table 14 presents the model assessment of the Stepwise Logistic Regression DIF procedure for the Picture Vocabulary sub-test for the two groups, while Table 15 presents the direction of DIF items for both these groups.

Table 14
Stepwise Logistic regression DIF procedure: Picture Vocabulary

Items	Stepwise R2			DIFF* square (2 degrees)	DIFF* R ²	Size of DIF
	Step 1	Step 2	Step 3			
1	1.00	1.00	1.00	00.00	0.00	No DIF
2	0.04	0.17	0.17	4.79	0.13	No DIF
3	0.24	0.27	0.27	2.04	0.03	No DIF
4	0.21	0.28	0.28	2.40	0.07	No DIF
	0.09	0.19	0.19	3.21	0.10	No DIF
6	0.14	0.22	0.22	1.82	0.08	No DIF
7	0.05	0.18	0.18	5.90	0.13	No DIF
8	0.07	0.08	0.08	0.39	0.01	No DIF
9	0.05	0.23	0.23	4.04	0.18	No DIF
12	0.07	0.15	0.15	1.06	0.08	No DIF
13	0.16	0.19	0.21	2.57	0.05	No DIF
14	0.23	0.26	0.26	6.45	0.03	No DIF
15	0.05	0.06	0.06	0.13	0.01	No DIF
16	0.15	0.24	0.24	2.71	0.09	No DIF
17	0.13	0.14	0.14	1.27	0.01	No DIF
18	0.15	0.15	0.15	0.61	0.00	No DIF
19	0.13	0.20	0.20	15.42	0.07	Large DIF
20	0.06	0.10	0.11	4.60	0.05	No DIF
21	0.21	0.26	0.26	13.40	0.05	Moderate DIF
22	0.17	0.17	0.18	1.07	0.01	No DIF
23	0.21	0.26	0.27	13.96	0.06	Moderate DIF
24	0.25	0.28	0.28	6.32	0.03	No DIF
25	0.22	0.23	0.23	1.78	0.01	No DIF
26	0.44	0.47	0.49	8.74	0.05	No DIF
27	0.47	0.48	0.49	4.58	0.02	No DIF
28	0.37	0.38	0.38	2.95	0.01	No DIF
29	0.21	0.30	0.32	16.96	0.11	Large DIF
30	0.36	0.41	0.43	12.12	0.07	Large DIF
31	0.16	0.19	0.19	2.51	0.03	No DIF
32	0.48	0.48	0.48	0.04	0.00	No DIF
33	0.21	0.33	0.34	21.58	0.13	Large DIF
34	0.22	0.22	0.22	1.22	0.00	No DIF

35	0.48	0.52	0.52	11.30	0.04	Moderate DIF
36	0.34	0.36	0.36	2.46	0.02	No DIF
37	0.32	0.52	0.52	7.96	0.20	No DIF
38	0.45	0.49	0.49	4.99	0.04	No DIF
39	0.46	0.46	0.47	0.40	0.01	No DIF
40	0.40	0.40	0.41	0.21	0.01	No DIF
41	1.00		1.00	0.00	0.00	No DIF
42	0.55	0.72	0.72	3.90	0.17	No DIF
43	0.21	0.32	0.32	1.41	0.11	No DIF
44	0.57	0.58	0.58	0.40	0.01	No DIF
45	0.54	.644	0.68	5.72	0.14	No DIF
46	0.11	0.25	0.24	3.03	0.13	No DIF
47	0.54	0.64	0.68	5.72	0.14	No DIF
48	0.26	.383	0.38	1.54	0.12	No DIF
49	0.02	0.14	0.14	1.44	0.12	No DIF
50	0.68	1.00	1.00	7.54	0.37	No DIF

*DIFF = difference

This table indicates that there were six items with DIF in the Picture Vocabulary sub-test, of which three (19, 29, and 33) had a large effect size of $R^2 \Delta > 0.06$ while the other three items (21, 23, and 35) displayed a moderate effect size of $0.035 < R^2 \Delta <= 0.060$. The null hypothesis of “no DIF” in these cases was thus rejected. Table 13 below shows the direction of the DIF of those items with a significant difference Chi-square (DIFF chi-square).

**UNIVERSITY of the
WESTERN CAPE**
Table 15
Direction of DIF: Picture Vocabulary

Item	Variables in equation	Beta	Direction
19	Total	0.18	Uniform
	RURAL_URBAN	-2.05	DIF Favours Urban
	Interaction term	0.04	isiXhosa learners
	Constant	-4.38	
21	Total	0.26	Uniform
	RURAL_URBAN	0.07	DIF Favours Rural
	Interaction term	0.04	isiXhosa learners
	Constant	-5.99	
23	Total	0.34	Non-Uniform
	RURAL_URBAN	4.73	DIF Favours HA Urban and LA
	Interaction term	-0.16	Rural
	Constant	-8.69	isiXhosa

29	Total	0.48	Non-Uniform DIF Favours
	RURAL_URBAN	8.80	HA Urban and LA Rural
	Interaction term	-0.26	
	Constant	-15.57	
33	Total	0.40	Non-Uniform DIF Favours
	RURAL_URBAN	-0.16	HA Urban and LA Rural
	Interaction term	-0.16	
	Constant	-13.36	
35	Total	0.45	Uniform DIF Favours
	RURAL_URBAN	-4.56	Urban isiXhosa learners
	Interaction term	0.12	
	Constant	-12.27	

*High ability Rural, Low ability Urban

The above analyses demonstrate that items 21 and 35 exhibited moderate uniform DIF. Item 21 favoured the rural isiXhosa learners while item 35 favoured urban isiXhosa learners. Item 23 from the moderate DIF items exhibited non-uniform DIF favouring high-ability urban test-takers and low-ability rural test-takers. The above table shows that only one of the large DIF items (19) exhibited uniform DIF which favoured urban isiXhosa learners. The other two large DIF items (29 and 33) displayed non-uniform DIF, with both items favouring high-ability urban learners and low-ability rural learners.

4.5.2. VERBAL ANALOGIES

Table 16 presents the model assessment of the Stepwise Logistic Regression DIF procedure for the Verbal Analogies sub-test for the two groups, while Table 17 presents the direction of DIF items for both these groups.

Table 16
Stepwise Logistic regression DIF procedure: Verbal Analogies subtest

Items	Stepwise R ²			DIFF* Chi square (2 degrees freedom)	DIFF* R ²	Size of DIF
	Step 1	Step 2	Step 3			
1	0.06		0.06	0.06	0.55	0.03 No DIF
2	0.24		0.25	0.25	1.54	0.01 No DIF
3	0.19		0.25	0.28	20.56	0.09 Large DIF
4	0.43		0.45	0.44	9.41	0.01 No DIF
5	0.06		0.10	0.14	4.88	0.08 No DIF
6	0.17		0.17	0.19	7.27	0.02 No DIF
7	0.04		0.12	0.12	16.74	0.08 Large DIF
8	0.07		0.07	0.11	8.29	0.04 No DIF
9	0.29		0.29	0.29	2.49	0.00 No DIF
10	0.45		0.45	0.46	3.37	0.01 No DIF
11	0.37		0.37	0.37	1.03	0.00 No DIF
14	0.24		0.27	0.27	6.33	0.03 No DIF
15	0.29		0.29	0.32	6.67	0.03 No DIF
16	.40		0.43	0.45	11.41	0.04 Moderate DIF
17	0.17		0.18	0.18	1.47	0.01 No DIF
19	0.39		0.40	0.40	0.33	0.01 No DIF
20	0.51		0.56	0.56	5.77	0.05 No DIF
21	0.27		0.28	0.29	2.03	0.02 No DIF
22	0.41		0.43	0.45	5.77	0.04 No DIF
23	0.64		0.68	0.69	3.16	0.05 No DIF
24	0.40		0.54	0.54	5.09	0.05 No DIF
25	0.50		0.51	0.62	6.54	0.12 No DIF
26	0.73		0.75	0.79	6.05	0.06 No DIF
27	0.47		0.47	0.49	1.76	0.02 No DIF
28	0.54		0.54	0.54	0.05	0.00 No DIF
29	0.36		0.36	0.37	0.38	0.01 No DIF
30	0.83		0.84	0.84	0.27	0.01 No DIF
31	1.00		1.00	1.00	-0.03	0.00 No DIF
32	1.00		1.00	1.00	0.00	0.00 No DIF
33	1.000		1.000	1.00	0.00	0.00 No DIF
34	.44		.45	0.51	2.14	0.07 No DIF

*DIFF = difference

This table indicate that there were three items with DIF, of which two items, 3 and 7, had a large effect size of $R^2 \Delta > 0.06$, while item 16 had a moderate effect size of $0.035 < R^2 \Delta \leq 0.060$. The null hypothesis of “no DIF” in these cases was thus rejected. Table 20 indicates the direction of the DIF of those items with a significant difference Chi-square (DIFF chi-square).

Table 17
Direction for DIF: Verbal Analogies subtest

Item	Variables in equation	Beta	Direction
3	Total	0.06	Uniform DIF Favours Urban isiXhosa learners
	RURAL_URBAN	-3.27	
	Interaction	0.22	
	Constant	1.04	
7	Total	0.10	Uniform DIF Favours Rural isiXhosa learners
	RURAL_URBAN	1.14	
	Interaction term	-0.00	
	Constant	-1.26	
16	Total	0.31	Non- Uniform DIF Favours HA Rural and LA Urban isiXhosa learners
	RURAL_URBAN	-3.65	
	Interaction term	0.21	
	Constant	-3.95	

*HA = High ability, LA = Low ability Urban

Of the items that displayed large DIF (3 and 7), item 7 displayed uniform DIF favouring rural learners and item 3, uniform DIF favouring urban learners. The moderate DIF item (16) displayed non-uniform DIF favouring high-ability rural learners and low-ability urban learners.

