IMPACT OF LINGUISTIC COMPLEXITY OF TEST ITEMS ON PUPILS' PERFORMANCE: THE CASE OF THE MALAWI PRIMARY SCHOOL LEAVING CERTIFICATE EXAMINATION MATHEMATICS WORD PROBLEMS

M.Ed (Testing, Measurement and Evaluation)

By

AMICABLE CHIKAKATE BANDA

B.Ed – (University of Malawi)

Submitted to the Department of Educational Foundations, Faculty of Education,

in partial fulfillment of the requirement for the degree of

Master of Education (Testing, Measurement and Evaluation)

University of Malawi

Chancellor College

November 2008

November, 2008

Master of Education (Testing, Measurement and Evaluation)

CHANCELLOR COLLEGE

DECLARATION

I hereby declare that the text of this dissertation entitled: IMPACT OF LINGUISTIC COMPLEXITY OF TEST ITEMS ON PUPILS' PERFORMANCE is substantially my own work

Amicable Chikakate Banda

_____

11TH July, 2008

University of Malawi

Chancellor College

Faculty of Education

The undersigned certify that they have read and recommend to the Postgraduate Studies and Research Committee and the Senate for acceptance of this thesis entitled: IMPACT OF LINGUISTIC COMPLEXITY OF TEST ITEMS ON PUPILS' PERFORMANCE submitted by AMICABLE CHIKAKATE BANDA in partial fulfillment of the requirements for the degree of Master of Education (Med).

Dr Bob W Chulu _____

Main Supervisor

Mr. JJ Bisika _____

Member, Supervisory Committee

Dr Richard Nyirongo _____

Deputy Head, Education Foundations Department

_____

Date

DEDICATION

I dedicate this work to my mother, Mrs. Vainet Kasela, who together with my late father, Mr. Henderson Kasela, guided me in the fear of the LORD and inspired me to like school. May the Good Lord bless you *Amayi*. To you and the rest I say, may the LORD "TEACH US TO NUMBER OUR DAYS ARIGHT, THAT WE MAY GAIN A HEART OF WISDOM" (Psalm 90:12)

ACKNOWLEDGEMENTS

I would like to acknowledge the help that a number of people rendered to me during the development of this work.

First, I would like to express my sincere gratitude to The Malawi National Examinations Board (MANEB), through the Executive Director, for funding the study. I would also like to thank Dr. Bob W. Chulu and Mr. J.J. Bisika for kindly and ably supervising the production of this thesis. I am also indebted to T.T. Zondola, S. Luhanga, V. Zimpita, L.K. Nteleka, A. Mpate, C. Nyasulu, S. Zuka and L. Mwale for the various roles they played during data collection and data analysis stages of this study. May the good Lord bless them richly.

Lastly, but by no means the least, I want to appreciate the support and love I got from my lovely wife, Irene, and our wonderful children, Chikhulupiliro and Shalom. Above all, I thank God for His loving-kindness to us all.

ABSTRACT

The purpose of this study was to examine the effect of language simplification on pupils' performance on the Primary School Leaving Certificate Examination (PSLCE) Mathematics word problems. The study was conducted in view of observations made regarding PSLCE pupils' poor performance in Mathematics in general and their differential performance on mechanical and word problems in particular. It sought to find out changes in difficulty levels of test items after simplifying their language, differences in pupils' performance on the two test forms and pupils' perceptions on the differences in the language of the two test forms.

The study made use of both quantitative and qualitative data obtained from a test and an interview conducted on equivalent groups of a sample of Standard 8 pupils drawn from six schools from the South East Education Division (SEED) using a two-stage cluster sampling procedure. It was found that pupils who took the simplified version of the test performed significantly better than those who sat for the original language version. The differences were statistically significant in 6 of the twelve items that became less difficult. Pupils who were interviewed observed that items in the simplified version were easier as they had shorter and more straight-forward sentences in plain language.

The findings of this study, therefore, support the view that complexity of language in word items is a critical factor in influencing examinees' performance more especially in circumstances where the language of the test is not the examinee's mother tongue. Hence, great caution on language should be taken during setting and moderation of word items in Malawi, where almost all examinees have English as their second or third language.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

BIC	Basic Interpersonal Communication
CALP	Cognitive Academic Language Proficiency
JCE	Junior Certificate Examination
LEP	Limited English Proficiency
MSCE	Malawi School Certificate of Education
MANEB	Malawi National Examinations Board
NAEP	National Assessment of Educational Progress
PSLCE	Primary School Leaving Certificate of Education
SEED	South Eastern Education Division
SMEs	Subject Matter Experts
SPSS	Statistical Package for the Social Sciences
TIMSS	Third International Mathematics and Science Study

CHAPTER 1

INTRODUCTION TO THE BACKGROUND

1.0 Chapter Overview

This chapter provides background information to the research problem, the statement of the problem, the purpose of the study and its research questions. The chapter also discusses the study's conceptual framework and its significance.

1.1 Background to the Study

Research reports have indicated that sometimes students fail to demonstrate what they know about in a given learning area in an examination not because they do not possess the knowledge and skills that are demanded, but rather because the test items that are set to assess these skills and knowledge are shrouded in materials that may have nothing to do with the characteristic the test is measuring (Abedi, 2002; Johnstone, 2003). There are several of such extraneous materials, one of which is the complexity of the language of the test items.

Language of test items plays a critical role in determining examinees' achievement. Mestre (1988) in Abedi, Lord and Hofstetter (1998) observes that in Mathematics, for instance, a particular concern has been on word problems where research has revealed that students in lower classes and those whose first language is not English, perform poorly as compared to those that are fully proficient in the language. Brown (2005) adds that English learners' poor performance at Mathematics problem-solving is a result of their level of English proficiency masking their mathematical knowledge. Learners need to fully comprehend Mathematics word problems in order for them to form a correct problem representation that will lead them to come up with a correct solution. Since almost all the pupils that sit for the Primary School Leaving Certificate Examination (PSLCE) in Malawi acquire English as a second language, it is necessary to investigate the extent to which complexity of the language of test items affects their performance.

1.1.1 Language issues in the PSLCE Examination

In Malawi, primary school pupils sit for their first public examination, the Primary School Leaving Certificate Examination (PSLCE), in Standard 8 upon completion of eight years at the primary school. The subjects which pupils sit for in this high-stakes examination are English, Mathematics, Chichewa, Primary Science and Social Studies. A candidate is deemed to have passed the PSLCE when he or she obtains a pass grade in at least four subjects, one of which must be English.

Except for Chichewa, the rest of the papers at PSLCE are assessed in English, which is a second or a third language to most Malawian pupils. Except for English, pupils start learning their subjects in English from Standard 5, following government's declaration on language of

instruction in primary schools (Malawi Government, Ministry of Education Circular, 1996). Otherwise, from Standard 1 up to Standard 4, instruction materials for the rest of the subjects on the curriculum are in Chichewa.

Introducing instruction through the medium of English quite late has increased teaching and learning challenges in primary schools in Malawi. One of these challenges is that most pupils come to Standard 8 without a firm grasp of the language to enable them engage in meaningful learning through the medium of English (Mchazime, 2001). Experts recommend that second language learners should be exposed to an academic language such as English for a period of between five and eight years if they are to acquire the necessary academic proficiency in the language (Brown, 2005). It is likely that Malawian pupils face language problems during examinations whose test items are in English. However, without empirical evidence, it is not possible to imagine the extent of this problem.

1.1.2 Observations from PSLCE Mathematics Chief Examiners' Reports

An analysis of the 1983 Chief Examiner's report of the PSLCE Arithmetic examination indicated that most pupils found the non-word items easier than word items (Kachaso, 1988). The suspicion here is that this failure was more than anything else, a result of the pupils' difficulties with English, the language of the test.

Similarly, the 2004 PSLCE Mathematics Chief Examiner observed that many pupils fared badly in the paper's word problems (Malawi National Examinations Board, Chief Examiner's Report, 2004). He attributed the pupils' failure in word items to the fact that several items in the paper had long phrases.

A study which Kachaso (1988) carried out showed that in Mathematics word items, language plays a significant role in determining pupils' performance. This conclusion was arrived at as candidates who had their test items translated into Chichewa performed significantly better than those who had their test items in English. It is obvious, however, that many pupils coming from linguistic backgrounds other than Chichewa would not benefit from test items that are translated to Chichewa. Again, translating mathematical concepts into Chichewa has been found to be dogged by serious problems since Chichewa does not possess enough lexical power to render various mathematical concepts (Kachaso, 1988).

1.2 Statement of the Problem

Hiddleston (1996), as cited in Susuwele-Banda (2005), has noted that candidates' performance in Mathematics in Malawi has characteristically been low at all levels of examinations. While research efforts that have been made to establish the reasons behind the underachievement in Mathematics have focused on such factors as the availability of teachers and their qualifications, school infrastructure and the availability of teaching and learning materials in schools. Very little has been done to investigate the impact of test items on pupils' achievement.

Commenting on the Chief Examiner's analysis of pupils' performance of the 1983 PSLCE Arithmetic examination, Kachaso (1988) observed that pupils found non-word items easier than the word problems. Kachaso suspected that this could be the effect of language, explaining that candidates were probably unable to fully comprehend word items and so were unable to successfully work them out because of language barriers.

Again, the 2004 PSLCE Mathematics Chief Examiner hinted in his report that many candidates performed poorly in word items because a good number of the items in that year's paper had long phrases (Malawi National Examinations Board, Chief Examiner's Report, 2004). This further strengthens the suspicion that complex language has an impact on candidates' performance in the PSLCE Mathematics word items. Therefore, this study set out to examine the extent to which use of complex language affects' pupils' performance in Mathematics word problems.

1.3 Purpose of the Study

The purpose of this study was to examine the extent to which simplification of the language of test items impacts on pupils' performance on PSLCE Mathematics word problems. The study sought to determine the achievement gains in primary school pupils' mathematical knowledge after removing language complexity, as a construct-irrelevant material. The researcher chose to base the study on Mathematics because it is relatively easier (in Mathematics) to distinguish between the construct of interest and other materials that are construct-irrelevant (Hambleton and Patsula, 1999; Johnstone, 2003).

1.4 Research Questions

The study sought to answer the following research questions:

1 Does the difficulty level of items change after their language is simplified?

2 Is there any significant difference in pupils' performance on the original test and on the simplified version of the same test?

3 What are students' perceptions about the changes in the linguistic complexity on the simplified version of the test?

1.5 The Study's Conceptual Framework

The study used adaptation (of test items) as its theoretical framework. Linguistic adaptation of a test means modifying the language of the items for the purpose of making them more comprehensible to the test takers. This could either be through translation or simplifying their linguistic complexity. The assumption for adapting tests, according to Johnstone (2003), is that the changes made to the tests will validate test results of the affected students. This assumption is made on the premise that the language of the original items prevents students from demonstrating their true levels of knowledge and skills in the subject area. Fuchs and Tindal (1999) in Johnstone (2003) explain that language simplification is there to level the playing field for all students because the construct-irrelevant variance is removed from the test.

Findings in several studies have shown that revising Mathematics items for the purpose of making them linguistically less complex, enhances students' performance (Abedi, Lord and Plummer, 1997; Hanson, et al 1998; Johnstone, 2003; Barton and Neville-Barton, 2005). Therefore, according to this framework, students' performance improves when the language of the test items is simplified. It is for this reason that the researcher chose adaptation as the study's theoretical framework. It is this framework that has helped in explaining the results of the study.

1.6 Significance of the Study

Research findings on the extent to which linguistic complexity of test items affects candidates' performance in Mathematics word items at PSLCE is of great significance not only to the nation's assessment body, the Malawi National Examinations Board, (MANEB), but also to the rest of the stakeholders in the education sector in Malawi. Some of the players that may benefit from the findings of the study are teacher trainers, education planners, methods advisors,

textbook writers and evaluators and pupils. These stakeholders may find the study significant because decisions to embark on education reform, which the Ministry of Education in collaboration with its respective institutions carries out from time to time, must be backed by research evidence. On issues regarding alleviating problems which students with limited English proficiency face in assessment, Cooley (1991) in Hafner (2000), argues that accommodating students with limited English proficiency accordingly, in assessment practices can inform policy, reform the curriculum and it can also increase accountability in assessment.

The results of the study may be significant to MANEB because the issues being discussed border on fairness in assessment and this may eventually help to improve the quality of education in the nation since it is believed that pupils' academic achievement improves according to CEED (2000) in Rivera and Stansfield (2001), if and when education systems identify what is to be learned, and then assess that material to determine the effectiveness of instruction.

It is also important to realize that results of high-stakes examinations such as the PSLCE have serious effects on pupils' lives since they are used for certification, selection and placement. Considering results of language learners in assessment may provide experts in the education sector with feedback to guide instructional practices (AERA Annual Meeting, 2001). Such attempts may clear any concerns from any quarters that may doubt as to whether or not the PSLCE is a reliable measure of candidates' knowledge and skills. The findings of the study will also add to the corpus of knowledge already existing on the impact of language of test items in learners' achievement in Mathematics.

1.7 Limitations of the Study

The study has only considered how language factors affect pupils' performance in Mathematics and not in other subjects on offer at the PSLCE level. It has also only looked at a selected number of language features that affect comprehensibility in textual test items, namely; item length, vocabulary, syntactical structure and abstract versus concrete presentation of word problems. Other features that research has shown to have impact on language learners' performance in word Mathematics include clause type and word length.

1.8 Chapter summary

This chapter has presented the background to this study by highlighting the important role language plays in Mathematics word items. The observations of the two PSLCE Mathematics Chief Examiners referred to in this chapter are an important indicator to the impact of the level of English used in Mathematics word items on pupils' performance. Since the PSLCE Mathematics examination is in English, findings from Kachaso (1988) and Mchazime (2001) which are based on translated items may not provide a fuller picture of the sort of challenges pupils in Malawi face as they deal with Mathematics word problems.

1.9 Structure of the thesis

The first chapter introduces the background to the study's problem. After presenting language issues in the PSLCE examination, the chapter presents the study's purpose, research questions, conceptual framework and limitations.

Chapter two reviews related literature and research. The review focuses on issues of second language acquisition as they are related to mathematics performance. The chapter then briefly describes the various linguistic features, which some studies have identified as affecting pupils' performance in word mathematics items. Afterwards, the chapter discusses linguistic simplification of test items as contrasted from translating them.