UNIVERSITY of the

4.5.3. LETTER-WORD-IDENTIFICATION RN CAPE

Table 18 presents the model assessment of the Stepwise Logistic Regression DIF procedure for the Letter-Word-Identification sub-test for the two groups, while Table 19 presents the direction of DIF items for both these groups.

Table 18
Stepwise Logistic regression DIF procedure: Letter-Word-Identification

Items	Stepwise R ²			DIFF*	DIFF* R ²	Size of DIF
	Step 1	Step 2	Step 3			
5	0.11	0.31	0.31	18.46	0.20	Large DIF
6	0.13	0.24	0.26	12.38	0.13	Large DIF
7	0.01	0.01	0.08	1.56	0.07	No DIF
8	0.13	0.27	0.27	8.61	0.14	No DIF
9	0.20	0.37	0.32	4.53	0.12	No DIF
10	0.25	0.39	0.39	8.66	0.14	No DIF
11	0.08	0.26	0.26	15.62	0.18	Large DIF
12	0.13	0.31	0.31	16.83	0.18	Large DIF
13	0.18	0.37	0.42	38.48	0.24	Large DIF
14	0.24	0.35	0.35	4.55	0.11	No DIF
15	0.15	0.26	0.28	10.90	0.13	Large DIF
16	0.24	0.34	0.34	3.19	0.10	No DIF
17	0.49	0.56	0.56	1.83	0.07	No DIF
18	0.00	0.10	0.10	1.32	0.10	No DIF
19	0.26	0.35	0.35	3.11	0.09	No DIF
20	0.08	0.20	0.24	21.25	0.16	Large DIF
21	0.16	0.34	0.39	38.61	0.23	Large DIF
22	0.49	0.56	0.56	1.59	0.07	No DIF
23	0.04	0.18	0.18	6.32	0.14	No DIF
24	1.00	1.00	1.00	0.00	0.00	No DIF
25	0.10	0.36	0.36	14.28	0.17	Large DIF
26	0.29	0.41	0.41	5.81	0.12	No DIF
27	0.14	0.23	0.24	8.89	0.10	No DIF
28	0.24	0.47	0.42	15.01	0.18	Large DIF
29	0.24	0.40	0.40	11.73	0.16	Large DIF
30	0.42	0.49	0.51	5.94	0.09	No DIF
31	0.31	0.32	0.38	8.15	0.07	No DIF
32	0.41	0.42	0.47	6.69	0.06	No DIF
33	0.46	0.46	0.46	0.47	0.00	No DIF
34	0.26	0.28	0.30	6.25	0.04	No DIF
35	0.36	0.37	0.38	6.40	0.02	No DIF
36	0.29	0.29	0.31	2.46	0.02	No DIF
37	0.38	0.40	0.42	6.03	0.04	No DIF
38	0.39	0.39	0.40	3.23	0.01	No DIF
39	0.34	0.35	0.35	1.14	0.01	No DIF
40	0.39	0.44	0.45	12.74	0.068	Large DIF
41	0.46	0.53	0.55	25.19	0.09	Large DIF
42	0.43	0.45	0.45	3.70	0.02	No DIF
43	0.44	0.51	0.51	12.09	0.07	Large DIF
44	0.47	0.48	0.48	2.86	0.01	No DIF
45	0.40	0.42	0.43	5.28	0.03	No DIF
46	0.43	0.44	0.45	3.52	0.02	No DIF
47	0.41	0.43	0.43	4.09	0.02	No DIF
48	0.27	0.27	0.27	1.54	0.01	No DIF
49	0.30	0.36	0.36	12.98	0.07	Large DIF
50	0.37	0.37	0.37	1.25	0.00	No DIF
51	0.35	0.35	0.35	1.71	0.00	No DIF
52	0.43	0.43	0.48	11.50	0.053	Negligible DIF
53	0.43	0.43	0.44	1.38	0.02	No DIF
54	0.42	0.42	0.44	5.75	0.02	No DIF
55	0.44	0.44	0.40	0.17	0.00	No DIF
56	0.41	0.41	0.42	0.19	0.01	No DIF
57	0.43	0.44	0.44	1.74	0.01	No DIF

*DIFF = difference

There were 16 items with DIF in the Letter-Word-Identification sub-test, of which 15 (5, 6, 11, 12, 13, 15, 20, 21, 25, 28, 29, 40, 41, 43, and 49) had a large effect size of $R^2 \Delta > 0.06$. The null hypothesis of “no DIF” in these cases was rejected. One item (52) had a DIF size that was negligible. Table 16 indicates the direction of the DIF of those items with a significant difference Chi-square (DIFF chi-square)

Table 19
Direction for DIF: Letter-Word-Recognition subtest

Item	Variables in equation	Beta	Direction
5	Total	0.10	Uniform DIF Favours Rural isiXhosa learners
	RURAL_URBAN	24.01	
	Interaction term	-0.10	
	Constant	-2.81	
6	Total	0.12	Non-Uniform DIF Favours HA Urban and LA Rural isiXhosa learners
	RURAL_URBAN	22.18	
	Interaction term	-0.38	
	Constant	-3.58	
11	Total	0.08	Uniform DIF Favours Rural isiXhosa learners
	RURAL_URBAN	22.80	
	Interaction term	-0.08	
	Constant	-1.60	
12	Total	0.11	Uniform DIF Favours Rural isiXhosa learners
	RURAL_URBAN	24.21	
	Interaction term	-0.11	
	Constant	-3.01	
13	Total	0.20	Non-Uniform DIF Favours HA Urban and LA Rural isiXhosa learners
	RURAL_URBAN	22.85	
	Interaction term	-0.39	
	Constant	-8.55	
15	Total	0.14	Non- Uniform DIF Favours both HA Urban and LA Rural isiXhosa learners
	RURAL_URBAN	16.40	
	Interaction term	-0.28	
	Constant	-4.10	
20	Total	0.11	Non-Uniform DIF Favours HA Urban and LA Rural isiXhosa learners
	RURAL_URBAN	18.88	
	Interaction term	-0.32	
	Constant	-3.82	
21	Total	0.18	Non-Uniform DIF Favours HA Urban and LA Rural isiXhosa learners
	RURAL_URBAN	17.92	
	Interaction term	-0.31	
	Constant	-8.03	

25	Total	0.15	Uniform DIF Favours Rural isiXhosa learners
	RURAL_URBAN	25.66	
	Interaction term	-0.15	
	Constant	-4.46	
28	Total	0.17	Uniform DIF Favours Rural isiXhosa learners
	RURAL_URBAN	27.01	
	Interaction term	-0.17	
	Constant	-5.80	
29	Total	0.16	Uniform DIF Favours Rural isiXhosa learners
	RURAL_URBAN	-26.23	
	Interaction term	0.16	
	Constant	-5.02	
40	Total	0.20	Non- Uniform DIF Favours HA Urban and LA Rural isiXhosa learners
	RURAL_URBAN	-6.84	
	Interaction term	0.13	
	Constant	-7.91	
41	Total	0.24	Non- Uniform DIF Favours HA Urban and LA Rural isiXhosa learners
	RURAL_URBAN	-12.46	
	Interaction term	0.22	
	Constant	-9.69	
43	Total	0.23	Uniform DIF Favours Urban isiXhosa learners
	RURAL_URBAN	-6.21	
	Interaction term	0.09	
	Constant	-7.52	
49	Total	0.21	Uniform DIF Favours Urban isiXhosa learners
	RURAL_URBAN	-2.59	
	Interaction term	0.02	
	Constant	-9.18	

*HA = High ability, LA = Low ability

Eight items displayed uniform DIF (5, 11, 12, 25, 28, 29, 43, and 49) and seven items (6, 13, 15, 20, 21, 40, and 41) displayed non-uniform DIF. Of the uniform DIF items, seven (5, 11, 12, 25, 28, 29, and 49) favoured rural learners and the other one DIF item (43) favoured urban learners. Of the non-uniform DIF items, all seven items (6, 13, 15, 20, 21, 40, and 41) favoured high-ability urban learners and low-ability rural learners.

4.5.4. DICTATION

Table 20 presents the model assessment of the Stepwise Logistic Regression DIF procedure for the Dictation sub-test for the two groups, while Table 21 presents the direction of DIF items for both these groups.

Table 20
Stepwise Logistic regression DIF procedure: Dictation subtest

Items	Stepwise R ²			DIFF* Chi Square (2 degrees freedom)	DIFF* R ²	Size of DIF
	Step 1	Step 2	Step 3			
2	0.19	0.21	0.21	0.35	0.02	No DIF
4	0.07	0.13	0.13	0.74	0.06	No DIF
7	0.01	0.11	0.11	1.16	0.09	No DIF
8	0.29	0.32	0.32	-13.23	0.04	No DIF
9	0.26	0.31	0.31	2.86	0.05	No DIF
10	0.20	0.22	0.22	2.40	0.02	No DIF
11	0.08	0.12	0.13	5.28	0.05	No DIF
12	0.13	0.20	0.23	14.21	0.11	Large DIF
13	0.08	0.15	0.18	10.19	0.10	Large DIF
14	0.11	0.16	0.16	6.65	0.05	No DIF
15	0.04	0.07	0.07	2.85	0.03	No DIF
16	0.08	0.11	0.13	3.18	0.05	No DIF
17	0.43	0.43	0.43	0.06	0.00	No DIF
18	0.03	0.11	0.13	18.21	0.10	Large DIF
19	0.15	0.18	0.20	9.66	0.05	No DIF
20	0.02	0.15	0.21	39.52	0.19	Large DIF
21	0.20	0.31	0.39	45.63	0.19	Large DIF
22	0.10	0.43	0.48	96.03	0.38	Large DIF
23	0.14	0.17	0.23	16.44	0.09	Large DIF
24	0.14	0.35	0.39	63.56	0.25	Large DIF
25	0.21	0.22	0.25	6.76	0.04	No DIF
26	0.07	0.36	0.39	77.37	0.32	Large DIF
27	0.14	0.18	0.19	8.94	0.05	No DIF
28	0.14	0.44	0.48	87.52	0.34	Large DIF
29	0.16	0.21	0.21	12.03	0.05	Moderate DIF
30	0.14	0.31	0.32	39.12	0.16	Large DIF
31	0.34	0.36	0.37	6.57	0.03	No DIF
32	0.23	0.25	0.26	6.21	0.03	No DIF
33	0.29	0.38	0.39	24.15	0.12	Large DIF
34	0.39	0.50	0.50	27.35	0.11	Large DIF
35	0.34	0.34	0.35	2.612	0.01	No DIF
36	0.23	0.27	0.28	11.73	0.05	Moderate DIF
37	0.33	0.36	0.38	9.66	0.05	No DIF

38	0.39	0.39	0.39	0.23	0.00	No DIF
39	0.32	0.41	0.41	22.26	0.09	Large DIF
40	0.31	0.34	0.36	8.95	0.05	No DIF
41	0.51	0.53	0.54	7.05	0.03	No DIF
42	0.39	0.39	0.39	0.00	0.00	No DIF
43	0.47	0.47	0.47	1.34	0.00	No DIF
44	0.36	0.39	0.39	7.85	0.03	No DIF
45	0.57	0.63	0.65	28.27	0.08	Large DIF
46	0.52	0.53	0.53	0.84	0.01	No DIF
47	0.55	0.55	0.56	2.66	0.01	No DIF
48	0.60	0.63	0.63	9.37	0.03	No DIF
49	0.51	0.55	0.56	12.70	0.05	Moderate DIF
50	0.60	0.66	0.66	19.07	0.06	Large DIF
51	0.60	0.63	0.63	9.13	0.03	Negligible DIF
52	0.50	0.54	0.54	6.73	0.04	No DIF
53	0.59	0.62	0.63	12.27	0.04	Moderate DIF
54	0.43	0.48	0.48	8.98	0.05	No DIF
55	0.57	0.58	0.58	3.30	0.01	No DIF
56	0.61	0.64	0.64	8.81	0.03	No DIF

*DIF = difference

There were twenty items with DIF in the Dictation sub-test, (12, 13, 18, 20, 21, 22, 23, 24, 26, 28, 29, 30, 33, 34, 36, 39, 45, 49, 50, and 53). Sixteen of these items had a large effect size (12, 13, 18, 20, 21, 22, 23, 24, 26, 28, 30, 33, 34, 39, 45, and 50), while four items (29, 36, 49, and 53) had a moderate effect size. The null hypothesis of “no DIF” in these cases was rejected. This sub-test also had one item in which DIF was negligible (item 45). Table 18 below indicates the direction of the DIF of those items with a significant difference Chi-square (DIF chi-square).

Table 21
Direction of DIF: Dictation subtest

Item	Variables in equation	Beta	Direction
12	Total	0.10	Non- Uniform DIF
	RURAL_URBAN	9.79	Favours both HA Urban and LA Rural
	Interaction term	-0.22	Rural isiXhosa learners
	Constant	-1.24	
13	Total	0.09	Non-Uniform DIF
	RURAL_URBAN	8.97	Favours HA Urban and La Rural
	Interaction term	-0.20	isiXhosa learners
	Constant	-0.43	
18	Total	0.04	Uniform DIF
	RURAL_URBAN	3.71	Favours Rural isiXhosa learners
	Interaction term	-0.07	
	Constant	-0.68	
20	Total	0.04	Non- Uniform
	RURAL_URBAN	6.42	Favours HA Urban and LA Rural
	Interaction term	-0.14	isiXhosa learners
	Constant	-1.23	
21	Total	0.17	Non-Uniform DIF
	RURAL_URBAN	8.53	Favours HA Urban and LA Rural
	Interaction term	-2.12	isiXhosa learners
	Constant	-5.14	
22	Total	0.10	Non- Uniform DIF
	RURAL_URBAN	8.46	Favours HA Urban and LA Rural
	Interaction term	-0.17	isiXhosa learners
	Constant	-4.58	
23	Total	0.14	Non- Uniform DIF
	Language group	7.49	Favours HA Urban and LA Rural
	Interaction term	-0.20	isiXhosa learners
	Constant	-2.85	
24	Total	0.10	Non- Uniform DIF
	RURAL_URBAN	7.56	Favours HA Urban and LA Rural
	Interaction term	-0.16	isiXhosa learners
	Constant	-3.81	
26	Total	0.05	Non- Uniform DIF
	RURAL_URBAN	6.58	Favours HA Urban and LA Rural
	Interaction term	-0.12	isiXhosa learners
	Constant	-2.75	
28	Total	0.12	Non- Uniform DIF
	RURAL_URBAN	7.81	Favours HA Urban and LA Rural
	Interaction term	-0.15	isiXhosa learners
	Constant	-5.24	
29	Total	0.10	Uniform DIF
	RURAL_URBAN	-2.49	Favours Urban
	Interaction term	0.05	isiXhosa learners
	Constant	-2.87	

30	Total	0.22	Non-Uniform Favours HA Urban
	RURAL_URBAN	0.38	and LA Rural
	Interaction term	-0.77	isiXhosa learners
	Constant	-4.62	
33	Total	0.19	Non- Uniform DIF Favours HA Rural
	RURAL_URBAN	-2.93	and LA Urban
	Interaction term	0.04	isiXhosa learners
	Constant	-4.94	
34	Total	0.34	Uniform DIF Favours Urban
	RURAL_URBAN	-2.35	isiXhosa learners
	Interaction term	0.01	
	Constant	-7.91	
36	Total	0.12	Non- Uniform DIF Favours HA Rural
	RURAL_URBAN	-3.72	and LA Urban
	Interaction term	0.07	isiXhosa learners
	Constant	-4.95	
39	Total	0.20	Uniform DIF Favours Urban
	RURAL_URBAN	-3.21	isiXhosa learners
	Interaction term	0.05	
	Constant	-5.25	
45	Total	0.28	Non- Uniform DIF Favours HA Rural
	RURAL_URBAN	-9.04	and LA Urban
	Interaction term	0.22	isiXhosa learners
	Constant	-8.58	
49	Total	0.37	Non-Uniform DIF Favours HA Urban
	RURAL_URBAN	2.87	and LA Rural
	Interaction term	-0.12	isiXhosa learners
	Constant	-12.81	
50	Total	0.37	Uniform DIF Favours Urban
	RURAL_URBAN	-1.83	isiXhosa learners
	Interaction term	0.00	
	Constant	-12.39	
53	Total	0.27	Non-Uniform DIF Favours HA Rural
	RURAL_URBAN	-7.95	and LA Urban
	Interaction term	0.18	isiXhosa learners
	Constant	-10.03	

*HA = High ability, LA = Low ability

Of the items with large DIF, four items displayed uniform DIF (18, 34, 39, and 50). One favoured the rural group (item 18) and the others, the urban group. The rest (12) of the large DIF items displayed non-uniform DIF (12, 13, 20, 21, 22, 23, 24, 26, 28, 30, 33, and 45). Nine of these items (13, 20, 21, 22, 23, 24, 26, 28, and 30) favoured high-ability urban learners and low-ability rural learners, while three

items (12, 33, and 45) favoured the high-ability rural learners and low-ability urban learners.

Four of the identified DIF items had moderate DIF (29, 36, 49, and 53). One item (29) displayed uniform DIF favouring urban learners, while the other three moderate DIF items displayed non-uniform DIF (36, 49, and 53), with two (36 and 53) favouring the high-ability rural learners and low-ability urban learners. The last of these items, item 49, favoured high-ability urban learners and low-ability rural learners.

4.5.5. SUMMARY OF DIF RESULTS

The table below is a summary of all the items which were identified as DIF items, and the groups that these items favoured.