Chapter three presents the study's design and methodology. It also describes the study's sample, methods and instruments used to collect data and how the data were analyzed. The study's results are presented in Chapter four. The chapter discusses these findings with reference to the study's research questions and conceptual framework. Chapter five concludes the study by shading light on the implications of the findings. Finally, recommendations and suggestions for further research are presented.

CHAPTER 2

REVIEW OF RELATED LITERATURE AND RESEARCH

2.0 Chapter Overview

Most of the research work on the issue of language and Mathematics achievement reviewed in this chapter has been done in predominantly English-speaking countries, particularly in the United States of America. These research works have focused on the differential performance of language minority students and language majority students, the impact of language background factors and the relative difficulty of linguistic structures in the language of test items (Abedi, et al. 1998). However, in the Malawi setting, there are no majority language- nor minority language students in as far as English language proficiency is concerned as almost all of them acquire English as a second or a third language. Therefore, this chapter reviews literature which mainly focuses on theories on the general role of language in influencing candidates' performance in Mathematics word problems. The chapter is briefly discussing issues of second language acquisition and the uniqueness of the language of Mathematics and then showing how these factors impact on students' performance. And more specifically, the review has focused on translation and simplification of test items as some of the test adaptation techniques to assist candidates whose first language is not English so as to overcome their perceived language-related challenges in Mathematics word problems.

2.1 Second language acquisition and the language of Mathematics

The importance of language in the teaching, learning and assessment of Mathematics cannot be over-emphasized. Language is a coherent set of symbols that are used to label concepts and provide a quick reference and meaning to concepts (Kachaso, 1988). This then means that students need to be proficient in the language of the test items for them to read and then be able to make sense of these mathematical concepts. This is necessary because students need to read,

decode and adequately comprehend a problem text in order for them to construct, from the verbal form of the problem, a correct conceptual representation upon which the problem-solving process can operate (Kintsch & Greeno, 1985). However, differences in English proficiency may be as a result of varying processes and experiences in acquiring the second language among language learners. The following section discusses second language acquisition and how it impacts on the language learners' attempts to comprehend the text of Mathematics word problems and then be able to solve them.

2.1.1 Second language acquisition and Mathematics performance

According to Ellis (1985), second language acquisition refers to the subconscious or conscious process by which a language other than one's mother tongue is learnt in a natural or a tutored setting. The author observes that the major factors in second language learning are the learner and his/her setting, regardless of the fact that the process is classroom tutored or not. The learner may acquire the language for academic needs or for social interaction.

Cummins (1980) as quoted in Cuevas (1984) identifies two levels of English language proficiency. One is called Basic Interpersonal Communication Skills (BICS). This is a proficiency level that one acquires for the purposes of achieving proficiency in social situations. It is acquired within a short time and it does not help much in academic settings. The other proficiency level, according to Cummins (1980) in Cuevas (1984) is called Cognitive Academic Language Proficiency (CALPS). This is the proficiency that is required in teaching and learning situations and its acquisition takes longer. This proficiency is said to call for a high degree of familiarity with words, grammatical patterns, styles of presentation and argument that is completely alien to ordinary and informal talk. It must be emphasized that it is the acquisition of

CALPS that is critical to the performance of pupils learning a second language in the assessment of content knowledge. This entails that most of what goes on in class will be particularly difficult to a second language learner because, as Fillmore (1982), in Cuevas (1984:135) asserts, second language learning is made more difficult when most of what is being learnt is the language of textbooks.

It may not be surprising therefore, that pupils in Malawian primary schools, most of whom acquire English as a second or a third language, face enormous problems to handle Mathematics word problems. It means that they have to acquire it to optimum levels of proficiency for them to be able to successfully solve Mathematics word problems. However, their efforts to acquire English to acceptable proficiency levels, as noted before, are hampered by the fact that they are introduced to Mathematics instruction in English in standard 5, and not right from the start.

Experts recommend that for one to master the linguistic skills necessary for academic pursuit, one requires to learn the language for a period between five and eight years (Brown, 2005).

In a study to investigate the impact of selected background variables on students' Mathematics performance, Abedi et al (1998) administered three versions of the same test in original English, simplified English and in Spanish to a sample of 1,394 students in grade 8 from an area with a large Spanish population. One of the study's findings was that length of time students had lived in the United States was the strongest predictor of performance in the test. This finding means that the longer language learners interact with the language of the test item at the CALP fluency level the higher their chances to perform better in Mathematics word items.

In Malawi, Mchazime (2001) carried out a study aimed at investigating the effects of English as a medium of instruction on primary school pupils' academic achievement in Standard 7 Social Studies. Two experimental groups received instruction through their mother tongues; Chichewa and Chiyao while the control group was taught in English. Pupils who received instruction through Chichewa and Chiyao obtained higher scores on a test than those who were taught in English. Again, when pupils instructed through Chiyao and Chichewa were taught together through Chichewa, as a medium of instruction, there was no significant difference in the performance of the two groups. The likely reason for similar performance here is that Chichewa and Chiyao are linguistically and culturally closer than they are with English.

Much as this study shows that in Malawi pupils would benefit if test items were in Chichewa or any other local language, it is practically difficult to translate test items in all the country's local languages. Moreover, it has already been pointed out that Chichewa and the rest of the local languages do not have the lexical capacity to fully and competently convey mathematical concepts (Kachaso, 1988). Therefore, it was important to explore the viability of language simplification as a test adaptation technique in alleviating language challenges faced by pupils in Malawi.

It can be argued that pupils' attempts to acquire acceptable levels of academic proficiency in English in Malawi are further frustrated by the fact that some teachers continue to use the vernacular as a medium of instruction even in classes where English is supposed to be used (Susuwele-Banda, 2005). This eventually reduces the pupils' opportunities to acquire more mathematical vocabulary and syntactical structures necessary for them to solve Mathematics word problems. It is not enough to just be aware of these challenges. What is important above all

is to investigate the extent to which the pupils' limited proficiency in academic English affects their performance.

2.1.2 The difficultness of the language of Mathematics

It has already been highlighted that pupils need to adequately comprehend a Mathematics word problem in order for them to be able to construct a correct conceptual representation of the problem. For this to happen, it requires pupils to be mathematically literate so that they are able to read, write and discuss Mathematics (Brown, 2005). However, this is not an easy thing for most pupils, regardless of their mother tongue because Mathematics language is inherently difficult.

The language of Mathematics is different from ordinary language as it is usually abstract, terse and specialized (Abedi et al, 1998). Top amongst the reasons why Mathematics language is very difficult to understand is the fact that it relies heavily on register (Cuevas, 1984). Here, register refers to a meaning given to some words from the ordinary language so as to convey a special meaning in that particular setting. The register is usually technical in nature, not commonly used in daily settings and that it is narrowly defined. Halliday (1975) as quoted in Cuevas (1984) argues that the language of Mathematics is further made difficult because its vocabulary and the way of presenting meanings is borrowed from the natural sciences. As such, understanding and solving Mathematics word problems can be considered an area for experts only (Kintsch and Greeno, 1985).

The fact that the language of Mathematics is inherently difficult creates an extra burden for second language learners. Pupils must struggle to understand and demonstrate their knowledge of Mathematics, whose language is already a difficult one, yet using their weak language. Llabre

and Cuevas (1983) in Cuevas (1984) warn that when interpreting Mathematics achievement by bilinguals, it is advisable to take into consideration variables such as pupils' level of reading proficiency, apart from the skill being measured; otherwise some scores may underestimate pupils' actual abilities in Mathematics.

Some research works have shed light on the difficultness of the language of Mathematics by showing a relationship between readability level and the ability to solve Mathematics word problems. Martin (1964) in Aiken (1972) found out that the partial correlation between reading ability and problem solving, with computational ability partialled out, was higher than the partial correlation between computational ability and problem-solving ability, with reading ability partialled out.

Research has also shown that training pupils specifically with the aim of enhancing their ability to comprehend the unique language of Mathematics enhances performance. Call and Wiggin (1966) in Aiken (1972) conducted an experiment on the effects of Mathematics vocabulary on students' achievement. The experimental group was taught by a teacher of English who stressed the understanding of meanings of Mathematics word problems. Emphasis was also put on how students could translate mathematical statements into symbols. An experienced Mathematics teacher taught the control group. The outcome of this study was that the experimental group did better on the construct in question in the test than the control group. The 2004 PSLCE Mathematics Chief Examiner, referred to in chapter One, would be right in his fears that pupils' poor performance in the paper may have to do with language problems. If this could be the case,

scores in such assessments could lack validity. Realizing the seriousness of how the complexity of test language can erode and confound the validity of test scores, some experts in psychometrics have hypothesized that modifying the language of the test items can increase the validity and reliability of test scores rather than forcing language learners to participate in tests that are linguistically complex for their readability levels (Hanson, et al 1998; Liu, Anderson, Swierzbin and Thurlow, 1999).

2.2 Linguistic features that affect pupils' performance in Mathematics word problems

Research findings have revealed that there are several aspects of the English language that cause the reading and understanding of Mathematics textual items to be difficult. It has also been shown through research that by modifying these features, comprehensibility is enhanced, thereby raising students' performance especially for those who acquire English as a second language (Abedi et al, 1998).

Through research, key English language features that cause difficulty for readers have been identified. These include vocabulary, word length, sentence length, long phrases, clause types, abstract as opposed to concrete presentations, passive voice construction and other language features (Hanson et al, 1998; Abedi et al, 1997; 1998). The effect of each language feature varies depending on the mathematical level or concept being considered, pupils' mother tongue and their level of English proficiency (Barton and Neville-Barton, 2005). Individually and/or cumulatively, these features affect text readers negatively. The language features may slow down readers, cause misrepresentations and sometimes they add to the readers' cognitive load

Barton and Neville-Barton (2005) conducted a study involving 42 Pasifika students learning English in Australia. The study was aimed at exploring the extent to which difficulties in

Mathematics are attributable to low proficiency in English. The study also wanted to discover particular language features that might cause these difficulties.

The results of this study showed that vocabulary on its own was not a serious problem but the combination of syntax and technical vocabulary caused difficulties.

While these results prove true to other English language learners such as the Pasifika in Australia, it would be important to find out the type of language features that cause difficulties for pupils in Malawi at the PSLCE level. The goal in linguistic simplification, which is the thrust of this study, is to reduce language demands placed on candidates so that their scores are valid and reliable estimates of their knowledge and skills in Mathematics. A few of these features are reviewed in the subsequent sections.

2.2.1 Sentence length

Sentence length is one of the commonest factors that affect pupils' ability to read and understand Mathematics word items. Abedi, et al (1998) observes that when sentences are unnecessarily lengthy, they affect pupils' reading skills. Such sentences tend to be more complex syntactically, and therefore, they are more difficult to comprehend. Consequently, they slow down pupils' reading pace, making them less efficient. Johnstone (2003) adds that generally, readability increases when sentence length is reduced. Abedi et al (1997) analyzed 8^{th} grade National Assessment of Educational Progress (NAEP) data from the 1990 and 1992 main assessment items. The results of the analysis showed lower Mathematics scores for the students who predominantly spoke a language other than English at home. Most importantly, though, is the fact that the relationship was more evident for longer items. In the same study, students who spoke a language other than English at home had much higher percentages of omitted and/or "not reached" items than those who spoke only English. This is likely to be so because limited

English proficient students were unable to complete answering all the test items due to their rather slow reading pace caused by a higher level of language demand in the longer items.

Abedi et al (1998) carried out another study, investigating background and linguistic factors which influence students' performance in Mathematics word problems. They administered original and simplified versions of one Mathematics test to a random sample of limited English proficient (LEP) students and those fully fluent in English. It was found that students registered a significant improvement in forty nine percent of the items that were simplified. Most importantly, though, the study found that item length had a stronger impact than other language complexity variables. The authors concluded that although item length may not be a cause of difficulty, it certainly serves as a convenient index for syntactic complexity and can be used to predict comprehension difficulty.

Lepik (1990) as cited in Abedi et al. (1997) found a significant relationship between the length of the problem and the time taken to solve it. This means that a test with many lengthy items will require more time for limited English proficient students to work them out. If the tests are speeded, as the case is with the PSLCE Mathematics, pupils' scores may not be a true representation of their ability in Mathematics as many of them will not manage to answer all the items on the test due to time constraints. This fear corroborates the views expressed by the 2004 PSCLE Mathematics Chief Examiner who felt that part of the poor performance in that year's examination, particularly on the word problems, could be explained on account of their being lengthy (MANEB Chief Examiner's Report, 2004). Without any empirical evidence, however, it would be difficult and dangerous to either accept or refute the Chief Examiner's claims.

2.2.2 Vocabulary

It has already been intimated that one of the reasons why Mathematics appears to be difficult is the fact that it uses unfamiliar language and words. Liu et al (1999) feel that the use of unfamiliar vocabulary in test items such as Mathematics can inhibit less English proficient pupils to answer an item correctly, even if they are capable of performing the task that is asked of them. Hanson et al (1998) point out that mysterious jargon, less frequent words and long words hinder readers from getting the intended meaning in word items. Therefore, the pupils may fail to perform satisfactorily as their attention is divided between employing Mathematics problem solving strategies and coping with difficult vocabulary (Abedi et al 1998). The authors also observe that length of a word is synonymous with word difficulty, saying long words are morphologically complex as more than one semantic unit are packaged together in a single word.

The use of register is another way in which Mathematics specialized language or vocabulary confounds performance. In Mathematics, many words from the ordinary language take on a different meaning when used in the mathematical context. Orton (1997) as cited in Kaphesi (2001) says the potential vocabulary-related problems of learning Mathematics concepts fall into such broad categories as words with specialized meaning, words with specialized emphasis, technical vocabulary and those with varied forms of abbreviations and specialized symbols. A pupil cannot correctly work out Mathematics word problems if he or she fails to understand and interpret these key vocabulary terms or symbols.

It is important to keep realizing that if Mathematics vocabulary is capable of impeding item comprehension for native speakers of English, then it is a double problem for second language learners of English. It is like learning two languages simultaneously. Aiken (1972), reports of a

study done by Linville (1970) which sought to examine the impact of the difficulty level of vocabulary and syntax of arithmetic word problems on students' performance. The results showed that both syntactic structure and vocabulary level were important variables in solving verbal arithmetic problems, with vocabulary level being more crucial.