Table 22
Summary results of the Logistic Regression method

Subtest	Effect size	Type of DIF	Direction of DIF favour	Number of items	Items
Picture Vocabulary	Large DIF	Uniform	Rural isiXhosa learners	0	0
			Urban isiXhosa learners	1	19
		Non-uniform	HA Rural LA Urban isiXhosa learners	0	0
			HA Urban LA Rural isiXhosa learners	2	29, 33
	Moderate DIF	Uniform	Rural isiXhosa learners	1	21
			Urban isiXhosa learners	1	35
		Non-uniform	HA Rural LA Urban isiXhosa learners	0	
			HA Urban LA Rural isiXhosa learners	1	23
Verbal Analogies	Large DIF	Uniform	Rural isiXhosa learners	1	7
			Urban isiXhosa learners	1	3
		Non-uniform	HA Rural LA Urban isiXhosa learners	0	0
			HA Urban LA Rural isiXhosa learners	0	0
	Moderate DIF	Uniform	Rural isiXhosa learners	0	0
			Urban isiXhosa learners	0	0
		Non-uniform	HA Rural LA Urban isiXhosa learners	1	16
			HA Urban LA Rural isiXhosa learners	0	

Subtest	Effect size	Type of DIF	Direction of DIF favour	Number of items	Items
Letter-Word-Recognition	Large DIF	Uniform	Rural isiXhosa learners	7	5, 11, 12, 25,28, 29,49
			Urban isiXhosa learners	1	43
	Non-Uniform	HA Rural LA Urban isiXhosa learners	0	0	
			HA Urban and LA Rural isiXhosa learners	7	6, 13, 15, 20, 21,40, 41
Dictation	Large DIF	Uniform	Rural isiXhosa learners	1	18
			Urban isiXhosa learners	3	34, 39,50
		Non-Uniform	HA Rural LA Urban isiXhosa learners	3	12, 33, 45
			HA Urban and LA Rural isiXhosa learners	9	13, 20, 21, 22, 23, 24, 26, 28,30
	Moderate DIF	Uniform	Rural isiXhosa learners	0	0
			Urban isiXhosa learners	1	29
		Non-Uniform	HA Rural LA Urban isiXhosa learners	2	36, 53
			HA Urban and LA Rural isiXhosa learners	1	49

In summary, for the first subtest, Picture Vocabulary, six items displayed DIF. Only three of these items displayed large DIF items, while the other three had moderate DIF effect sizes. Only one of these items favoured rural isiXhosa learners, while the other items favoured either urban isiXhosa learners, or high-ability urban isiXhosa learners and low-ability rural isiXhosa learners.

The second subtest (Verbal analogies) displayed two large DIF items and one moderate DIF item; from the large DIF items, one favoured the rural learners and the one favoured urban learners, while the moderate item favoured high-ability rural and low-ability urban learners.

In the third subtest, (Letter-Word-Identification), 15 large DIF items were identified. Of these items, seven favoured rural learners and one favoured urban learners. In addition, seven items favoured high-ability urban learners and low-ability rural learners.

Compared to the other subtests of the WMLS, the last sub-test (Dictation) had more DIF (20) items. From these items, sixteen were large DIF items, while four were

moderate DIF items. Three of these large DIF items favoured urban learners, while one favoured rural learners. Three items were identified as favouring the high-ability rural learners and low-ability urban learners, while nine items were identified as favouring the high-ability urban learners and low-ability rural learners.

The results of the logistic regression analysis for each of these subtests revealed an occurrence of DIF in the items identified in the above table. The reason for the occurrence of these DIF items between the rural and urban isiXhosa learners may be attributed to various factors, and these factors may differ for each of these subtests. These factors, the explanations about the occurrence of the DIF items in these subtests, and the implications which these DIF items may have on the overall interpretation of the test, will be discussed in detail in Chapter 5.



CHAPTER FIVE

DISCUSSIONS AND CONCLUSIONS

5.1. INTRODUCTION

The overall aim of this study was to evaluate the item bias of the isiXhosa version of the WMLS across rural and urban isiXhosa learners. Various analyses were undertaken in pursuance of this aim, namely the comparison of mean score differences, reliability, and item-bias analysis across rural and urban isiXhosa learners. These processes were also undertaken to evaluate, in the end, the scalar equivalence on the isiXhosa version of the WMLS across these two groups of learners. In this type of equivalence, as discussed in Chapter 2, full-score comparability of the test across these groups is of main concern (Van de Vijver & Rothmann, 2004), with item-bias analysis contributing to an understanding of scalar equivalence at an item level.

The results in Chapter 4 revealed mean-score group differences across rural and urban learners on at least two of the subtests of this test. This has thus led to further investigations, and an item-bias analysis was conducted on the different subtests of the WMLS. In all the subtests DIF items were identified. DIF was thus established as a contributing factor to the difference in the performance of these groups on this test. In this chapter, the possible factors contributing to the occurrence of the DIF items between the rural and urban learners will be discussed in detail, after a discussion of the results. The implications of the findings for the overall interpretation of results on this test will also be discussed.

5.2. SUMMARY AND DISCUSSIONS

The results of the analysis of specific aim 1 on evaluating mean-score differences on the WMLS subtests between rural and urban isiXhosa learners demonstrates that there were significant mean-score differences on two of the subtests of the WMLS (Verbal Analogies and Dictation sub-tests) across these groups, with the rural group

doing better than the urban group on the Dictation subtest, and the urban group scoring higher than the rural group on the Verbal Analogies sub-test. The results on Picture Vocabulary and Letter-Word-Identification sub-tests of the WMLS showed no mean differences between the two groups. The null hypotheses of no difference were not rejected for the Picture Vocabulary and Letter-Word-Identification subtest, while for Verbal Analogies and Dictation subtests of the WMLS the null hypotheses were rejected.

In an investigation into the psychometric characteristics of the subtests across groups, a reliability index analysis of the Cronbach's Alpha was undertaken. The results of this analysis indicated that the Picture Vocabulary, Letter-Word-Identification and Dictation subtests had good internal consistency for both groups, while the Verbal Analogies had an acceptable Cronbach's Alpha for the rural learners but not the urban learners. The null hypothesis which stipulates that "there are no differences between the Cronbach's Alpha of the two groups in the different subtests of the WMLS", was rejected for the Verbal Analogies subtest. The null hypothesis of no difference was not rejected for the Picture Vocabulary, Letter-Word-Identification and Dictation subtests.

The variation in the reliability index of the Verbal Analogies subtest across the two groups pointed in the direction of further investigation into the items of this subtest. Even though the equality of reliability results of the other subtests were promising, it remained important to conduct further analyses into the other subtests.

Analyses of the item difficulty and item discrimination indexes on these subtests for both these groups revealed several possible problematic items on all the subtests, even though the mean item characteristics were very similar across the two groups. There was more evidence of problematic items on the Letter-Word-Identification and Dictation subtests than on the Verbal Analogies and Picture Vocabulary subtests, especially with regard to differences on item discrimination values across groups, which is surprising, given the fact that the reliability coefficients of the two subtests

were at a satisfactory level for both groups and equal across the two groups (except for the urban group on Verbal Analogies).

The item bias analyses confirmed concerns about the scalar equivalence of all four subtests across the two groups. However, both the Picture Vocabulary and Verbal Analogies subtests displayed few DIF items (a small proportion of the overall number of items per sub-test), supporting tentatively the claim that mean-score differences on these subtests could be interpreted as real differences on the construct of interest. The identified DIF items on these subtests did not overlap with the items that were identified as possibly problematic with the individual item analyses (section 4.5), supporting the need to use efficient DIF identification methods in addition to item analyses in bias or fairness research. A methodologically interesting finding was that the DIF items that were identified, displayed differences across the two groups on item difficulty, but presented with similar discrimination values across the groups in these subtests. This was true even for the non-uniform DIF items (see for example PV29, PV33, and VA16).

The result of at least three non-uniform DIF items on the Picture Vocabulary subtest requires further investigation into the content of those items, as this may be an indication of different constructs being measured in the two groups in these items. This investigation falls beyond the scope of this study, however, and specific recommendations will be discussed later.

The result of only three DIF items on the Verbal Analogies subtest was surprising, given the differences in the reliability of this subtest across these two groups. This finding (the differences in reliability) also needs to be investigated further, possibly by means of exploratory factor analysis, to evaluate construct equivalence in these rural and urban isiXhosa learners.

The finding of a large number of DIF items on Letter-Word-Identification and Dictation was unexpected. The result supports the findings of the individual item analysis across the two groups, but was not reflected in differences in reliability. In

the case of the Letter-Word-Identification subtest, it also did not lead to mean-score group differences. However, differences were identified across the two groups on their mean scores of the Dictation subtest; furthermore the item discrimination values of both these subtests displayed differences across these groups which thus displayed differences in the interpretation of the construct being measured, for example item (LWI 20).

The differences between these two groups may have resulted from differences in the exposure of these groups to the construct being measured by the Letter-Word-Identification subtest. The DIF items of this test showed the rural group to be favoured by most of the uniform DIF items in this subtest, for example items (LWI 5, LWI 11, and LWI 12). In other words, this subtest identified urban isiXhosa learners as being the group with minimal exposure to the construct being measured by the Letter-Word-Identification subtest. Mbatha and Pluddemann (2004) reveal that second-language learners of isiXhosa in urban areas indicated that they wanted to learn isiXhosa, but they also indicated that they felt less enthusiastic about using isiXhosa inside the classroom. These learners were only exposed to isiXhosa as a subject inside the classroom and had minimal exposure in learning it in a social environment. Hendricks (2003) shares this opinion, saying that second- and third-language learners of isiXhosa, who are not socially exposed to this language, are less likely to be proficient in it. The consequences of language exposure for first-language speakers is, however, more severe. Goldsworthy (2001) notes that learners with poor exposure to phonological awareness skills will most likely struggle with reading and spelling. This is one of the challenges confronting isiXhosa learners in urban areas; the low status of isiXhosa further lessens the enthusiasm of learners to learn this language in both rural and urban areas dominated by isiXhosa. Generally the orthographic structure of the isiXhosa language is fairly easy, resulting in the high number of correct responses in the Letter-Word-Identification.

Bekker (2005) recognises some concerns regarding the isiXhosa language, namely that speakers of this language may be shifting to a stable form of bilingualism (English and isiXhosa) or that there could be a genuine shift from isiXhosa to

English, which is a status-dominant language in this region. Furthermore, Bekker (2005) acknowledges that this shift to English in a diverse society may be the result of this language belonging to a dominant group, which encourages assimilation. These factors may be additional aspects impacting on the findings of this study.