Shaftel et al (2003) conducted a study to evaluate how some linguistic features affect the performance of different student population groups in Mathematics word items. The items were sourced from the Kansas Mathematics Assessment at grades 4, 7 and 10. The results indicated that while a few language characteristics had unique effects depending on grade level, vocabulary affected all groups. It is interesting to note in the results of this study that effects of individual Mathematics and linguistic characteristics were greater at the lower grade (grade 4) than in the higher grades. The likely explanation for this finding is that students at grade 4 were less proficient in English than were for example, students at grade 10. In view of this, there was no doubt that PSLCE students in Malawi face challenges in Mathematics word problems and there was need to investigate the extent to which this language challenge affects pupils' performance.

2.2.3 Syntactic Structure

Another linguistic feature that affects the comprehensibility of an item is its syntactic structure. Brown (2005) says Mathematics uses language structure which is very specific, different from everyday language. As a result, language learners of English find such syntactic structures very cumbersome and confusing. Some of these syntactic structures include the use of comparatives, passive voice, conditional clauses, and others structural forms (Abedi et al, 1997; 1998; Hanson et al, 1998; Barton and Neville-Barton, 2005). These experts claim that when minor syntactic

changes are made to the wording of Mathematics problems students' performance is enhanced. This is possible because semantic relations are made more explicit, without affecting the underlying semantic mathematical structure.

In a study to determine the effect of syntactic complexity in Mathematics word items on students' performance, Larsen, Parker and Trenholme (1978) in Abedi et al(1997), prepared three tests at different levels of syntactic complexity but equal mathematical difficulty. It was found that scores of low-achieving students in grade 8 on word items were significantly lower on problems containing structures of greater complexity. This shows that language structure in Mathematics word problems effects specific groups of students, most especially those struggling to acquire a second language which is in many ways different syntactically from their mother tongue.

De Corte and Verschaffel (1987) conducted a study to find the effect of semantic structure on first graders' strategies for solving addition and subtraction word problems. They found that at the material and at the verbal level, children at this stage are strongly influenced by the semantic structure of the problems.

Other studies have also shown that there is a major difference between Mathematics language structures and structures in day-to-day English. It has been observed for example, that there is reduced redundancy in mathematical narratives unlike in literary narratives. Brown (2005) argues that it becomes difficult for English language learners to quickly and easily comprehend mathematical structures because these language learners lack the built-in contextual cues found in language arts. Kintsch and Greeno (1985) add that English language learners find difficulties

because in the process of solving Mathematics word problems, they are called upon to infer information that is needed for solving the problem but is not included in the text. It appears to be clear here, that syntactical structure of a textual test item affects the item's potential comprehensibility. Therefore, it was important to investigate the extent to which such language aspects affect pupils' performance in Malawi's at PSLCE as most of the examinees have acquired English as a second or a third language.

2.3 Adapting tests for pupils with limited English proficiency

What has been presented so far is the argument that language is an essential variable not only in the teaching and learning processes of Mathematics, but also in the subject's assessment practices. Kaphesi (2001) thinks that (most) of the problems pupils face in learning Mathematics are language-related. The reason behind this thinking has been advanced earlier on that for one to be able to solve a Mathematics word problem; one has to first of all comprehend fully the word problem so as to form a correct problem representation from the text. Several aspects of the English language have been identified that contribute in making Mathematics language complex. Hence, pupils with less than adequate English proficiency, such as the case is in Malawi, will be treated unfairly if test items place an undue, or worse still, an uninformed emphasis on pupils' reading skills.

There is, therefore, a great need to identify appropriate, valid and reliable means of ameliorating the linguistic challenges faced by Malawian pupils with limited English proficiency in assessment practices. This is necessary because, in essence, every test is inevitably a measure of not only what the learner knows about the particular subject matter, but also the learner's proficiency in the language of the test (Alanis, 2000).

The use of accommodations has been widely proposed as a means of including English language learners in assessments (Rivera and Stansfield, 2001). Several ways of accommodating students with limited English proficiency in assessments have been tested. These include giving students extra time, providing them with bilingual dictionaries, glossaries or any combination of these ways (Abedi, 2002). However, there are broadly two ways where the text of the test is modified linguistically. These are translation and linguistic simplification of the test. Each of these test modification techniques has varying effects in alleviating the linguistic barriers which students face. There is need for the Malawi Government and the MANEB to be guided by credible research findings to decide whether at all there is any need to adopt a particular technique so as to address the perceived linguistic challenges which pupils face in Malawi. The two modes of linguistic adaptation are discussed in detail in the subsequent sections.

2.3.1 Translation

One of the common ways of alleviating language challenges in assessment faced by students with limited English proficiency discussed in literature is translation. The goal in translating test items is to increase the validity of assessment of limited English proficient students (Rabinowitz, Ananda and Bell, 2004). Hambleton and Patsula (1999) add that translating tests from one language and culture to another is advantageous and safe when, among other reasons, the purpose of the test is cross-national or cross-cultural and also when there is no expertise to develop a new test in a second language. Again, test items may be required to be translated if there is need to ensure fairness to examinees in the event that there are several languages.

Otherwise, the two authors warn against translating test items for no compelling reasons, bearing in mind the challenges associated with this adaptation technique.

Hambleton and Patsula, (1999) advise that apart from knowing the source language and the target language, people to be entrusted with translation of test items should be conversant with the cultures of the two languages, be familiar with the construct being assessed, and also have some knowledge of test development. It is like this because, according to Rabinowitz, Ananda and Bell, (2004) simple translation of assessments into various native languages represents a naïve approach to the concern of improving reliability, validity and accessibility of core academic assessments to students less fluent in English.

Proponents of translating test items insist that for the process to serve its purpose well, the process must be sensitive to differences in which different communities use language. This can be achieved by ensuring that translators from the target linguistic communities do participate in the exercise (Solano-Flores et al 2005). Translators also carry out what they call ‘localization’ in order for the language of the test to reflect the language usage of the target community. This, according to Tunick (2003) in Solano-Flores, Speroni and Sexton (2005), is an attempt to adapt the text of the test to the language and cultural requirements of the target environment.

However, it has been observed that translation as a technique to accommodate limited English proficient students is prone to subtle and serious errors. In spite of any careful adherence to translation guidelines and methodical review by content and language experts, translated items may still function differently from the original ones (Rabinowitz, Ananda and Bell, 2004). The process is bound to be dogged by several errors. Erican (1998) in Solano-Flores et al (2005)

reports that studies have shown that even a slight inaccurate translation of a word may influence an item's differential functioning.

Translation of test items brings about inaccuracies and biases. The items are often qualitatively different from the original ones such that test scores are not comparable across the two language forms of the test (Rabinowitz et al 2004). One of the major biases associated with translating test items comes about due to differences in the characteristic the test is measuring, technically referred to as 'construct'. This is common especially when the test is trans-cultural and also when the construct in question shows considerable differences across culture, in both conceptualization and in behaviours associated with it (Van de Vijver and Hambleton, 1996). The construct being measured should remain unchanged when the circumstances of assessment have been changed.

The bias to do with construct can be a real threat to the validity and reliability of assessment in the Malawian context if assessment authorities can adopt it. Kachaso, (1988) acknowledges that one of the major challenges in his study was to find suitable terms in the vernacular to describe various mathematical concepts. And because of this, he relied mostly on loan words from English. For example, the concept of ***ratio*** did not have a suitable rendering in Chichewa; as a result, it was presented as ***resho***.

The problem can be more insurmountable if tests were to be translated into several other local languages. The reason could be what Hambleton and Patsula, (1999) advance that it is not always true to claim that the constructs are universal, especially when tests are cross-cultural. They advise that it is safer to translate tests when the target language and/or culture are closer to

the source language and/or culture. It is important to note that linguistically and culturally, English and Chichewa are not linguistically close such that translation may not be a viable method of adapting test texts.

Another serious threat associated with translating test items from one language to another is item bias, which is also referred to as differential item functioning. Van de Vijver and Hambleton, (1996) explain that it does occur that in one language version, an item can be more difficult than in the other language version due to problems with translation. Therefore, test scores may not be comparable across the language versions of the same test.

In their error analysis of the Mexican translation of the Third International Mathematics and Science Study (TIMSS-1995) test, Solano-Flores et al (2005), observed that negative correlations were consistently higher than positive correlations. They felt that this supported the notion that translation errors tend to bias against the target students. The authors concluded that the combined effect of the frequency and the severity of the translation errors limit its usefulness as a viable means to alleviate problems students face as a result of them not possessing adequate proficiency in English. For this reason, this study does not use translation as a method of adapting linguistically complex word items.

2.3.2 Linguistic simplification/Plain language

In view of the limitations associated with translation as a means of alleviating language challenges faced by pupils with limited English proficiency, especially in the context of Malawi, it is worthwhile to investigate how linguistic simplification can be used to alleviate language-related problems which pupils in Malawi face in Mathematics word items.

In linguistic simplification, as the term intimates, the complexity of the language of the test item is simplified so as to bring it to the level that is no more complex than is necessary to test the examinee's knowledge of the subject matter (Livingston, 1973). Difficult non-mathematical language and non-essential details are removed from test items. Grammatical structures are simplified for the purposes of enhancing not only the item's comprehensibility but also pupils' performance (Abedi, Hofstetter and Baker, 2001). In other quarters, simplified language is referred to as plain language.

Proponents of the use of plain language in testing situations argue that unnecessary language load, difficult vocabulary and syntactical complexity, as already advanced, may disadvantage those pupils who are not fully proficient in the language of the test, in spite of them possessing the required knowledge and skills demanded in test items. The difficult language interferes with the language learners' potential ability to demonstrate their knowledge and skill in the subject are under examination (Shaftel, 2003). It has also been argued that in assessment situations, if a pupil lacks mastery of language in which a test is given then the test becomes, in part, a test of language proficiency. Zoani (2002) observes that if a language used in a test is beyond their comprehension, examinees will often produce a myriad of irrelevant responses, since each one of them will interpret the question differently from the way the examiner had intended. Therefore, there is a great need to reduce the verbal load and syntactical complexity of test items in order to increase the chance of the students taking the test in their second language to do well.

Proponents of linguistic simplification claim that with this technique, there is no concern that it can negatively affect the construct of the test. Rivera and Stansfield (2001) argue that in fact reducing the level of English language proficiency needed to comprehend a test item is good

because it reduces the amount of construct-irrelevant variance in test scores of language learners. It is like this because in Mathematics, for instance, language is not the construct of interest.

There have been some research findings in Mathematics assessment where pupils' performance improved after word problems were simplified. Johnstone (2003) carried out a study of 231 sixth grade students from traditionally underperforming schools where they were tested on the original and the simplified versions of one Mathematics test. The outcome was that students scored significantly higher on the linguistically simplified version of the test.

Abedi et al (1997) conducted a study in which three versions of the same Mathematics test were administered to 1,400 students in grade 8. The sample comprised English language learners and native speakers of English. The test versions were in Spanish, original English and simplified English. One of the findings in the study was that both the native speakers and language learners performed significantly better in the simplified version of the test, leading the researchers to conclude that linguistic clarification of Mathematics items is beneficial to all students. In the same study, eighty percent of the students interviewed preferred the linguistically modified items, arguing that it was easier for them to understand these items rather than those in original English.

Another study carried out by Abedi et al (2000) as cited in Abedi (2002) is on various accommodation strategies for English language learners. The accommodation strategies tested in this study included provision of extra time, glossary, linguistically modified items and glossary with extra time combined. It was found in the study that the provision of linguistically modified

items was the only accommodation strategy that reduced the performance gap between English language learners and native speakers

2.4 Chapter Summary

Studies reviewed in this chapter have shown that language is a critical aspect in determining pupils' performance in Mathematics word items. It has further been shown that comprehensibility of word items is enhanced, particularly to language learners, by simplifying and reorganizing some key English language features. However, most of these studies have been done in societies with marked differences in terms of linguistic configuration with Malawi. For example, there are no language minority groups in Malawi in as far as English proficiency is concerned. Almost all pupils in Malawi acquire English as a second or a third language.

Some studies done in Malawi, which have been reviewed in this chapter, have focused on translation as a technique of adapting test items (Kachaso, 1988; Mchazime, 2001). These studies have shown that pupils in Malawi perform better when assessment items are in the vernacular. It has been noted; however, that translation as a test adaptation technique is bound to face enormous challenges if it were to be implemented in Malawi. Apart from the inherent weaknesses of construct bias and item bias associated with translation, making the decision of which local language(s) tests should be translated into would have serious political and economic implications. It was for these reasons that the current study set out to examine the extent to which linguistic simplification of Mathematics word items affects students' performance in Malawi.

CHAPTER 3

METHODOLOGY

3.0 Chapter Overview

This chapter describes the methodology of the study. This includes the design of the study, instruments employed to collect data, the sample and sampling procedures. The chapter also describes procedures for data analysis. Ethical considerations in the study have also been outlined.

3.1 The Study's Design

The study's approach was predominantly quantitative although it also made use of some qualitative data. The study used the experimental design as its main strategy and specifically, it used a modified posttest only with equivalent groups. The experimental design was chosen to be the study's main research strategy because for one to determine the extent to which the complexity of language of test items affects students' performance presupposes a causal relationship. And to establish such a relationship requires manipulating the variables in the study. This allows the researcher to carry out the most rigorous test of causal hypotheses as the

researcher keeps constant all other extraneous variables so that differences to be noted on the dependent variable should be explained only as a result of the manipulated independent variable (Gall, Borg and Gall, 1996; McMillan, 1992).

3.2 Study Sample

Gall et al (1996:769) define sampling as a process of selecting members of a research sample from a defined population for the purpose of representing the population. A sample is therefore, a subset of the units that make up the population. Sampling is mostly used when the total population cannot be reached for various reasons, such as funds and time. It is for this reason that sampling was preferred in this study.

3.2.1 Sample population and its characteristics

This study targeted Standard 8 pupils in three districts of the South Eastern Education Division (SEED). These districts are Balaka, Zomba Urban and Zomba Rural. The three districts have a population of 231 primary schools altogether. The participants' age range was from 10 to 15 years. At the time of the study, most of these had been in school for eight years and had learnt Mathematics in English for four years.

3.2.2 Sample size and sampling procedure

The study involved a total of 483 students. Of these, 247 were boys, thus 51 percent whilst 236, representing 49 percent, were girls. This sample was drawn from six schools that were randomly selected from three educational districts in the South East Education Division (SEED). The SEED has five educational districts, namely; Zomba Rural, Zomba Urban, Balaka, Machinga and Mangochi. Of the six schools, three were located in the urban area and the other three are located in the rural area. One of these schools was private while five were public schools.

The pupils involved in the study were selected using a two-stage cluster sampling procedure. First, the three educational districts were randomly picked. Then some schools within these selected educational districts were also randomly selected. All pupils in Standard 8 in the selected schools participated in the study. In this sampling procedure, a list of approved PSLCE examination centres maintained and updated at the Malawi National Examinations Board (MANEB) served as the sampling frame.

3.2.3 Subject Matter Experts

The study also involved three subject matter experts (SMEs), whose major task was to develop and compile the two test forms used in the study. The subject matter experts were practising teachers of Mathematics and English language. These were well-qualified teachers with more than ten years of teaching experience at primary school. They had also been involved in assessment activities with MANEB. These activities include item writing, moderation of test items and scoring national examinations. Their experience with MANEB activities was a vital element in their task of modifying the language of the test items. An English language expert who was incorporated in the team was a graduate and he was conversant with the PLSCE Mathematics curriculum.

3.3 Instruments for data collection

The study used a test and an interview guide as instruments to collect data. The test, in two language versions, was meant to examine the effect of complexity of language on pupils' performance. The structured interview guide was designed to gauge the perceptions of pupils on the complexity of language on the two test forms. (Refer to Appendix A and Appendix B)

3.3.1 The test

The test in the study had 25 multiple choice mathematics word items in total covering the following broad topics in the PSLCE syllabus: area; rate, ratio and proportion; profit and loss, measurements; number and numeration. Of these items, 20 were subjected to linguistic simplification. These items were identified for their potential difficult language for Standard 8 pupils in Malawi. Some of these items were the released PSLCE MANEB examinations from the years 2001 to 2006. More problems were sourced from the pupils' Standard 8 Mathematics textbook. On top of the twenty linguistically complex items, additional five items were identified, whose language was considered non-complex. These acted as anchor items. Three Subject Matter Experts, all professionally qualified and experienced in PSLCE Mathematics and English Language compiled the test.

There are two major reasons why the study used items from the students' textbook besides the ones released from MANEB. The first reason was to ensure that there was a wider coverage of the curriculum as items from the released MANEB papers were selected on criteria other than curricular representative-ness. Again, it is common practice with many Malawian schoolteachers, especially during the time the data for the research were being collected, (July) that they practise on past papers in readiness for the PSLCE national examinations that normally take place in early September. In that case the outcome of the research was going to be confounded if some students were to get the items correct due to practice and not as a result of the language factor which the study sought to establish.

3.3.2 Linguistic simplification of test items

The SMEs were briefed on what research and literature say on how the various language features affect the comprehensibility of mathematics word problems. The SMEs also drew on their professional knowledge and their experience to draft items which could be deemed simplified in order to enhance the items' comprehensibility. The simplification exercise focused on language that could present difficulties, language that could confuse pupils or language that could be misunderstood (Abedi et al, 1997).

The features that were targeted were sentence and item length, vocabulary, syntactical structure and abstract as opposed to concrete presentation in mathematics word problems. Only non-technical language was simplified while preserving the underlying mathematical knowledge. The particular changes made to respective items were noted so as to reckon the particular language feature that might have been 'key' to comprehension. (Refer to Appendix A and Appendix B)

Table 3.1: Language Features in Items

Language Feature	Item Number
Item length	7, 17, 18, 21 and 23
Difficult vocabulary	6, 9, 10, 16 and 22
Complex syntactic structure	1, 12, 14, 15 and 19
Abstract presentation	8, 13, 20, 24 and 25

Consequently, two test forms, A and B were created (Refer to Appendix A and Appendix B). Twenty original items together with the five anchor items constituted test Form A, while the

linguistically simplified version plus the anchor items too, constituted test Form B. This is shown in Table 3.2 below:

Table 3.2: Test Forms

ITEM TYPE	FORM A	FORM B
Original English	20	-
Simplified English	-	20
Anchor items	5	5
TOTAL	25	25

3.3.3 Test administration

The two test forms, A and B were randomly assigned to pupils using the spiralling procedure, thus the two test forms were alternately placed to pupils within each class. The reason for the randomization was to create two randomly equivalent groups. Random assignment of test forms to pupils was designed to reduce sources of bias and other threats to internal validity due to selection (Abedi et al, 1998).

Since one school was involved per day, the researcher sought the assistance of other professionals to administer the tests with the help of willing teachers in the participating schools.

An experienced Subject Officer at MANEB and three experienced primary school teachers helped with the administration of the test.

3.3.4 The interview

The study made use of the interview to gauge students' perceptions about the changes in the linguistic complexity of the 20 test items which were linguistically modified. A total of 24 students were purposefully selected for the interviews from the participating schools. The interviewees were purposefully selected in order to capture both girls and boys. Besides, there was a need to have pupils who could express themselves quite well. Each student was interviewed for about ten minutes and the interview was tape-recorded and the researcher also took some notes to complement the information that was audiotape-recorded. In cases where students had difficulties to express themselves, interview procedures were somewhat modified. For example, some questions were asked using the vernacular. Again, where necessary, two or more pupils were brought together to respond to the questions. In such situations the more timid pupils were encouraged to talk and they became more eager to air out their views.

Pupils were asked to study selected pairs of test items from the two test forms and then respond to the interview questions. They were also asked to judge which items, between those in Form A and those in Form B were easier for them. They were asked to explain why they thought such items were easier. Pupils were also asked to state which of the two test forms was easier to read. Again, they had to explain why they thought so. They were also asked about the language(s) they spoke at home and at school and to explain if at all the language(s) they spoke had any effect on their efforts to improve proficiency in English. Except for one school, the researcher carried out the interviews in the rest of the schools.

3.4 Pilot Study

The study's instruments for data collection were pilot-tested at Ndangopuma Primary School. The school lies on the outskirts of the City of Zomba. This school was chosen to pilot-test the instruments because it caters for pupils from both the urban and rural setting. Specifically, the exercise focused on time for the test, instructions on the test, keys for the items and wording of questions on the Interview Guide. After the pilot-test, independent individuals knowledgeable in mathematics, including a school leaver and an experienced Subject Officer at MANEB, revised the test. These individuals provided helpful insights to improve on the wording of the items and on keys or answers to some of the items. One item, for example, did not have a correct answer. The duration of the test was also revised. The test was initially meant for 1 hour 30 minutes and this was reduced to 1 hour 15 minutes as test takers in both forms finished well before time.

3.5 Data Analysis

The study generated both the quantitative and qualitative types of data. The study's quantitative data was generated from the two forms of tests that were created. Qualitative data emanated from the interviews the pupils were engaged in. This section describes several methods and techniques that were followed in analyzing the different modes of data generated from the test and the interview guide.

3.5.1 Group Equivalence

In order to establish that the two groups of pupils taking the two forms of tests were equivalent, their performance on the anchor items was considered. Group means on the anchor items were compared using an independent samples t-test at 0.05 level of significance. This was meant to test the equivalency of groups A and B that took the two test forms.

3.5.2 Difficulty level of items

Several item analyses were performed so as to establish the reliability of the test items. These included calculating difficulty level and item discrimination indices of the two sets of tests. The difficulty of an item, also known as its p-value, refers to the proportion of examinees that answer the item correctly (Oosterhoof, 1990). Although in some quarters the item difficulty level is given as a percentage, in most cases it is expressed as a proportion.

Since each item was administered in both the original and the simplified version of the same test, it was deemed important to examine the differences in the difficulty of these pairs of items. It was further felt necessary to determine whether or not such differences were statistically significant. This is important because when the items with significant differences are qualitatively studied, they provide insight about the sort of language features and changes that significantly affect pupils' performance in word items (Rivera and Stansfield, 2001). The difference on item difficulty between original and simplified items was examined using a t-test at 0.05 level of significance.

Another important aspect in trying to understand the quality of test items is to analyze their discrimination power. This is the item's ability to discriminate between those examinees that possess the knowledge sought after from those that do not. The item discrimination index is also known as the item's D value. Therefore, it was deemed necessary that further to determining the items' difficulty, it would be important to understand the items' discrimination power. Further to this, the study used 'a rule of thumb' proposed by Ebel and Frisbie (1986) as quoted in Backhoff, et al (2000) to determine the quality of test items with reference to their discrimination indices.

Thus, items with D values above 0.39 are rated Excellent, those from 0.30 to 0.39 are said to be Good, those from 0.20 to 0.29 are rated as Mediocre, those from 0.0 to 0.20 are rated as Poor and items less than –0.01 are rated as Worst.

3.5.3 Differences in students' performance

In order to understand the differences in pupils' performance on the two test forms, a number of analyses were carried out and compared. Descriptive statistics of the scores on the two test forms were computed. These include mean, standard deviation, range, minimum and maximum scores. The groups' distributions of scores were also compared by plotting their graphs on the same axis. The group means were also compared through an independent samples t-test at 0.05 level of significance. This was meant to establish whether or not the difference in performance between the two groups of pupils was statistically significant.

3.5.4 Pupils' perception on language complexity differences

For the qualitative data, the researcher transcribed the interviews and read through students' responses to the questions. The categories and themes that were emerging from the data were noted and coded so that sensible stories could be woven through. The categories and themes were compared to the research question(s) for interpretation. The qualitative and quantitative data were integrated to seek larger meaning(s) and/or alternative understanding

3.6 Data management

Pupils' scores on the two test forms were into the computer in order to perform various data analysis methods. Responses of pupils that were involved in the interview were transcribed. Notes which the researcher was taking down as the interviews were being conducted were also put alongside the transcriptions. Interviews that were done in the vernacular were translated into English. (Refer to Appendix E for a sample of an interview with a pupil.

3.7 Ethical considerations

The researcher sought permission from responsible authorities to administer tests and conduct interviews with students in the schools selected. Permission was sought from District Education offices in SEED and Head teachers of schools involved. The researcher assured these responsible officers and the participants themselves that pupil identities, their scores on the tests and their responses on the interview would be treated with utmost confidentiality.

3.8 Chapter Summary

Although the study's design was predominantly quantitative, qualitative data was also used. Using a modified posttest only experiment as its strategy, the study used a test and an interview to collect data from Standard 8 pupils drawn from six schools randomly sampled from the South East Education Division. Qualified and experienced primary school teachers were used to design a linguistically simplified version of a test from released MANEB items. The major data analysis technique for the test scores was comparison of group means through the independent samples t-test.

Chapter 4

DISCUSSION OF THE FINDINGS

4.0 Chapter Overview

This chapter presents the results of the study. It discusses the findings with reference to the research questions, which were presented earlier in the study. Findings on the test for group equivalence are presented first. Then the chapter goes on to discuss changes in the level of difficulty of the test items in the simplified version. Differences in the pupils' performance in the

original and simplified versions of the same test are examined. Finally, the chapter presents pupils' perceptions about the changes in the linguistic complexity of the test.

4.1 Test for group equivalence

It should be stated at this stage that the results of pupils' performance presented in this section are from equivalent groups in as far as ability in mathematics is concerned. This assertion is based on the results of pupils' performance on the five anchor test items. Figure 4.1 below depicts graphically that the two groups performed more or less the same. No group is dominating the other in terms of frequencies on any score. In the figure, frequencies appear to be different for scores 0 and 1 but from score 2 through to 5 the frequencies are close.

Fig. 4.1: Pupils' Scores on Anchor Items

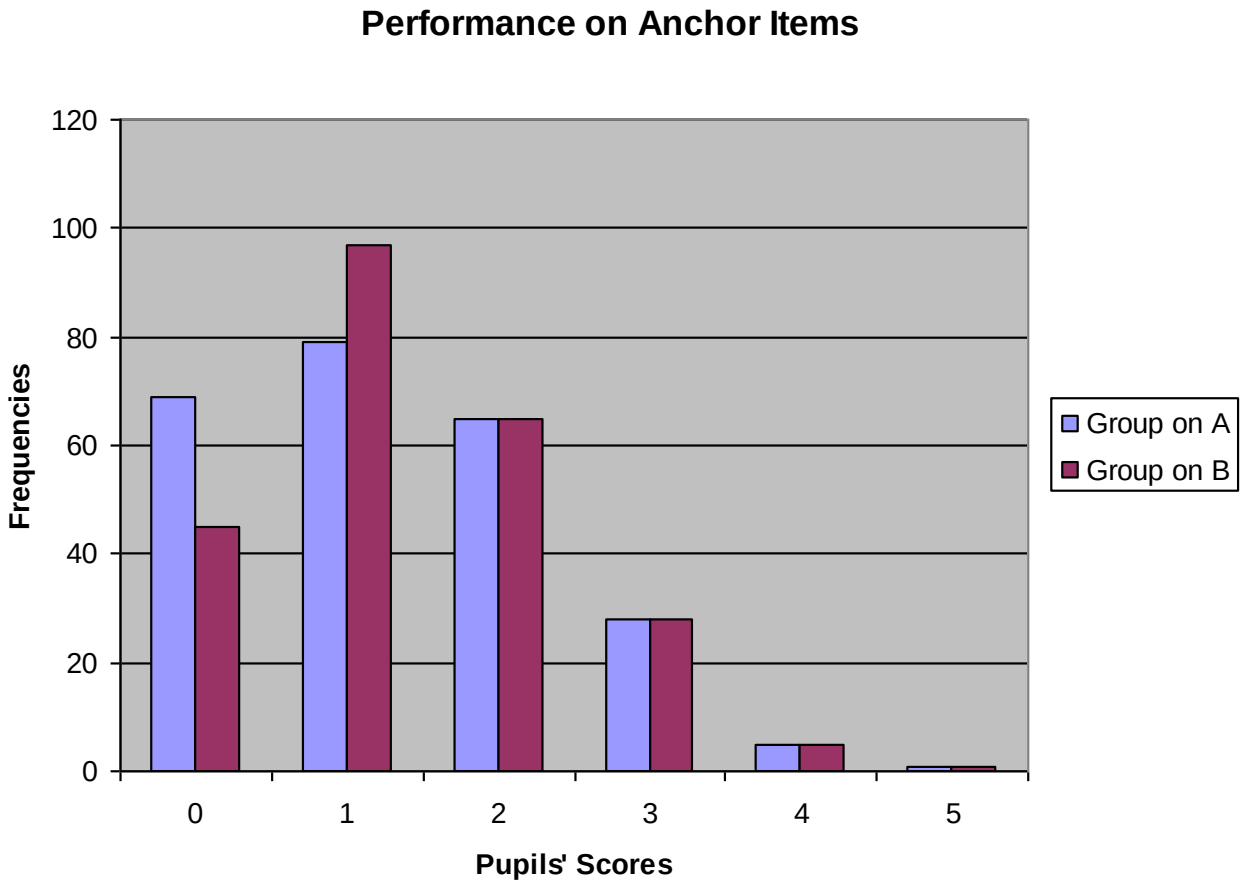


Figure 4.1 shows that the performance of the two groups of pupils on the anchor items was not different. This suggests that the two groups that were created as a result of the spiralling procedure were equivalent in as far as mathematics ability is concerned. The scores for the two groups of examinees on the anchor items were entered into the SPSS software to test for significant statistical difference. The results here also show that the difference in the performance of the two groups of pupils on the five anchor items was not statistically significant. Table 4.2 below, presents the summary of an Independent Samples t-test of the scores of the two groups of pupils' on the five anchor items.