The interesting and unexpected finding in the Dictation subtests was that most of the uniform items favoured the urban children. However, even though that is true, the large number of items with non-uniform DIF points towards the need for further investigation into constructs bias and the nature of the construct in this subtest. The direction of the DIF in these non-uniform items will also tend to inflate the scores of low-ability rural learners, which may then lead to a misidentification of low ability or competence in these low-ability groups. These items will also lead to inflated scores in the high-ability urban groups.

The manner in which the different groups interpreted the different items can be identified as a possible contributing factor to these findings. Huysamen (2002) reveals that differences in the interpretation or understanding between different groups of test-takers on the same test may be identified as construct irrelevance. Huysamen (2002) adds that, in instances where language proficiency is assessed mainly in diverse contexts, test-takers' home languages cannot be ignored, as home language can be a potential source of the occurrence of construct irrelevance variation. Therefore it is advised that school learners should be assessed in their dominant language (Huysamen, 2002). As was discussed in Chapter 1, the isiXhosa language has various identified dialects, such as a currently identified urban dialect. However, only two of the dialects of isiXhosa form part of the standard isiXhosa language used in formal resources in isiXhosa such as books and articles, and in formal sectors such as the education sector. Educational assessment tools also use this standardised language, which is seldom spoken by urban learners in their social interaction and home environment. Furthermore, more of the deep rural isiXhosa learners are identified as using more of this standardised isiXhosa when compared to the urban isiXhosa learners (Bekker, 2005). Bekker (2005) also emphasises that the difference between rural and urban isiXhosa is vague, as more urbanisation has

occurred because of past apartheid laws and poverty in the rural areas. The state of development in the education system of both rural and urban schools has been slow, though rural school are still defined by poverty and minimal school resources. Slight mean-score differences were identified in the Dictation sub-test across the rural and urban learners.

According to Hambleton, Marenda, & Spielberger (2005), the problem of dialects within a language can become a threat to validity in an adapted test. To prevent this from happening, Hambleton et al. (2005) recommends that test users should choose which dialect is of interest, or whether the goal of testing is to produce an adaptation that will apply across dialects within a language. In the test used in this study, the scalar equivalence is of high priority, and thus all isiXhosa dialects should be included.

Things that may contribute to differences (Bekker, 2005):

- Members of the same culture group may vary widely in their degree of acculturation and proficiency in the language of the test;
- Familiarity with words and syntax in the native language;
- Educational background;
- Familiarity with tests and test-taking skills.

5.3. LIMITATIONS OF THE STUDY

- One of the identified limitations of this study was the sample size. As additional data was collected to equalise the sample for the purpose of the study, this data increased the sample size. The impact of this large sample on the findings of DIF can lead to an over-identification of DIF.
- The other limitation that can be identified in this study is in the distribution of the sample by grade, where there were more grade 7 learners in the urban group and more grade 6 learners in the rural group. These differences could have

contributed to the responses given by these learners in the different groups as they differed in their educational levels. This could also have also led to the recognition of variances in the difficulty level of either items or sub-tests across the different groups.

5.4. CONCLUSION

Based on the findings on the WMLS, one can conclude that the scalar equivalence of the test for the two groups has been approximately attained for some of the sub-tests, for example Picture Vocabulary and Verbal Ability in this instrument. Meanwhile the scalar equivalence of the Letter-Word-Identification and Dictation sub-tests is still questionable, based on the identified number of DIF items. However, recommendations will be provided to ensure that these tests attain their scalar equivalence. Therefore only the Picture Vocabulary and Verbal Analogies sub-tests can be tentatively used across the two groups, and the Letter-Word-Identification and Dictation sub-test need to be further investigated.

5.5. RECOMMENDATIONS

In the Letter-Word-Identification and Dictation sub-tests, further investigation needs to be undertaken on the content of those items which displayed differences in the response patterns of rural and urban learners, as this may be an indication of different constructs being measured in the two groups in these items. It is also recommended that DIF items in comparisons between rural and urban groups on Picture Vocabulary and Verbal Analogies be excluded, while a factor analysis of the Verbal Analogies subtest across the two groups should be conducted to explore reasons for the lower reliability of this test in the urban group.

REFERENCE LIST

- Alexander, N. (2005). *Language, class and power in post-apartheid South Africa: Project for alternative education*. Cape Town: University of Cape Town.
- Allalouf, A., Hambleton, R. K., & Sireci, S.G. (1999). Detecting the causes of differential item functioning in translated verbal items. *Journal of Educational Measurement*, 36, 185-198.
- American Psychological Association (APA). (2003). Ethical principles of psychologists and code of conduct. American Psychologist, Retrieved September 30, 2009, from <http://www.apa.org>.
- Bedell, B., Van Eeden, R., & Van Staden, F. (1999). *Culture as a moderator variable in psychological test performance: Issues and trends in South Africa*.
- Bekker, I. (2005). *Language Attitudes and Ethnolinguistic Identities in South Africa: A critical review*. Somerville, MA: Cascadilla Press 233-239.
- Calteaux, K. C. (1996). *Standard and non-standard language varieties in the urban South Africa*. Main Report for the STANON Research Programme. Pretoria, HRSC Publishers.
- Camara, W.J. (2007). *Standards for Education and Psychological Testing: Influence in Assessment Development & Use*. Unpublished Paper. November 8th.
- Carter, J.A., Lees, J.A., Murira, G.M., Gona, J., Neville, B.G.R., & Newton, R.J.C. (2005). Issues in the development of cross-cultural assessments of speech and language for children. *International Journal of Language & Communication Disorders*. 40(4), 385-401.

Classen, N.C.W. (1997). Cultural differences, politics and test bias in South Africa. *European Review of Applied Psychology*, 47(4), 297-307.

Classen. N.C.W., Krynauw, A.H., Peterson, H., & Mathe, M. (2001b). *A standardization of the WAIS-III for English-speaking South Africans*. Pretoria: Human Sciences Research Council.

Creswell, J.W. (2003). *Research Design. Qualitative, Quantitative, and Mixed Methods Approaches*. 2nd (Ed). London: SAGE Publications.

Cole, N.S., & Moss, P.M. (1993). *Bias in test use*. R.L. Linn (Ed.), Educational measurement (3rd ed.). Phoenix, AZ: Oryx.

De Klerk, G. (2008). *Cross-cultural testing*. In M. Born, C. D. Foxcroft, & R. Butter (Eds), Online Readings in Testing and Assessment, International Test.

Department of Labour (2009, July 14). Government Gazette. *Republic of South Africa*, 400, No. 19370. Cape Town, 2006.

De Vellis, R.F. (2003). Research design and methods: A process approach. Mountain View: Mayfield Publishing.

Fick, M.I. (1929). Intelligence test results of poor white, native (Zulu), coloured and Indian school children and the educational and social implications. *South African Journal of Science*, 26, 904-920.

Field, A. (2005). *Discovering statistics using SPSS*. London: Sage Publications.

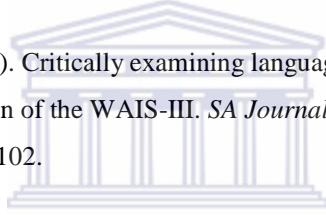
Finlayson, R., & Slabbert, S. (2003). What turns you on: An exploration of urban South African Xhosa and Zulu youth text. *Language, Culture and Curriculum*, 16(2), 165-172.

Foxcroft, C.D. (1997). Psychological testing in South Africa: Perspective regarding ethical and fair practices. *European Journal of Psychological Assessment*, 13 (3), 229-235.

Foxcroft, C. D. (2002a). Ethical issues related to psychological testing in Africa: What I have learned (so far). *Online Readings in Psychology and Culture* (Unit 5, Chapter 4),

(<http://www.ac.wvu.edu/~culture/index-cc.htm>)

Foxcroft, C.D., & Aston, S. (2006). Critically examining language bias in the South African adaptation of the WAIS-III. *SA Journal of Industrial Psychology*, 32 (4), 97-102.



Foxcroft, C.D., Paterson, H., Le Roux, N., & Herbst, D. (2004). *Psychological assessment in South Africa: A needs analysis. The test use patterns and needs of psychological assessment practitioners*. Final report. Pretoria: HSRC.

Foxcroft, C. D., & Roodt, G. (2001). *An introduction to psychological assessment in the South African context*. Cape Town: Oxford University Press.

Geisinger, K.F. (2000). Psychological testing at the end of the millennium: A brief historical review. *Professional Psychology: Research and Practice*, 31 (2), 117-118.

Gierl, M.J. (2000). Construct equivalence on translated achievement tests. *Canadian Journal of Education*, 25, 280-296.

Goldsworthy, C. (2001). *Comprehension of language*, (3rd Ed.). New York: Carrow-Woolfolk Publishing.

Graziano, A.M., & Raulin, M.L. (2000). *Research methods: A process of inquiry*. Boston: Allyn and Bacon.

Grieve, K.W. (2001). Factors affecting assessment results. In: Foxcroft, C. & Roodt, G. (Eds). *An introduction to psychological assessment in the South African context*. Cape Town: Oxford University Press.

Hambleton, R.K. (1994). Guidelines for adapting educational and psychological tests. *Bulletin of the International Test Commission*, 10, 229-244.

Hambleton, R.K., Merenda, P.F., & Spielberger, C.D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Lawrence Erlbaum. 26.

Harrison, M.G., Malaka, D., Amoateng, A.Y., & Toldson, I. (2005). Perceived competence of rural South-African and rural African-American families: A cross-cultural assessment of structural validity. *Ethnicity & Disease*, 15(3), 379–386.

Hau, K., & Chang, H. (2001). Item selection in computerized adaptive testing: should more discriminating items be used first?. *Journal of Educational Measurement* 38, 249-266.

Health Professions Council of South Africa. (1998). *Policy on the classification of psychometric measuring devices, instruments, methods and techniques*. Pretoria. HPCSA.

Hendricks, M. (2003). Institution for the study of English in South Africa: The school policy documents for isiXhosa L3 and L2. Pretoria: Department of Education.

Heugh, K. (2002). The Case Against Bilingual and Multilingual Education in South Africa: Laying Bare the Myths. *Perspectives in Education*, 20(1), 171-196.

Heugh, K., Diedericks, G.A.M., Herbst, D.L., & Winnaar, L. (2007). *Assessment of the language and mathematics skills of grade 8 learners in the Western Cape in 2006*. Pretoria: Human Sciences Research Council.

Horne, T.J. (2001). Education and language transferees. *The Education African Forum*, 5. In press.

Huysamen, G.K. (2002). The relevance of the new APA standards for educational and psychological testing in employment testing in South Africa. *South African Journal of Psychology*, 32(3) 26-33.

International Testing Commission (ITC). (2001). International guidelines for test use. *International Journal of Testing*, 1, 93-114.

Kanjee, A. (2007). *Using logistic regression to detect bias when multiple groups are tested*. Pretoria: Human Sciences Research Council.

Kaplan, R.M., & Saccuzzo, D.P. (1997). *Psychological Testing: Principles, Applications, and Issues*. New York: Brooks/Cole Publishing Company.

Kerswill, P. In Mattheier, K., Ammon, U. & Trudgill, P. (2006). Migration and language. *Sociolinguistics/Soziolinguistik. An international handbook of*

the science of language and society, (2nd Eds), 3. Berlin: De Gruyter.

Koch, E. (2005). *Evaluating the equivalence, across language groups, of a reading comprehension test used for admissions purposes*. Unpublished doctoral thesis, Port Elizabeth: Nelson Mandela Metropolitan University.

Koch, E. (2007a). Die evaluering van ‘n eentalige toelatingstoets wat vir toelating tot hoër onderwys in ‘n meertalige konteks gebruik word. *SA Joernaal vir Industriële Sielkunde*, 33 (1), 90-101.

Koch, E. (2007b). *The monolingual testing of competence: acceptable practice of unfair exclusion*. Cuvelier, P., Du Plessis, T., Meeuwis, M. & Teck, L. (Eds). 2007. Multilingualism and exclusion. Policy, practice and prospects. Pretoria: Van Schaik Publishers.79-103.

Koch, E. (2009). The case of bilingual tests: a study of test adaptation and analysis. *South African Linguistics and Applied language Studies*, 27(3). 301-317.

Kunnan, A.J. (2000). *Studies in Language Testing: Fairness and validation in language assessments*. New York: Press Syndicate of the University of Cambridge,

Mantsha, T.R. (2002). *Psychometric properties of a Venda version of the Sixteen Personality Factor questionnaire (16PF) fifth edition*.MA Thesis. Pretoria: University of South Africa.

Manyike, T.V. (2007). *A comparative assessment of L1 and L2 writing, performance among grade 7 learners in English and Xitsonga*. Unpublished doctoral thesis. Pretoria: University of South Africa.

Mauer,K.F. (2000). *Psychological test use in South Africa*. Department of Industrial Psychology: University of South Africa.

Mbatha, T., & Pluddemann, P. (2004). The status of isiXhosa as an additional language in selected Cape Town secondary schools. *Occasional Papers. No. 18*. PRAESA. Rondebosch.

McGorry, S.Y. (2000). Measurement in a cross-cultural environment: survey translation issues. *Qualitative Market Research: An International Journal*. 3 (2).74-81.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.) Washington, DC: American Council on Education; New York: Macmillan.

Meiring, D. (2007). *Bias and equivalence of psychological measures in South Africa*. Netherlands: Labyrinth Publishing.

Mungazi, D.A., & Walker, L.K. (1997). *Educational reform and the transformation of Southern Africa*. Westport: Praeger Publisher.

Ntantiso, M. (2009). Unpublished M.A. Thesis, Port Elizabeth: Nelson Mandela Metropolitan University.

Owen, K. (1991). Test and item bias: The validity of the Junior Aptitude Tests (JAT) for various pupils in South Africa regarding constructs measured. *South African Journal of Psychology*, 21, 112- 118.

Owen, K. (1998). *The role of psychological tests in education in South Africa: Issues, controversies and benefits*. Pretoria: Human Sciences Research Council.

Payne, K.T., & Taylor, O.L. (2002). Multicultural influences on human communication. In G. H. Shames and N. B. Anderson (Eds), *Human Communication Disorders. An Introduction*, 6th (ed) Boston: Allyn & Bacon

Psychological Society of South Africa (PsySSA). (2004). *Code of conduct*. Retrieved 30 June, 2004 from www.psyssa.com

Robin, F., Sireci, S. G., & Hambleton, R.K. (2003). Evaluating the equivalence of different language versions of a credential exam. *International Journal of Testing*, 3 (1), 1-20.

Rosselli, M., & Ardila, A. (2003). Acalculia. *Med link Neurology*. www.medlink.com.

Sayed, Y. (1998). *The Segregated information highways: An analysis of information literacy in Higher Education*. London: University of Cape Town.

Schaap, P., & Basson, J.S. (2003). The construct equivalence of the PIB/SPEEX motivation index for job applicants from diverse cultural backgrounds. *SA Journal of Industrial Psychology*, 29(2), 49-59.

Schaap, P., & Vermeulen, T. (2008). The construct equivalence and item bias of the PIB/SPEEX conceptualisation-ability test for member of members of five language groups in South Africa. *SA Journal of Industrial Psychological*, 34(3), 29-38.

Schultz, K.S., & Whitney, D.J (2005). *Measurement theory in action*. Oaks, CA: Sage Publications.

Sireci, S.G., & Allalouf, A. (2003). Appraising item equivalence across multiple languages and cultures. *Language Testing*, 20 (2), 148-166.

Statistics South Africa (2007). *The people of South Africa population census: census in brief*. Pretoria: Statistics South Africa.

<http://www.southafrica.info/about/people/population.htm>.

Taylor, T.R. (1987). Test bias: *The role and responsibilities of test user and test publisher*. Johannesburg: National Institute for Personnel Research.

Tyler, L.E., & Walsh, W.B. (1979). *Tests and measurements* (3rd ed). Prentice-Hall INC, New Jersey.

Uiterwijk, H., & Vallen, T. (2006). Linguistic sources of item bias for second generation immigrants in Dutch tests: *Language Testing*. 22(2). 211-234.

Van de Vijver, F.J.R. (1998). Towards a Theory of Bias and Equivalence. *ZUMA-Nachrichten Spezial*; Retrieved September 30.

Van de Vijver, F.J.R., & Lueng, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks: Sage Publication.

Van de Vijver,F.J.R., & Rothmann, S. (2004). Assessment in multicultural groups: The South African case. *S.A. Journal of Industrial Psychology*. 30 (4),1-7.

Van de Vijver, F.J.R., & Tanzer, N.K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*. (in press).

Van Zyl, T. (1996). *HSRC test catalogue, Part 1: and proficiency tests*. Pretoria: HSRC.

Walsh, W.B., & Beltz, N.E. (2000). *Test and assessment* (4th ed). Prentice Hall. New Jersey.

Webb, V. (2001). Language in South Africa: *The role of language in national transformation, reconstruction and development*. Pretoria: University of Pretoria.

Woodcock, R.W., & Sandoval,A.F. (2005). *Woodcock-Munoz Language Survey: Normative Update*. Itasca, IL:River Side Publishing. Revised.

Zumbo, B.D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modelling as a unitary framework for binary and Likert type (ordinal) item score*. Ottawa: Directorate of Human Resources Research and Evaluation, Department of National Defense. Retrieved August 12, 2009, from <http://www.educ.ubc.ca/faculty/zumbo/DIF/handbook.pdf>

APPENDIX A
Description of test (WMLS)

TEST	TEST REQUIREMENTS	MEASURES	RESPONSE STYLE
Picture Vocabulary	Test takers name familiar and unfamiliar pictured objects that involve easy and difficult school related knowledge and experience.	Oral language, including, language development and lexical knowledge.	Oral (word)
Verbal Analogies	Test takers complete oral analogies requiring verbal comprehension and reasoning.	Verbal reasoning using lexical knowledge.	Oral (word)
Letter-Word Identification	Test takers reads out familiar and unfamiliar letters and words. Which differ in difficulty level	Letter-Word Identification skills.	Oral (letter, word, name)
Dictation	Test takers respond to questions by writing verbal comprehension, knowledge of letter forms, spelling, punctuation, capitalisation, and word usage.	Prewriting Skills (for early items), Ability to respond in writing to a variety of questions and grammar.	Motoric (Writing)



UNIVERSITY *of the* WESTERN CAPE

DEPARTMENT OF PSYCHOLOGY

Private Bag X 17, Bellville 7535, South Africa, Telephone:(021) 959-2283/2453
Fax: (021) 959-3515 Telex: 52 6661

APPENDIX B

INFORMATION SHEET

Project Title:

An evaluation of group differences and items bias, across urban and rural isiXhosa learners on the isiXhosa version *Woodcock Muñoz Language Survey* (WMLS)

What is this study about?