Table 4.2: Summary of an Independent Samples t-Test of Scores on the Anchor Items

Statistic	Test Group		Independent Samples t-Test
	Group with Form A	Group with Form B	
Number of Pupils	247	236	
Mean	1.29	1.43	
Standard Deviation	1.08	1.11	
Standard Error of Mean	6.87	7.23	
t-Statistic			-1.453
Degrees of Freedom			481
Probability-Value			0.147

It is evident from the results above, that although there are some differences in some statistics regarding performance of the two groups of test takers on the anchor items; the performance is statistically not significant. This result, therefore, indicates that the spiral procedure worked to create two equivalent groups with respect to mathematics ability.

4.2 Item difficulty level after language simplification

This section presents results of various item analyses performed on the scores in order to understand the 'behaviour' of items after linguistic simplification. These analyses include the items' difficulty level and discrimination power.

4.2.1 P-Values for items

After scoring the items in each test form, p-values for each item were calculated. As already stated in Chapter 3, p-values are proportions of examinees that get the item correct. Table 4.3 below shows the p-values for each item on the original and on the linguistically simplified items.

In Table 4.3, it is evident that simplifying the language of test items somewhat reduces the difficulty of the items. In their simplified form, twelve items, representing 60 percent of the test, became less difficult. This means that after simplifying the language of the items, a larger proportion of pupils were able to read, understand and solve correctly these items than the case was in the original English version of the test. These are item numbers 6, 8, 9, 11, 12, 15, 17, 18, 21, 22, 23 and 24. In three items; 10, 16 and 19 there was no change in their level of difficulty. Five items, thus 25 percent of the test, became more difficult after the linguistic modification. These are item numbers 7, 13, 14, 20, and 25.

Table 4.3 Difficulty Level Indices for Forms A and B

Item	p-value (Original)	p-value (Simplified)	Difference
6	0.09	0.12	+0.03
7	0.19	0.18	-0.01
8	0.11	0.19	+0.08
9	0.49	0.80	+0.31
10	0.79	0.79	0
11	0.42	0.58	+0.16
12	0.31	0.44	+0.13
13	0.10	0.09	-0.01
14	0.36	0.35	-0.01
15	0.66	0.68	+0.02
16	0.63	0.63	0
17	0.79	0.83	+0.04
18	0.48	0.61	+0.13
19	0.30	0.61	0
20	0.37	0.36	-0.01
21	0.41	0.46	+0.05
22	0.53	0.61	+0.07
23	0.33	0.38	+0.05
24	0.33	0.40	+0.07
25	0.21	0.19	-0.02

To a large extent, the results in Table 4.3, support the view that simplifying the language of mathematics word problems reduces the level of difficulty of the items in question as 60 percent of the items became less difficult after their language was simplified. The various linguistic 'barriers;' such as difficult non-mathematical words, complex syntactical structures and non-essential details which impeded pupils' comprehension of these items were removed. These language barriers disadvantaged pupils with less English proficiency, as the case is with Standard 8 pupils in Malawi. Pupils that took the test in the original English failed to demonstrate their knowledge and skill in mathematics because the language barriers interfered with their efforts to comprehend the items (Shaftel et al 2003). The results in this study are in agreement with what Abedi et al (1998) found in a study investigating background and linguistic factors that influence pupils' performance in Mathematics word problems. In the study, pupils registered a significant improvement in 49% of the linguistically simplified items.

Item numbers 7, 13, 14, 20 and 25 (Refer to Appendix B) became more difficult after their language was simplified. One reason to explain this could be that some of these items were already very difficult and so modifying their language did not help anything at all. The fact that all the items that became more difficult after language modification have p-values of not more than 0.36 in Form A, attests to the suspicion that these items were already difficult due to reasons other than language. Barton and Neville-Barton (2005) argue that the effect language simplification can have in reducing item difficulty depends on a number of factors one of which is the mathematical level or concept being considered. It is likely therefore that these items were difficult to students because of other reasons such as the construct in question, the number of steps involved in solving the problem and other factors other than linguistic complexity. Item

numbers 7 and 14 (Refer to Appendix B), for example, require several steps for one to arrive at the answer. Such items are normally more difficult than those that would require a single step. It is also possible that for unknown reasons, the simplification process made some items difficult. Rivera and Stansfield (2001) observe that linguistic simplification can inadvertently produce a more difficult item.

4.2.2 Items' Discrimination Power

According to Thorndike (1997) an item's discrimination index shows the item's ability to discriminate between those examinees that possess the knowledge in question from those who do not have. One of the several procedures of determining an item's discrimination index is splitting examinees into two, thus those scoring on the higher side and those scoring on the lower side of the median. Then the discrimination index of an item is computed by subtracting the number of pupils who get the item correct in the lower group (N_L) from those that get it correct in the upper group (N_U) and then divide the result by the number in the upper group (N_U) as shown in the formula below:

$$D = N_U - N_L/N_U$$

Table 4.5 shows the discrimination index for the items on both test forms using the formula above.

Table 4.4 Discrimination Power of Items

ITEM	FORM A	FORM B
6	0	0.48
7	0.32	0.59
8	0.56	0.69
9	0.52	0.36
10	0.23	0.32
11	0.46	0.61
12	0.42	0.54
13	0.68	0.60
14	0.17	0.30
15	0.25	0.34
16	0.32	0.56
17	0.19	0.30
18	0.18	0.32
19	0.45	0.58
20	0.69	0.67
21	0.52	0.56
22	0.30	0.46
23	0.42	0.60
24	0.54	0.73
25	0.57	0.71

In Table 4.5 above, 17 items of the experimental 20, thus 85 percent, had their discrimination indices improved in the simplified version of the test. This suggests that linguistic simplification made these items more able to discriminate the examinees who had the knowledge and skills called for in the respective items. This is further evidence that complexity of language of test items inhibits capable pupils' efforts to demonstrate their ability in mathematics word items.

Ebel and Frisbie (1986), as quoted in Backhoff et al (2000) proposed a rule of thumb in determining the quality of test items, in terms of their discrimination indices, where items with D values above 0.39 are rated excellent, those from 0.30 to 0.39, are rated as Good, those from 0.20 to 0.29 are rated as Mediocre, those from 0.0 to 0.20 are rated as Poor, while those less than -0.01 are rated as Worst. Table 4.2.3 below shows how these items in the two test forms are rated with reference to Ebel and Frisbies (1986) 'rule of thumb'.

Table 4.5 Rating of Item Quality According to Discrimination Power

ITEM QUALITY	TEST FORM	
	Form A	Form B
Excellent	11	14
Good	3	6
Mediocre	2	-
Poor	4	-

The figures in Table 4.6 suggest that linguistic simplification of items helped to enhance the quality of test items as all the items in test Form B are rated from good to excellent whereas

some in test form A are mediocre and poor. It can be argued here that linguistic simplification transformed the poor and mediocre items in test Form A, into good and excellent ones in test Form B.

4.2.3 Language features' influence in reducing item Difficulty

The various language features applied when linguistically simplifying the items had different levels of influence in reducing the items' levels of difficulty. Below, is a brief description of how each language feature fared in reducing the difficulty level of the respective items to which it was applied.

4.2.3.1 Item length

Of the four language features that were applied, it appeared that length of items came out to be the strongest influence in reducing the level of difficulty of items. All the five items that were initially long became less difficult after their language was simplified through length reduction. For example, item number 21 is seventy-seven words long in the original form while the simplified form it is reduced to thirty-one words.

These five items were more difficult in Form A because long items become more complex syntactically. As such these items became very difficult to comprehend. This slowed down pupils' reading pace, thereby making these pupils less efficient in solving the word problems. Readability increases when sentence length is reduced (Johnstone, 2003).

4.2.3.2 Syntactic structure

Complex syntactic structure and difficult vocabulary came second to item length in influencing difficulty level of test items in this study. In this category, 60 percent of the items that had vague

and complex syntactic structures in the original test form became less difficult in the simplified form. For example, in item number 14 in Form A, the structure “workers at an estate received K700 each for each two weeks” did not clearly convey the much needed information to help a pupil solve the problem correctly. In Form B, this structure was changed to “workers received K700 each in two weeks”.

It is a recognised fact that mathematics uses language structure that is different from everyday language (Brown, 2005). However, when small changes are made to the wording of mathematics word problems, performance is enhanced. Items in the simplified version of the test became less difficult as most complex syntactic structures were removed and replaced with more familiar structures. Abedi et al (1997) report of findings of a study by Larsen et al (1978) in which scores of low-achieving students in grade 8 on word items were significantly lower on problems containing structures of greater complexity. Therefore, the use of unfamiliar syntactic structures in mathematics word problems, especially to learners whose first language is not English, poses a great challenge if reliable scores are to be realised.

4.2.3.3 Vocabulary

In this study, 60 percent of items that had difficult vocabulary terms in the original form became less difficult. This shows that it is possible for learners to possess the ability to correctly work out a mathematics word problem but this ability can be inhibited if they are unable to

comprehend the item on account of a difficult word. In the simplified version of the test, a number of items became less difficult because the unfamiliar words in the original test form were replaced with more familiar ones. For example, in item number 6, the word “dripping” was changed to “coming out”, whereas in item number 9, the term “silver jubilee” was changed to “twenty-five-years”. Similarly, the word “commences” in item number 22 was changed to “starts”. To a large extent, these changes enabled many pupils that took the simplified version of the test to comprehend the items better and consequently to correctly solve them.

However, vocabulary, as already noted, was not as influential as length of items was in reducing items’ difficulty level in this study. The fact that 60 percent of the items that had difficult vocabulary terms had their difficult indices reduced, compared to 83 percent, for those with item length, shows that not as many pupils were bothered with vocabulary as much as they were with long items, possibly because in most schools pupils are drilled on key and common vocabulary terms associated with mathematics word items. Again, most of the supposedly difficult words included on the test, are found in their books. As such, these words were not very strange.

Another reason why difficult vocabulary was not as influential as item length was in reducing item difficulty level could be that the difficult vocabulary terms in some of the items were not critical to students’ reading and comprehension of the gist of the items. As such most students were able to ramble through them and still understand the items regardless of the difficult words. A good example in this regard is the replacement of the unfamiliar term “auctioneer” with a more familiar term “shopkeeper” in item number 10. There was therefore, no difference in the level of difficulty on the item in the two test forms with .79 apiece. Another example is in item number 16, where the word “purchased” was replaced by “bought”. This change was not “key”

enough to bring about any reduction in item's difficulty level, as the term in question is not critical enough to solving the problem.

4.2.3.4 Abstract vs. Concrete Presentation

On abstract versus concrete presentation, only two items; numbers 8 and 24 became less difficult in the simplified version of the test, representing 40 percent of all the items in this category. In item number 24, "a school purchasing three drums of paraffin" is changed to "Alinafe has three drums of water". Otherwise, the other items that were changed in this manner ended up becoming more difficult. A good example is item number 13, where the presentation of "three different contractors; one from Africa, another from America and the third one from Asia" was changed to "three girls; Chitsanzo, Mayamiko and Ulemu digging a garden". Similar changes made on item number 20, were not strong enough to make the item less difficult as these changes only targeted measurements; thus, knots were changed to kilometres and less familiar names of places were replaced with more familiar ones.

4.2.4 Test for statistical significance in P-Value differences

It must be noted, however, that although 12 items on test Form B became less difficult as a result of linguistic simplification, not all differences in p-values between these pairs were statistically significant. The test for statistical significance was set at .01 level of significance because figures involved are small. Table 4.2.4 below, shows the outcome of the test for statistical significance for the pairs of the p-values of items that had become less difficult after linguistic simplification.

From Table 4.7, it is noted that 6 out of the 12 items that had become less difficult in the simplified version have their differences statistically significant. It is possible, from these results, to identify the type of language feature that had significant influence in determining the level of difficulty of items. For example, although all the items under item length had become less difficult, only one item did so significantly. Table 4.8 below, presents a summary of the items that significantly changed for each category of language feature.

Table 4.6 Results of the Test for Statistical Significance for p-values

Item	p value (Original)	p value (Simplified)	Significance at .01 level
6	0.09	0.12	Not Significant
8	0.11	0.19	Significant
9	0.49	0.80	Significant
11	0.42	0.58	Significant
12	0.31	0.44	Significant
15	0.66	0.68	Not significant

17	0.79	0.83	Not significant
18	0.48	0.61	Significant
21	0.41	0.46	Not significant
22	0.53	0.61	Significant
23	0.33	0.38	Not significant
24	0.33	0.40	Not significant

Rivera and Stansfield (2001) emphasize that information in Table 4.7 is important to item developers to be aware of the specific language features that significantly reduce the difficulty level of items. They argue that such items can be examined qualitatively in order to determine specific reasons for the differences in p-values. This information would be useful to test developers as they review items for subsequent examinations.