This study stems from a bigger project, namely, the Additive Bilingual Education project that aims at the promotion of additive bilingual education through the medium of both English and isiXhosa languages which forms part of the University of the Western Cape with Prof Elize Koch as a senior researcher. This study aims to evaluate the isiXhosa version of the WMLS which is used to assess language proficiency, on the bias, across rural and urban isiXhosa learners on the subtests of this version. We are inviting you to participate in this research project because you would understand the language of the test and respond to the questions asked on the test. The principal at your school also identified you, together with a teacher, as a possible participant in this research project because of your gender and that you are a grade 6 or 7 learner as was requested in this study. Your parents consent to take part in this study will be asked, including your consent. Under no circumstances will you be forced to participate in this study.

What will I be asked to do if I agree to participate?

You will be asked to answer some questions on the test. The test is a language test and you will be given questions related to vocabulary, analogies, reading and dictation. You will be tested by a Psychology student from the University of the Western Cape. You will be tested individually in a classroom or venue at the school. This student will explain everything you need to do, and will write down your answers to the questions. The testing will take place at your school during school hours, with the permission of the Department of Education, Eastern Cape and the principal of your school, and will take about 50 minutes to complete. The kinds of questions that will be asked are, for example, to show you a picture of a cat and then to ask you what it is. Some of the questions will be easy and some of the questions will be more difficult. In some of the questions you will be requested to read a word, in others to complete a sentence, or to write a word or a

sentence. You must please do your best on the test as the research will only be successful if we get the best possible answers from each child. Participation in the research is NOT a requirement of the school.

Would my participation in this study be kept confidential?

We will do our best to keep your personal information confidential. To help protect your confidentiality, your test answer sheet will be assigned a number. The number will be used to store the results of the test on the computer. Your name will not be used anywhere on any computer records of the research or in any reporting of the results. The researcher, Prof Koch, and her research assistants will be the only people who will have access to your results. Your results on the test will not be made available to anybody. The only personal information that we will store is your age, your first language and your gender. We will also record the name of your school. If we write a report or article about this research project, your identity and the identity of the school will be protected to the maximum extent possible.

What are the risks of this research?

There are no known risks associated with participating in this research project. We are not doing research on you as a person, but on the test. You are being tested so that we could collect data in order to do the research on the test.

What are the benefits of this research?

This research is not designed to help you personally, but the results may help the investigator learn more about the test and whether this test can actually be used to improve learning in the schools of SA. We hope that, in the future, other people might benefit from this study through improved understanding of important aspects of language that need to be tested, and how.

Do I have to be in this research and may I stop participating at any time?

Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may withdraw at any time of the study. If you decide not to participate in this study or if you stop participating at any time, you will not be blamed or punished.

Is any assistance available if I am negatively affected by participating in this study?

You will not be negatively affected by this research. However, if you feel that you are, you can contact Prof Elize Koch who will do everything possible to refer you for support and assistance.

What if I have questions?

This research is being conducted by Prof Elize Koch at the University of the Western Cape. If you have any questions about the research study itself, please contact Prof Elize Koch at: The University of the Western Cape, 021-9592482, skoch@uwc.ac.za

Should you have any questions regarding this study and your rights as a research participant or if you wish to report any problems you have experienced related to the study, please contact:

Head of Department: Prof K. Mwaba

Dean of the Faculty of Community and Health Sciences: Prof R. Mpofu

University of the Western Cape

Private Bag X17

Bellville 7535

This research has been approved by the University of the Western Cape's Senate Research Committee and Ethics Committee.





UNIVERSITY *of the* WESTERN CAPE

DEPARTMENT OF PSYCHOLOGY

Private Bag X 17, Bellville 7535, South Africa, Telephone:(021) 959-2283/2453
Fax: (021) 959-3515 Telex: 52 6661

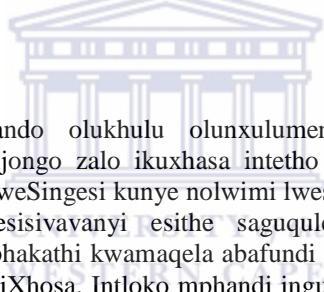
APPENDIX B

Incwadi yesazizi

Isihloko sophando:

Uhlalutyo kokwahlukana kwamaqela, kunye nokhetho lwee items ezithe zafumaneka zino khetho kumaqela abafundi besiXhosa basezilalini kunye nabo basezi dolophini kwi *Woodcock Muñoz Language Survey* (WMLS).

Lungantoni oluphando?



Oluphando luyinxalenye yophando olukhulu olunxulumene ne (Additive Bilingual Education Project), iinjongo zalo ikuxhasa intetho zolwimi ezibhini ngokufanelekileyo iquka ulwimi lweSingesi kunye nolwimi lwesiXhosa. Injongo zoluphando kukuhlalutywa kwesisivavanyi esithe saguqulelwa esiXhoseni, kujongwa ukuba asinakhetho na phakathi kwamaqela abafundi basezilalini kunye nabo basezidolophini abathetha isiXhosa. Intloko mphandi ingu Prof Elize Koch, oluphando luqhutywa kwiDyunesiti yase Tshona Koloni. Simema wena ubeyinxalenye yoluphando ngokuba uthetha ulwimi oluphando luyakuthi luqhutywa ngalo kwaye uzokwazi ukuphendula nemibuzo koluvavanyo. Inqununu iquka nomfundisitsapho wesikolo sakho bakukhethile njengomthathi nxaxheba ofanelekileyo koluphando, ngenxa yesini sakho kunye nebanga ofunda kulo grade 6-7 njengokufanelekileyo koluphando. Imvume yokuba uthathe inxaxheba iyakucelwa emzalini wakha, nawe uyakuthi ucelwe imvume yokuthatha inxaxheba. Akuyakubakho mntu onokunyanzelisa ukuba uthathe inxaxheba koluphando.

Yintomi ekufuneke ndiyeenzile ukuba ndivumile ukuthatha inxaxheba?

Uyakuthi ucelwe ukuba uphendule imibuzo koluvavanyo. Oluluvavanyo lolwimi lwesiXhosa kwaye uyakunikwa imibuzo enqamelana namagama, analogie, ukufunda kunye nezibizelo. Uzokuvavanywa ngumfundsi owenza iPsychology ephuma kwiDyuniveti yase Tshona Koloni. Uvavanyo olu luyakwenzelwa bucala kwigumbi lokufundela kwelinje lamagumbi esikolweni sakho. Lomfundi

uyakucacisa konke ekufanele ukwenzile, kwaye uyakubhala pantsi iimpendulo zakho zalemibuzo. Oluvavanyo luyakuqhutyelwa esikoloni sakho ngexesha lesikolo, ngemvume ezsuka kwi Department of Education, Eastern Cape kunye nenqununu yesikolo, oluvavanyo luyakuthatha imizuzu eyi55 ukuze luqghitywe. Umzekelowemibuzo enokuthi uyibuzwe, uyakuboniswa umfanekiso wekati wena uyakuchaza ukuba yintoni. Eminye yalemibuzo izakubanobulala eminya izokuba nobinzima. Kweminye yale mibuzo uzofuneka ufunde igama, kweminye uzokuqqhibezela isivakalisi okanye ubhale igama okanya isivakalisi. Sicela uphendule kangangoko unako ukuze oluphando lubeyimpumelelo umntana ngamye funeka aphendule kangangoko anako. Ukuthatha inxaxheba koluphando ayisosinyanzekiso esiphuma kwisikolo sakho.

Ukuthatha kwam inxaxheba koluphando luyakugcinwa luyimfihlelo na?

Siyakuthi senze ekusemandleni wethu ukugcina ulwazisi lakho lufihlakele. Ukunceda ufhla isazisi sakho sifhlakele, incwa yakho yokuphendula koluvavanyo luyakunikwa inani. Elinani liyakuthi lisetyenzisewa ukugcina iziphumo zoluphando kwikhompyutha. Igama lakho aliyikusetyenziswa nakweyiphi ikhompyutha koluphando okanye kwiziphumo zophando. Intloko mphandi uProf Koch, kunye nabaphandi bakhe koluphando bayakuthi ibengabo bodwa abanokufikelela kwezizophumo. Akekho omnye umntu onokufikela kwezizophumo. Ulwazi oluyakuthi lugcinwe ngawe yiminyaka yakho, ulwimi olwimi olusebenzisayo kunye nesinyi sakho. Igama lesikolosakho naso siyakuthi sigcinwe. Xa kubhalwa iziphumo zoluphando akukho nenyi indlela onokuveza isazisi sakho okanye eso sesikolo sakho. to the maximum extent possible.

WESTERN CAPE

Zeziphi iingqxaki endinohlangana nazo koluphando

Akukho zingxaki endinokuthi ndihlangana nazo njengomthathi nxaxheba Koluphando. Oluphando alwenziwa kuwe kodwa kuvavanyo ubuqu. You are being tested so that we could collect data in order to do the research on the test.

Yeyiphi imivuzo endinokuthi ndiyifumane koluphando?

Iimivuzo yoluphando ayenzelwanga ukunceda umntu omnye., kodwa ezizophumo ziayakuthi zizame ukunceda abaphandi bafunde lukhulu ngessi vavanyi sizetyenziswa koluphando, kunye oluphando luwa kuthi lincede kuhlumo lwemfundo Yase Mzantsi Afrika . Siyathemba ukuba banintsi abantu abanokuthi bazuze koluphando kwiminyaka ezayo, kwaye bafunde lukhulu ngezolwimi..