Table 4.7 Summary of Significant Changes

Language Feature	Changed Items	Change with Significance	Percentage
Syntactic Structure	3	2	67
Vocabulary	3	2	67
Abstract/Concrete Presentation	2	1	50
Item Length	5	1	20

4.2.5 Omitted items

It has been observed that difficult items also cause pupils with less English proficiency to omit them on a test. Abedi et al (1997) analyzed 8th grade NAEP data from the 1990 and 1992 main assessment items and found that pupils who predominantly spoke a language other than English at home, had a larger percentage of omitted items. In this study too, a larger percentage of omitted items were in Form A, most of whose items were difficult on account of their language as compared to those in the simplified version. Table 4.8 below, shows the percentages of omitted items in the two test forms.

Table 4.8: Number and Percentage of Omitted Items

Test Form	Number of Items Omitted	Percentage of Omitted Items
A	114	61
B	74	39

This means that pupils that took test items in Form A had more problems reading the long and linguistically complex items than the case was with those who took test Form B. Therefore they

just abandoned these items as they did not fully comprehend them and again as they were certainly against time.

4.3 Differences in pupils' performance

This section presents results of pupils' performance on the two test forms. Since the pupils' results on the anchor items have already been presented in section 4.1, the results presented here are those on the 20 items that were linguistically simplified.

4.3.1 Descriptive statistics of pupils' scores

After scoring the two test forms, several analyses of the students' scores indicated that on the whole, pupils who took the linguistically simplified version of the test (Form B) performed better than those who took the test in original English (Form A). Table 4.9 shows the Frequency Distribution of scores on both test forms.

Table 4.9 Frequency Distribution of Scores on test Forms A and B (i)

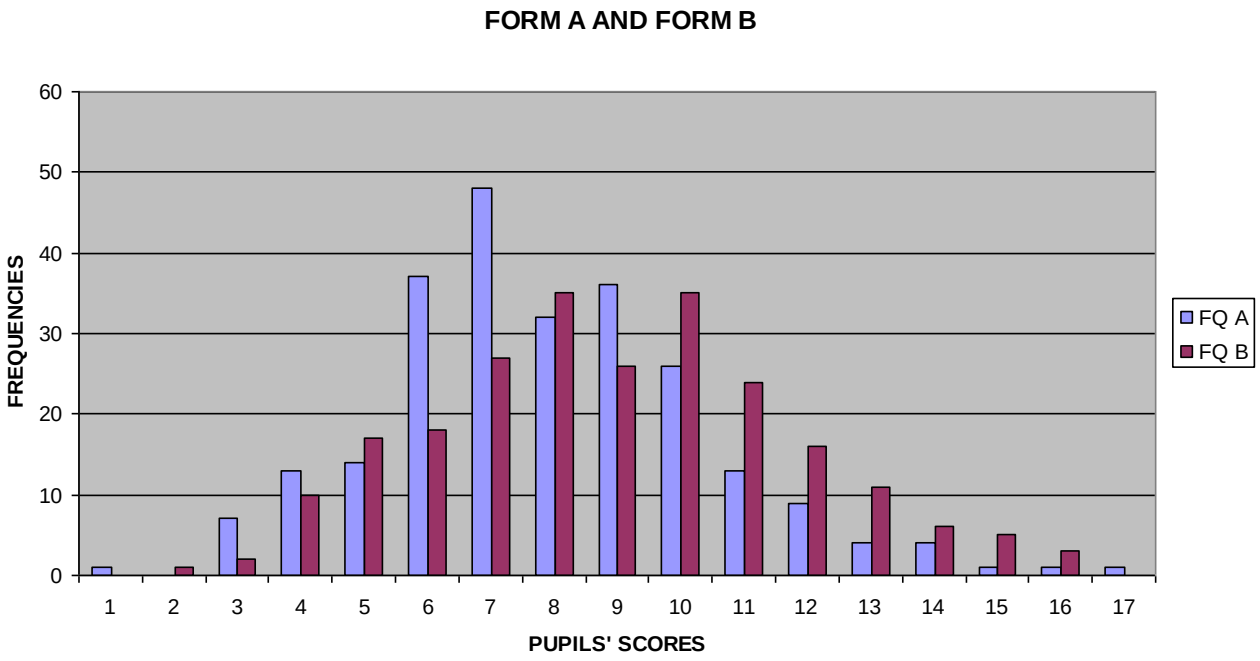
	FORM A		FORM B	
Test score	Frequency	Valid %	Frequency	Valid %
1	1	0.4	-	-
2	-	-	1	0.4
3	7	3.2	2	1.3
4	13	8.5	10	5.5
5	14	14.2	17	12.7

6	37	29.1	18	20.3
7	48	48.6	27	31.8
8	38	61.5	35	46.6
9	36	76.1	26	57.6
10	25	86.6	35	72.5
11	13	91.9	24	82.6
12	9	95.5	16	89.4
13	4	97.2	11	94.1
14	4	98.8	6	96.6
15	1	99.2	5	98.7
16	1	99.6	3	100.0
17	1	100.0	-	-

It is evident from the frequency distributions above that it is pupils that took test Form B that had a better overall performance. If the cut-score for this test were to be fixed at 6, which is slightly one standard deviation below the mean of test Form A, 217 pupils who took Form A would pass the test, representing 88 percent while 206 pupils of those who took Form B would pass it, representing 87 percent. However, if the credit pass were to be fixed at 11, almost one standard deviation above the mean for test Form A, only 33 pupils who took test Form A, representing 13 percent of the examinees would pass with credit while 65 pupils, almost double the number of the pupils in Form A, would pass it with credit in Form B, representing 28 percent. This suggests that linguistic simplification made on the items helped pupils in test Form B to perform better

than their counterparts taking test Form A. Figure 4.10 below, presents this information graphically.

Fig. 4.10: Frequency Distributions of Scores in test forms A and B. (ii)



Although scores for both test forms would be classified as normally distributed, it can be noted from Figure 4.10 that the two distributions do not necessarily superimpose each other. In test Form A more scores are on the lower end of the graph, hence, positively skewed, than the case is with scores in test Form B. Scores for Form B, on the other hand, display more of a symmetric distribution, with equal halves from the median (9). Since it has already been shown that the two groups taking the tests were equivalent with respect to mathematics ability, this observed difference in the distributions can be attributed to the linguistic differences in the test forms. Table 4.11 presents the summary of the descriptive statistics of the scores in the two test forms.

Table 4.11: Descriptive Statistics of Pupils' Performance on test Forms A and B

Statistic	Original Test Form (A)	Simplified Test (Form B)
N	248	236
Mean	7.89	8.90
Std Error of Mean	0.16	0.18
Median	8.00	9.00
Std Deviation	2.54	2.82
Variance	6.46	8.10
Range	16	14
Minimum score	1	2
Maximum score	17	16

To underscore the point that pupils that took items in test Form B performed better on the test, all the statistical indicators in Table 4.11 attest to this fact. The mean for the sample in test Form A is 7.89 with a standard deviation of 2.54 while that for Form B is 8.90, with a standard deviation of 2.82, representing a mean difference of 1.01 units on the score scale.

4.3.2 Test for significance in pupils' performance

Further to the descriptive statistics indicated and described above, all the test scores were entered into the SPSS software and a t-test for independent samples was performed in order to determine whether or not the differences in the performance of pupils on the two test forms as observed in the descriptive statistics above were statistically significant. The result of this statistical test

showed that the difference in the two means is statistically significant at .05 level of significance. Table 4.12, shows the result of this statistical test.

This result means that the probability that the difference in the performance of pupils on the two test Forms occurred by chance is less than .05. It can therefore be argued that for the pupils that were sampled for this study, differences in the complexity of language of test items, as a factor, had a major influence on their performance. This conclusion is most probable considering the fact that the groups of pupils that sat for the two forms of the test were equivalent in as far as differences in mathematics ability are concerned.

This result means that the language of the test items for the students that were sampled in this study had a major influence on their performance.

Table 4.12: Summary of an Independent Samples t-Test on test Forms A and B

Statistic	Test Form		Independent Samples t-test
	Original English	Simplified English	
Number of Pupils	247	236	
Mean	7.89	8.90	
Standard Deviation	2.54	2.82	

Standard Error of the Mean	0.16	0.18	
t-Statistic			-4.179
Degree of Freedom			481
Probability-Value			0.000

4.4 Pupils' perceptions on the complexity of language of test items

In the interview (Refer to Appendix C), 21 out of the 24 pupils that were interviewed, representing 87.5 percent, indicated that test items in Form B were easier than those in Form A. They, therefore, said they preferred items in Form B because most of them (items) were shorter and had language that was straightforward, hence, easier to understand. They argued that the longer and wordier statements in Form A confused them and distracted their efforts to comprehend the task in the items. The major issues that pupils took into consideration in making their preference were length of the item and difficult vocabulary.

4.4.1 Length of items

On the length of items, pupils felt that the long items in Form A impeded their attempt to comprehend the problems as most of the words in the items were unnecessary and did not in any way contribute to the process of coming up with a correct answer. As such reading through them only consumed their precious time. They observed that the shorter and more concise statements of the same items in Form B not only had the necessary key words to solve the items but also presented the task(s) more clearly. Only one interviewee said that in some mathematical problems, longer statements are necessary so as to give a clearer picture of the task.

4.4.2 Vocabulary

On vocabulary again, 21 pupils out of the 24 interviewed, representing 87.5 percent indicated that Form B had more readable and more familiar vocabulary terms which could not impede their efforts to understand the task in the test items. According to the pupils interviewed, the following were some of the difficult vocabulary terms in test Form A: “commence”, “silver jubilee”, “annually”, “purchase”, “dripping” and “value”(Refer to Sample Transcript in Appendix D). This claim was attested to during the interviews as some of them had great difficulties to read these and other difficult vocabulary terms. Pupils said that although they might have come across some of these vocabulary words, the examination setting made the situation worse. One interviewee put it like this,

> “Yes I know what the term “silver jubilee” means but the tension brought about by examinations caused me to forget it altogether”.

On the same issue of vocabulary, another interviewee confessed that he got the item wrong just because he did not know that “silver jubilee” refers to twenty five years. He said had he taken test Form B where the task was clearer, he could have solved the item correctly.

4.4.3 Test version and time

In terms of time, 10 of the 12 pupils who took test Form A and were interviewed indicated that test items in Form B were easier than those in Form A. They further said reading through and trying to understand the long items in Form A took them longer than the case was with items from test Form B. Regarding vocabulary, they said the difficult words in Form A impeded them and so slowed down their efforts and pace to read and understand the word items. On the other

hand, they said test items in Form B had shorter sentences which were put in more straightforward language therefore taking shorter time to read through and work out the answers.

4.4.4 Language used at home and at school

On language used at home, all students said they spoke Chichewa. Three of them said they also spoke English apart from Chichewa because they had parents/wards and relatives that were educated. Except for the interviewees from the private school who said speaking English was compulsory at school, all the interviewees from public schools said they mostly spoke Chichewa at school. English was spoken occasionally in class. Incidentally, most of the students that go to private schools in Malawi come from homes where English is sometimes used.

4.4.5 Advice to item setters

Regarding advice to setters of Mathematics test times, all the interviewees said test items should not contain unnecessarily long sentences. They said test makers should try as much as possible to remove all unnecessary words that contribute nothing to the essence of the task. They also said that item setters should not include difficult words. They observed that longer sentences and difficult words prevent them from fully getting the meaning from the items as they wander in the "forest" of words thereby wasting time.

Since Malawian pupils are still in the early stages of English language acquisition at the PSLCE, setters need to be extra cautious of the language they use in mathematics word items as De Corte et al (1985) as quoted in Abedi et al (1997:47) observe that inexperienced problem solvers, who lack highly developed semantic schemata for problem solving, rely more on the text.

4.5 Chapter Summary

This chapter has presented the findings of the study with respect to the study's research questions. The various analyses of the quantitative data presented in this chapter have shown that largely, pupils performed better in the linguistically simplified items than they did in the items in original English. Pupils that were interviewed also showed their preference for the linguistically simplified test items. Taking into consideration that the results in the study have shown that the spiralling procedure succeeded to create two equivalent groups of pupils in as far as mathematics ability is concerned, the conclusion that one can make is that the difference in pupils' performance is as a result of the difference in the language of the two test forms. Therefore, the results in this chapter have shown that language of mathematics word items significantly affects pupils' performance.

Chapter Five

CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS

5.0 Chapter Overview

This chapter provides the conclusions, implications and recommendations of the study based on its findings and in light of the study's theoretical framework. Although the conclusions and recommendations made in this chapter are drawn from the understanding of the entire study, they follow specific research question of the study. In order to determine the effect of language complexity on pupils' performance in Mathematics word problems, the study sought to find out changes in difficulty level of test items after simplifying their language, differences in pupils' performance on the two versions of the test. Finally, the study gauged pupils' perceptions on the differences in the language of the two test forms.

5.1 Difficulty level of test items

The results of this study support the view that simplifying the linguistic complexity of mathematics word items reduces the difficulty levels of items in question. In this study, 60 percent of the items whose language was simplified had their difficulty level reduced. Thus, these items became easier than the case was before the linguistic simplification was effected. This implies that a larger proportion of pupils correctly answered the items as the language of the items was simplified.

This underscores the fact that language is a key factor in influencing pupils' performance because for one to know the task in the item, one has to first of all be able to read and fully comprehend the text of the item. Only then is one able to formulate a correct problem representation that will lead to a correct answer. Hanson et al (1998) say successful performance in tests often depends on students' ability to read, decode and comprehend the written text. Therefore the simplification exercise which involved reducing the length of the items, removing difficult words and simplifying syntactic structures, enhanced pupils' reading of the word items and made the items more comprehensible.

5.2 Differences in pupils' performance

The differences in pupils' performance on the two forms of tests as observed from the preliminary analyses of the pupils' scores (Table 9 and Fig. 10) and also from the Independent samples t-test (Fig. 4.2) are not there due to differences in the basic abilities between the two groups of pupils. It has already been proven that these two groups were equivalent. Therefore, this suggests that the differential performance is there due to the differences in the complexity of the language of the items in the two test forms. The linguistic simplification process enhanced

the pupils' ability to comprehend the items more clearly. As such they were more capable of constructing a proper problem representation and thereby solving the problem correctly. Pupils' comprehension for items in test Form B was enhanced because some items were made shorter, others had their difficult vocabulary replaced with more familiar terms, and others had less complex syntactical structures.

The findings of this study are consistent with several others which have shown that there is a high correlation between mathematics performance and reading skills. Mc Ghan (1995) in Brown (2005) reported a correlation of .84 between fourth graders' reading comprehension and mathematics scores. Abedi et al (1998) state that studies have shown that students' performance in mathematics word problems is impeded when statements are long and also when problems are presented in abstract rather than concrete terms. The simplification of the items' complex language in this study improved the pupils' reading and comprehension skills and consequently, their performance.

5.3 Pupils' perception of differences in language complexity

Responses of pupils involved in the interview clearly indicate that they are aware of the negative effects of complex language on their performance in mathematics word problems. Pupils' preference of the linguistically simplified items means that it is possible to come up with word problems in mathematics that can enable pupils to demonstrate their genuine competence in the subject.

5.4 Summary

The major purpose of the study was to examine the impact of linguistic simplification of Mathematics word items on pupils' performance. Specifically, the study sought to find out whether difficulty levels of items change after linguistically simplifying them. The study also wanted to examine whether there would be any significant differences in the performance of pupils taking the original test and those taking a simplified version of the same test. Finally, the study wanted to understand pupils' perceptions on the differences in linguistic complexity between the two forms of the test.

The results of the study have shown that students who took the simplified version of the test performed significantly better than those who took the test in its original English. The results of this study indicate that clarifying the language of test items, thus, reduction of item length, removing difficult vocabulary, simplifying syntactic structure and using concrete rather than abstract presentations of word items reduces items' difficulty levels, thereby, helping pupils to improve their performance.

5.5 Recommendations

From the findings of this study a number of recommendations can be made when one considers the language of mathematics word items to be included on the PSLCE national examination and indeed other lower level classes. This can be both at setting and moderation stages of test development. These recommendations are listed below:

1. Since language complexity in mathematics word problems has been found to be a critical aspect in determining the performance of pupils, teachers and item setters should always be wary about the kind of language they employ in crafting items for students in lower level classes. This caution should be both on national examinations and classroom-

based assessment. The language used in mathematics word items should be commensurate with pupils' level of English language proficiency.

2. MANEB should emphasize issues of clarity of language in mathematics word items when training setters and moderators for national examinations at the PSLCE level. Brown (2005) observes that research has shown that even writers who possess considerable knowledge about a particular topic may overestimate their target audience's ability to comprehend. Hanson et al (1998) add by saying that items that are more comprehensible to their readers yield more valid estimates of students' skills.

3. It is necessary for teachers to prepare students on key and common language features. In view of pupils' comments, and also from the "behaviour" of some items, it is evident that when pupils are trained about specific mathematical language features such as vocabulary terms, they face fewer challenges in tackling word problems than when they are not prepared This may not be the sole responsibility of English Language and mathematics teachers in schools but also for all the teachers in all subjects so that pupils' linguistic proficiency is enhanced.

4. Looking at the difficulties Malawian pupils face with the language of Mathematics, it would be worthwhile for Government and other stakeholders to come up with policies to ensure pupils start being introduced to mathematics in English as early as Standard 1 so as to accord them with ample time to acquire sufficient levels of English proficiency as they sit for the PSLCE in Standard 8.

5.6 Suggestions for further research

In order to answer some related questions on the effect of the language of test items on pupils' performance and also to contribute more knowledge on the issue, the following areas are recommended for further research:

1 Comparing performance of Malawian students in Mathematics between mechanical and word problems.

2 Doing a study similar to this one but at an advanced examination level such as at the Junior Certificate Examination (JCE) or at the Malawi School Certificate of Education (MSCE) examination.

3 Carrying out a similar study with a larger sample drawn from all parts of Malawi.

4 Doing a study similar to this but using performance based assessment instead of multiple choice items

5 Finding out pupils' language backgrounds and how these affect their performance in learning and assessment activities at school.

6 Carrying out a study similar to this but with a pre-test and posttest design.

REFERENCES

AERA Meeting, (2001) *Position Statement Concerning High-Stakes Testing* Retrieved on 15-10-2006, from http://www.jstor.org/sici?sici=0013-189X(200011)29%3A8%3C27%3AIRTAPS%3E2.0

Abedi, J. (2002) *Assessment and accommodations of English language: Issues, concerns and recommendations.* Retrieved on 06-12-2006, from http://cehd.um.edu/NCEO/LEP/Accommodations/AccomLepresource.htm

Abedi, J., Lord, C. & Plummer, J.R. (1997) *Final report of language background as a variable in NAEP mathematics performance.* Retrieved on 13-10-2006 from http://www.cse.ucla.edu/Reports/TECH429.pdf

Abedi, J., Lord, C., & Hofstetter, C. (1998) *Impact of selected background variables on Students' NAEP math performance.* Retrieved on 4-12-2006 from http://www.cse.ucla.edu/Reports/TECH478.pdf

Abedi, J., Hofstetter, C. & Baker, E. (2001) *NAEP math performance and test Accommodation: Interactions with student language background.* Retrieved on 5-12-2006 from http:www.cse.ucla.edu/Reports/newTR536.pdf

Allanis, I. (2000) A Texas two-way bilingual program*:* Its effects on linguistic and Academic achievement, *BilingualResearch Journal*, Vol.24, No.3 Retrieved on 4-12-2006 from http://www.brj.asu.edu/v243/artclesar2.html

Aiken, L.R. (1972) Language factors in learning mathematics *Review of Educational Research* Vol.42, No.3. Retrieved on 26-01-2006 from http://www.jstor.org

Backhoff, E., Larrazolo, N., & Rosas, M. (2000). The level of difficulty and discrimination power of the Basic Knowledge and Skills Examination (EXHCOBA). *Revista Electronica de Investigacion Educativa,* 2 (1). Retrived on 27-07-2007 from http://redie.uabc.mx/vol2no1/contents-backhoff.html

Barton, B. & Neville-Barton, P. (2005) *The relationship between English language and Mathematics learning for non- native speakers.* Retrieved on 06-12-2006 from http://www.tlri.org.nz/pdfs/9211-summaryreport.pdf

Brown, P.J. (1999) *Findings of the 1999 plain language field test.* Retrieved on 17-12-2006 from http://www.doe.k12.de.us/AAB/aTCH204Findingsofthe1999 PlainLangugeField20Testpdf

Brown, C.L. (2005) Equity of literally-based math performance assessments for English language learners in *Bilingual Research Journal,* Vol. 29 No2. Retrieved on 06-12-2006 from http://brj.asu.edu/content/vol29-no2art5.pdf

Chimwenje, C. (1992) *Evaluation of the Primary School Leaving Certificate Examination In Malawi: How the examination meets its educational and selection goals* Unpublished doctoral thesis, University of Sussex

Cohen, L., Manion, L. & Morrison, K. (2000) *Research Methods in Education,* London Routledge Falmer

Cuevas, G.J., (1984) Mathematics learning in English as a second language in *Journal for Research in mathematics Education,* Vol.15, No.2. Retrieved on 17-01-2007 from http://www.jstor.org

De Corte, E & Verschaffel, L. (1987) The effect of semantic structure on first graders' Strategies for solving addition and subtraction word problems in *Journal for Research in Mathematics Education,* Vol.18, No.5

Ellis, R. (1985) *Understanding second language acquisition,* Oxford, Oxford University Press

Gall, M.D., Borg, W.R. & Gall, J.P. (1996) *Educational Research: An Introduction,* New York, Longman

Hanna, E.I. (2005) *Inclusive design for maximum accessibility: a practical approach to Universal Design*. Retrieved on 13-10-2006 from http://www.pearsonedmeasurement.com/research.htm

Hafner, A.L (2000`) *Evaluating the impact of test accommodations on test scores.* Retrieved on 17-12-2006 from http://ww.doe.k12.de.us/AAB/Atch13 EvaluatingtheImpactofTestAccommodationsonTestScores.pdf

Hanson, M.R., Hayes, J.R., Schriver, K., LeMahieu, P.G. & Brown, P.J. (1998) *A Plain Language approach to the revision of test items.* Retrieved on 04-12-2006 From http://www.doe.k12.de.us/ABB/Atch1Plainlanguageapproachtothe revisionoftestitems.pdf

Hambleton, R.K. & Patsula, L. (1999) *Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices.* Retrieved on 06-10-2006 from http://www.testpublishers.org/journal1.htm

Hayes, J.M. (1970) *Educational Assessment of immigrant pupils.* London, National Foundation for Educational Research

Hinkle, D.E., Wiersma, W. & Jurs, S.G. (1998) *Applied statistics for the behavioural Sciences,* New York, Houghton Mifflin Company

Johnstone, C.J. (2003) *Improving validity of large-scale tests: Universal Design and student performance (Technical Report 37).* Minneapolis, University of Minnesota, National Centre on Educational Outcomes. Retrieved on 12-10-06 from http://www.education-umn.edu/NCEO/OnlinePubs/ Technical37.htm

Kachaso, L. (1998) Language effects on pupils' performance in mathematics' word Problems in Standard Seven, Unpublished Med thesis, Zomba, University of Malawi

Kaphesi, E.S. (2001) The use of language in mathematics teaching in primary schools in Malawi, Unpublished doctoral dissertation, University of Nottingham

Khanje, C.K., (1987) Effects of school resources availability on secondary school Students' examination achievement in Malawi, Unpublished Med thesis, Zomba, University of Malawi

Khisty, L.L. (1992) *A nationalistic look at language factors in mathematics teaching in Bilingual classrooms.* Retrieved on 06-12-2006 from http://www.ncela.gwu.edu/pubs/symposia/third/khisty.htm

Kintsch, W. & Greeno, J.G. (1985) Understanding and solving word arithmetic problems

In *Psychological Review*, Vol. 92, No.1

Liu, K. Anderson, M.E., Swierzbin, B. & Thurlow, M (1999) *Bilingual accommodation For limited English proficient students on statewide reading tests: Phase 1 (Minnesota Report No.20).* Minneapolis, MN: University of Minnesota, National Centre on Educational Outcomes. Retrieved on 23-03-2007 from http://www.education.umn.edu/NCEO/OnlinePubs/MnReport20.html

Liwewe, S. (1987) An examination of the relationship between teacher variables and Pupil academic achievement in MSCE English and Mathematics in secondary schools in Malawi, Unpublished Med thesis, Zomba, University Of Malawi

Livingston, S.A. (1973) Verbal overload and achievement tests: a replication, *American Educational Research Journal*, Vol.10, N0.10. Retrieved on 06-10-2006 From http://www.jstor.org

Malawi Government, Ministry of Education (1996) Circular on language of instruction in Primary schools in Malawi. Ref. No. IN/214

Malawi National Examinations Board (2004) PSLCE Mathematics Chief Examiner's Report

Mchazime, H.S. (2001) Effects of English as a medium of instruction on pupils' Academic achievement in social studies in Malawi. Unpublished doctoral dissertation, University of South Africa

McMillan, J.H. (1992) *Educational Research: Fundamentals for the consumer* New York, Happer Collins Publishers

Oosterhof, A., (1990) Classroom Applications of Educational Measurement, Prentice Hall, New Jersey

Paul, D.J., Nibbelink, W.H. & Hoover, H.D. (1986) The effects of adjusting readability On the difficulty of mathematics story problems in *Journal for Research In Mathematics Education*, Vol.17, No.3. Retrieved on 06-10-2006 from http://www.jstor.org

Rabinowitz, S., Ananda, S. & Bell, A. (2004) *Strategies to access the core academic knowledge of English language learners.* Retrieved on 17-01-2007 from http://www.testpublishers.org/Documents/JATTdec2004-ELL-strategies -papepdf

Rivera, C. & Stansfield, C.W. (2001) *The effects of linguistic simplification of science Test items on performance of limited English proficient and monolingual English-speaking students.* Retrieved on 17-10-2006 from http://www.ceee.gwu/Products-Assess.Accountability/simplification.pdf

Shaftel, J., Belton-Kocher, E., Glasnapp, D.R. & Poggio, J.P. (2003) *The differential Impact of accommodations in statewide assessment:* Research summary. Minneapolis, MN: University of Minnesota, National Centre on Educational Outcomes. Retrieved on 10-12-2006 from http://www.education.umn/NCEO/TopicAreas/AccommodationsKansas.htm

Solano-Flores, G., Speroni, C. & Sexton, U. (2005) *Test translation: Advantages and Challenges of a socio-linguistic approach.* Retrieved on 17-01-2007 from http://www.air.org/documents/AERA2005TesttranslationAdvantages.pdf

Solano-Floes, G., Contreras,-Nino, L.A. & Backhoff, E. (2005) *The Mexican translation Of TIMSS-1995: Lessons on test translation from a post-mortem study.* Retrieved on 17-01-2007 from http://www.air.org/news/documents/ NCME2005MexicanTranslation.pdf

Susuwele-Banda, W.J. (2005) Classroom assessment in Malawi: Teachers' perception and practices in mathematics. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University

Thorndike, R.M. (1997) *Measurement and evaluation in psychology and education* New Jersey, Prentice Hall

Van de Vijver, F. & Hambleton, R.K. (1996) Translating Tests: Some practical Guidelines. Retrieved on 17-10-2006 from http://arno.uvt.nl/show.cgi

Zoani, A.J.R., (2002) Moderators: A paper presented at a training of trainers workshop Zomba, 18th-20th November, 2002

APPENDICES

Appendix A: Test in original English (Form A)

1. The basic salary of an individual is K6 740. If a 10 percent increase is given, how much is the new salary?

A. K7 127 B. K7 324 C. K674 D. K7 414

2. A company pays a commission of K60 for every K800 sales. How much commission does it pay for sales worth K32 000?

A. K240 B. K400 C. K2 400 D. K1 800

3. A school has enough food for 300 pupils for 20 days. How many pupils will eat the same amount of food in 15 days?

A. 400 pupils B. 115 pupils C. 225 pupils D. 40 pupils

4. A boy sold 150 oranges. Find the amount he paid as a market fee if he was charged 20t on each orange sold.

A. K300 B. K75 C. K3 D. K30

5. There are 128 poles equally spaced in a straight line . If the distance from one pole to the next pole is 20 m, calculate the distance form the first pole to the last pole.

A. 2560 m B. 2040 m C. 2540 m D. 2440 m

6. A tank holds 70 litres of water. If the tank has a hole and water is dripping at the same rate of 0.5 litres per minute, how much water would be in the tank after one hour?