Kunyanzelekile ukuba ndithathe inxaxheba koluphando, kwaye ndingarhoxa nanyaliphi na ithuba?

Ukuthatha inxaxheba koluphando ayisosinyanzeliso. Ukuba uthathe isiqgibo sokurhoxa koluphando ungakwenza oko nangaliphina ithuba, awuyukubuzwa mibuzo okanye utholwe.

Ikhona inxaso endinokuthi ndiyifumane ukuba ndichaphazeleke kakubi ngokuthatha inxaxheba koluphando?

Awuyikuthi achaphazeleke kakubi koluphando. Kodwa ukuba oko kwenzekile ungathi uqhagamishelane no, Prof Elize Koch oyakuthi enze konke okusemandleni ukuba ufumane uncedo.

Ukuba ndinemibuzondenza njani?

Oluphando lweziwa ngu Prof Elize Koch *ekwi Dyunivesithi yase Tshonakoloni*. Ukuba unemibuzo, ungatsalela kut Prof Elize Koch *kwiDyunivesithi yaseTshona Koloni*, 021-9592482, skoch@uwc.ac.za

Ukuba uneminye imibuzo ngoluphando nceda utsalele intloko mphamndi, okanye ufunu ukuchaza ingxaki othe wayifumana koluphando, nceda utsalele :

Head of Department: Prof K. Mwaba

Dean of the Faculty of Community and Health Sciences: Prof R. Mpofu

University of the Western Cape

Private Bag X17

Bellville 7535

This research has been





UNIVERSITY *of the* WESTERN CAPE

DEPARTMENT OF PSYCHOLOGY

Private Bag X 17, Bellville 7535, South Africa, Telephone: (021) 959-2283/2453
 Fax: (021) 959-3515 Telex: 52 6661

APPENDIX C

ASSENT FORM

Title of Research: An evaluation of group differences and items bias, across urban and rural isiXhosa learners on the isiXhosa version of the *Woodcock Muñoz Language Survey* (WMLS)

The study has been described to me in a language that I understand and I freely and voluntarily agree to participate. My questions about the study have been answered. I understand that my identity will not be disclosed and that I may withdraw from the study without giving a reason at any time and this will not negatively affect me in any way.

Participant's name.....

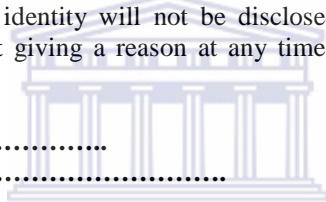
Participant's signature.....

Date.....

Witness' name:.....

Witness' signature:

Date:



Should you have any questions regarding this study or wish to report any problems you have experienced related to the study, please contact the study coordinator:

Study Coordinator's Name: Prof Elize Koch

University of the Western Cape

Private Bag X17, Belville 7535

Telephone: (021) 959-2842

Cell: 0824439311

Email: skoch@uwc.ac.za



UNIVERSITY *of the* WESTERN CAPE

DEPARTMENT OF PSYCHOLOGY

Private Bag X 17, Bellville 7535, South Africa, Telephone: (021) 959-2283/2453
Fax: (021) 959-3515 Telex: 52 6661

APPENDIX C

INCWADI YESIVUMELWANO (MZALI)

Isihloko sophando: Uhlalutyo kokwahlukana kwamaqela, kunye nokhetho lwee items ezithe zafumaneka zino khetho kumaqela abafundi besiXhosa basezilalini kunye nabo basezi dolophini kwi *Woodcock Muñoz Language Survey* (WMLS) yesiXhosa.

Oluphando lucacisiwe kum ngolwimi oluvakalayo kwaye ndiyavumelana nokuthatha inxaxheba ngokukhululekileyo. Imibuzo endinayo malunga noluphando iphendulekile. Ndiyaqonda ukuba akukho ndlela inokundiveza ukuba ndingubani kwaye nokuba ndinga rhoxa nangaliphi na ixesha koluphando, ngaphandle kwesizathu kwaye akuyi kundichaphazel a kakubi.

Igama lomzali	Utyikityo lomzali
Umhla	
Igama lengqina	
Utyikittyo lengqina	
Umhla	

Ukuba uthe wanayo imibuzo malunga noluphando okanye ufunu ukwazisa ngengxaki othe wahlangana nazo koluphando, tsalela umququzeleli loluphando :

Igama loomquluquzeleli wophando: Prof Elize Koch
University of the Western Cape
Private Bag X17, Belville 7535
Telephone: (021) 959-2842
Cell: 0824439311
Email: skoch@uwc.ac.za



UNIVERSITY *of the* WESTERN CAPE

DEPARTMENT OF PSYCHOLOGY

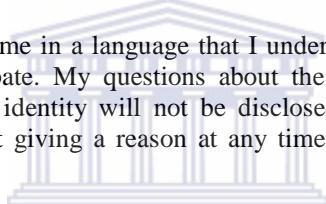
Private Bag X 17, Bellville 7535, South Africa, Telephone: (021) 959-2283/2453
Fax: (021) 959-3515 Telex: 52 6661

APPENDIX D

CONSENT FORM

Title of Research: An evaluation of group differences and items bias, across rural and urban isiXhosa learners on the isiXhosa version of the *Woodcock Muñoz Language Survey* (WMLS)

The study has been described to me in a language that I understand and I freely and voluntarily agree to participate. My questions about the study have been answered. I understand that my identity will not be disclosed and that I may withdraw from the study without giving a reason at any time and this will not negatively affect me in any way.



Parents name:.....

Parents signature:.....

Date:

Witness' name:.....

Witness' signature:

Date:

Should you have any questions regarding this study or wish to report any problems you have experienced related to the study, please contact the study coordinator:

Study Coordinator's Name: Prof Elize Koch

University of the Western Cape

Private Bag X17, Belville 7535

Telephone: (021) 959-2842

Cell: 0824439311

Email: skoch@uwc.ac.za

UNIVERSITY *of the* WESTERN CAPE



DEPARTMENT OF PSYCHOLOGY

Private Bag X 17, Bellville 7535, South Africa, Telephone: (021) 959-2283/2453
Fax: (021) 959-3515 Telex: 52 6661

APPENDIX D

INCWADI YESIVUMELWANO (YOMFUNDI)

Isihloko sophando: Uhlalutyo kokwahlukana kwamaqela, kunye nokhetho lwee items ezithe zafumaneka zino khetho kumaqela abafundi besiXhosa basezilalini kunye nabo basezi dolophini kwi *Woodcock Muñoz Language Survey* (WMLS) yesi Xhosa.

Oluphando lucacisiwe kum ngolwimi oluvakalayo ndaye ndiyavumelana ngokuthatha inxaxheba ngokukhululekileyo. Imibuzo endinayo malunga noluphando iphendulekile. Ndiyaqonda ukuba akukho ndlela inokuveza isazisi sam, nokuba ndinga roxa nangaliphi na ixesha koluphando, ngaphandle kwesizathu kwayi kundichaphazelka kakubi.

Igama lomthathi nxaxheba	WESTERN CAPE
Utyikityo lomthathi nxaxheba	
Umhla	
Igama lengqina	
Utyikittyo lengqina	
Umhla	

Ukuba uthe wanayo imibuzo malunga noluphando okanye ufunu ukwazisa ngengxaki othe wahlangana nazo koluphando, tsalela umquuzeleli loluphando :

Igama lomqluquzeleli wophando: Prof Elize Koch
University of the Western Cape
Private Bag X17, Belville 7535
Telephone: (021) 959-2842
Cell: 0824439311
Email: skoch@uwc.ac.z

APPENDIX E

Ethics statement:

The following steps are taken during this research in accordance with the International Ethical Principles of Psychologists and Code of Conduct (2002), the SA Code of Conduct for Psychologists (www.psyssa.com) and the UWC, Faculty of CHS's Policies and Procedures for Ethical Review of Research Projects (May, 2003):

Secondary Date:

1. Permission to conduct this study with grade 6 and 7 learners both males and females was requested from the Department of Education in the Eastern Cape.
2. The principals of the schools in which the data was collected from for this study were asked permission to conduct this study in their schools. They were under no circumstances forced to take part in this study.
3. The isiXhosa version of the information sheet, informed form and assent forms was given to both the participant and the parent.
4. Participants were informed about the purpose of the study, and that participation is voluntary and they were informed that under no condition will their identity be reflected by their results and that these results can only be used in for the purpose of the bigger project.
5. Clear instructions were given to participants before the beginning of each test in a language that they understand.
6. Results of this test were stored in a safe place in outlook for the confidentiality of results.

Primary data

1. The same procedures were undertaken as that of the ones of the secondary data which was used in this study.
2. In this study the ethical procedures was followed as required from the Ethical Board in UWC.

APPENDIX F**VERIFICATION**

MRS HELEN ALLEN, APED
(Accredited Professional Text Editor, SATI)

Date: 4 December 2009

90 Buitekant Street
 SWELLENDAM
 6740

ID 3101260009082
 Tel. 028 514 2006
 e-mail helenallen@mweb.co.za

TO WHOM IT MAY CONCERN:



I am a professional text editor, accredited with the South African Translators' Institute (SATI).

On 4 December 2009, I completed the editing of an academic thesis written by

Unathi Lucia Silo

Titled

An evaluation of group differences and items bias, across rural isiXhosa learners and urban isiXhosa learners, of the isiXhosa version of the Woodcock Muñoz Language Survey (WMLS).

HP Allen

A historical overview of the study of family business as an evolving field.

Helen Allen