A. 40 litres B. 35 litres C. 30 litres D. 100 litres

7. A rectangular plot measuring 14 m by 12 m is to be fenced with barbed wire. Find the cost of fencing the plot if the wire is sold at K70 per metre.

A. K1 400 B. K3 640 C. K1 820 D. K3 540

8.A suit marked K10 000 was sold at a cash discount of 7.5%. Calculate the cash price of the suit.

A. K750 B. K925 C. K8 250 D. K9 250

9. When will a boy celebrate his silver jubilee birthday if he was born in 1997?

A. 2012 B. 2022 C. 2032 D. 2015

10. An auctioneer is paid a commission of 20% for goods sold. How much commission did the auctioneer receive if K112 000 worth of goods were sold?

A. K8 9600 B. K1 120 C. K2 400 D. K22 400

11. Find the least length of cloth that can be cut into pieces of 9 cm, 15 cm and 24 cm long.

A. 360 cm B. 325 cm C. 125 cm D. 3 cm

12. For every three pieces of bread that Natowa eats, Chimwemwe eats one piece. How many pieces of bread did Chimwemwe eat if they both ate a total of 100 pieces?

A. 33 pieces B. 75 pieces C. 25 pieces D. 50 pieces

13. A contractor from Africa takes 48 months to construct a road. An Asian contractor takes 72 months to do the same job while an American contractor takes 36 months. Find the number of months the three would take to finish constructing the road if they were working together.

A. 156 months B. 52 months C. 16 months D. 26 months

14. Workers at an estate receive K700 each for each two weeks. Calculate the total amount of money which two workers will receive after working for 42 days.

A. K2 100 B. K3 200 C. K4 200 D. K8 400

15. A bicycle valued at K7 000 is charged tax at the rate of 20t in every kwacha. How much tax does the buyer pay?

A. K14 B.K5 600 C. K1 400 D. K3 200

16. Mrs. Dzino purchased a refrigerator from a manufacturer for K75 000. After selling it to Mr. Lilime, she made a loss of K1 500. How much did Mr. Lilime pay for the refrigerator?

A. K76 500 B.K72 000 C. K74 500 D. K73 500

17. The world record of the highest and lowest temperatures of places are 580ºC above the freezing point of water, in Libya, Africa, and 890ºC below freezing point of water, in Vistol, Antarctica. What is the difference between the two temperatures?

A. -310ºC B. +103ºC C. -1470ºC D. +470ºC

18. The temperature of a metal placed in a very cold place was 30ºC below the freezing point of water. If the metal was warmed until there was a rise of 19ºC, what was the reading on the scale of the thermometer?

A. +19ºC B. +21ºC C. -49ºC D. -11ºC

19. A tobacco company paid 40% of its total annual profit in one year as tax. The following year, tax rate had been reduced to 35% and the company had made an annual profit of K8 500. How much money did the company save from the tax reduction?

A. K4 400 B. K3 025 C. K2 975 D. K425

20. Ilala II cruises at 4 knots per hour. It takes 8 hours to travel from Chipoka to Nkhotakota, 6 hours from Nkhotakota to Chizumulu Island, and 3 hours from Chizumulu to Nkhata Bay. How far is Nkhata Bay from Chipoka in knots?

A. 36 knots B. 72 knots C. 68 knots D. 60 knots

21. Thokozani was riding an ox-wagon. She was helping her father carry maize from their garden to their home. Using a mark on the wheel, she counted the number of turns which the wheel made all the way from the garden to her home. She found that the wheel made 34 turns

and when she measured the circumference of the wheel, she found that it was 1.5 m. What was the distance from her home to the garden?

A. 22 m B. 51 m C. 35.5 m D. 32.5 m

22. Suzgo is a girl who goes to a school that commences at 7:30 am. If she arrived at the school 35 minutes late, at what time did she get there?

A. 8:05 am B. 8:00 am C. 7:55 am D. 8:55 am

23. Chikondi, Mphatso and Kondwani have 18 goats, 15 goats and 12 goats respectively. All these goats are looked after by one person whose pay of K450 per month is a total contribution by the three people according to the number of goats each one has. How much should Chikondi contribute?

A. K150 B. K100 C. K120 D. K180

24. A school purchased 3 drums of paraffin. Each drum contained 250 litres of paraffin. The head teacher wanted to fill the paraffin in 25 smaller tins. How much paraffin would each tin hold?

A. 30 litres B. 10 litres C. 3 litres D. 75 litres

25. Thirty people were asked to clear a piece of land. Each person was given 15 m^2 of land to clear. If the number of people was reduced from 30 to 20, how much land would each clear?

A. 450 m^2 B. 10.5 m^2 C. 300 m^2 D. 22.5 m^2

Appendix B: Test in Simplified English (Form B)

1. The basic salary of an individual is K6 740. If a 10 percent increase is given, how much is the new salary?

A. K7 127 B. K7 324 C. K674 D. K7 414

2. A company pays a commission of K60 for every K800 sales. How much commission does it pay for sales worth K32 000?

A. K240 B. K400 C. K2 400 D. K1 800

3. A school has enough food for 300 pupils for 20 days. How many pupils will eat the same amount of food in 15 days?

A. 400 pupils B. 115 pupils C. 225 pupils D. 40 pupils

4. A boy sold 150 oranges. Find the amount he paid as a market fee if he was charged 20t on each orange sold.

A. K300 B. K75 C. K3 D. K30

5. There are 128 poles equally spaced in a straight line . If the distance from one pole to the next pole is 20 m, calculate the distance form the first pole to the last pole.

A. 2560 m B. 2040 m C. 2540 m D. 2440 m

6. A pot has 70 litres of water. It has a hole where ½ a litre of water is coming out every minute. How much water would be in the pot after one hour?

A. 40 litres B. 35 litres C. 30 litres D. 100 litres.

7. A plot is 14 m long and 12 m wide. Find the cost of fencing it at K70 per metre.

A. K1 400 B. K3 640 C. K1 820 D. K3 540

8. The price for cow is K10,000. Mwale bought it at a discount of 7.5%. How much did Mwale pay for the cow?

A. K750 B. K925 C. K8 250 D. K9 250

9. A boy was born in 1997. When will he become 25 years old?

A. 2012 B. 2022 C. 2032 D. 2015

10. A shopkeeper gives John a commission of 20% on sweets he sells. How much commission does he receive if he sells sweets worth K11 2000?

A. K89 600 B. K1 120 C. K2 400 D. K2 240

11. Find the L.C.M. of 9 cm, 15 cm and 24 cm.

A. 360 cm B. 325 cm C. 125 cm D. 3 cm

12. As Natowa eats 3 pieces of bread, Chimwemwe eats one piece. How many pieces of bread did Chimwemwe eat if both ate a total of 100 pieces?

A. 33 pieces B. 75 pieces C. 25 pieces D. 50 pieces

13. Chitsanzo takes 48 hours to dig a garden. Mayamiko takes 72 hours while Ulemu takes 36 hours to do the same work. How many hours would it take to dig the garden if they were working together?

A. 156 hours B. 52 hours C. 16 hours D. 26 hours

14. Workers receive K700 each in two weeks. How much will two workers receive after 42 days?

A. K2 100 B. K3 200 C.K4 200 D. K8 400

15. Moyo buys a bicycle. Its price is K7000. Tax is charged at 20t in every kwacha. How much tax does Moyo pay?

A. K14 B.K5 600 C. K1 400 D. K3 200

16. Puna bought a garden from the chief at K75 000. She sold it to Ganizani at a loss of K1 500. How much did Ganizani pay for the garden?

A. K76 500 B. K72 000 C. K74 500 D. K73 5000

17. The highest temperature ever recorded is 580˚C above the freezing point of water and the lowest is 890˚C below the freezing point of water. Find the difference between the two.

A. -310˚C B. +103˚C C. -1470˚C D. +470˚C

18. The temperature of a knife is 30˚C below the freezing point of water. Find its new temperature, if it rose by 19˚C.

A. +19˚C B. +21˚C C.-49˚C D. -11˚C

19. A company pays 40% of its profit as tax in one year. In another year, the tax is reduced to 35%. How much money does the company save if it made a profit of K8500 in that year?

A. K4400 B. K3025 C. K2975 D. K425

20. Maria travels at 4 km per hour on a bicycle. It takes 8 hours for her to travel from Blantyre to Zomba, 6 hours from Zomba to Liwonde and 3 hours from Liwonde to Balaka. How far is Balaka from Blantyre?

A. 36 km B. 72 km C. 68 km D. 60 km

21. Thokozani was playing with a wheel whose circumference was 1.5 m. The wheel made 34 turns from Thokozani's garden to her home. Find the distance from Thokozani's garden to her home.

A. 22 m B. 51 m C. 35.5 m D. 32.5 m

22. A school starts at 7:30 am. Suzgo is late by 35 minutes. At what time does Suzgo arrive at the school?

A. 8:05 am B. 8:00 am C. 7:55 am D. 8:55 am

23. Chikondi has 18 goats, Mphatso has 15 goats and Kondwani has 12 goats. If they pay K450 to a worker according to the number of goats each has, how much does Chikondi pay?

A. K150 B. K100 C. K120 D. K180

24. Alinafe has 3 drums of water. Each drum contains 250 litres. Alinafe wants to fill this water into 25 smaller tins. How much water would each tin hold?

A 30 litres B. 10 litres C. 3 litres D. 75 litres

25. There are 30 boys in a class. Each boy slashed an area of 15 m^2. The number of boys has been reduced form 30 to 20 boys. How much area will each boy now slash?

A. 450 m^2 B. 10 m^2 C. 300 m^2 D. 22.5 m^2

End of Questions

Appendix C: Interview Guide

Let the student(s) have a look at the two sets of the test and let him/her read a few selected test items from each set:

1. Which set of questions is easier for you?
    - What things make the items in this set easier for you?
2. Which set of test is easier to read?
    - Why do you think this set is easier to read?
    - Are there some words or any thing that confuses you as a reader?
3. In which set of test items do you think you can perform better? Why?
4. Suppose you did not have enough time for the test, which set would you choose?

- Do you think having more time would make you do well on any of these tests?
    - Why do you think so?

5. Which language(s) do you often speak?

- At home?
- At school?
    - Why is it like this?
- In which language are you comfortable/fluent?
- Why?

6. What advice about language of test items can you give to those who set Mathematics test items? ***ZIKOMO KWAMBIRI***

Appendix D: Sample Interview Transcript

An interview with a pupil at School B
July, 2007 11:00 – 11:20am
[AB = Amicable Banda; P1 = Pupil]

AB: Which pair of questions is easier for you?
P1: I think those in Test Form B,

AB: **What is it about questions in test Form B that make them easier for you?**
P1: The questions are shorter. Although the questions are shorter in test Form B, they contain everything one requires to be able to come up with the correct answer.

AB: **Any other reason?**
P1: Form A has difficult questions because one is made busy to read through long sentences. These long sentences do not help one to come up with the correct answer. 'Key' words necessary for one to come up with the correct answers are available in B.

AB: Is that all?
P1: Yes, you need 'key' words for you to come up with a correct answer, not worthless words, winding sentences which do not benefit you at all.

AB: Which test Form contains questions which are easier for you to read?

P1: Those in Test Form B

AB: Why do you think this set (Form B) is easier for you to read?
P1: The problems are straightforward. If the sum is long, it can easily confuse you.

AB: What else?
P1: When a problem is long or when it has difficult words, it is not easy to follow what is being demanded in it.

AB: Are there some words or anything that confuse you as a reader in test Form A?
P1: Yes, there are.

AB: Like what, can you give some examples?
P1: Words like 'commence', 'silver jubilee', 'purchase'. For me words that are not very strange, like 'silver jubilee, but the tension brought about by examinations caused me to forget it altogether".

AB: In which set do you think you can perform better?
P1: I think it is in B

AB: Why do you think so?
P1: It has straightforward questions which are not difficult to understand

AB: Suppose you did not have enough time for the test, which set would you prefer?
P1: Questions in Form B

AB: Can you give reasons for your choice? Do you think having more time would help you do well?
P1: Shorter sentences do not take much of your time to read through them. Longer sentences will consume much time. It's like going through the forest, not knowing where one is going. Big words only confuse you.

AB: Which language(s) do you speak at home?
P1: At home, we all speak Chichewa.

AB: What about English or any other language?
P1: Yes, I sometimes speak English but not very much.

AB: Why is that so?
P1: I have no one to speak English with at home. All my relatives did not go to school.

AB: Anything more?
P1: English is very difficult, so we do not speak it at home.

AB: In which language are you comfortable?
P1: Chichewa, definitely!

AB: Tell me why it is like that.
P1: This is the language I have known all my life. Everyone at home including my friends speak Chichewa.

AB: What advice concerning language used in Mathematics word items would you give to those responsible for setting the PSLCE exams?
P1: They should think about us. They should set questions that do not confuse us in any way. They should remember that we are primary school pupils. They should not set questions as if they are meant for secondary school learners.

AB: Any thing else?
P1: Item setters should not use difficult words or unnecessary words in questions which just end up confusing us. Reading through long questions takes much of our time.

AB: Zikomo kwambiri, Thank you very much
P1: Zikomo, Thank you.

End of Interview

Appendix E: Formula for testing differences in P-values

The procedure used to determine whether or not the difference between a pair of p values was 'hypothesis testing about differences in proportions, thus: H_o: $P_1 - P_2 = 0$

Ha: $P_1 - P_2 > 0$

Where P_1 = proportion of pupils getting an item correct in test Form B

P_2 = proportion of pupils getting the item correct in test Form A

The first step was to calculate the standard error of the difference between the two proportions (p values), then divide the difference between the p-values by the standard error of the difference to obtain a normal deviate, z which is then referred to a table of areas under the normal curve so as to determine the level of significance (Rivera & Stansfield, 2001; Hinkle, Wiersma & Jurs, 1998). The formulae for obtaining the standard error of the difference between p-values and for obtaining a test statistic or a z are shown below:

1. Standard error of the difference (S_{p1-p2})

$S_{p1-p2} = \sqrt{pq}\ (1/n_1 + 1/n_2)$

where

$p = (f_1 + f_2)/(n_1 + n_2)$

$q = 1 - p$

f_1 = frequency of occurrence in the first sample

f_2 = frequency of occurrence in the second sample

2. Normal deviate (z)

$Z = (p_1 - p_2)/S_{p1-p2}$