Investigating Mathematical Proficiency Testing in Namibian School High Stakes Mathematics Examinations: An exploratory study

A thesis submitted in fulfilment of the requirements for the

DOCTOR OF PHILOSOPHY (PhD) DEGREE
(Mathematics Education)

at

RHODES UNIVERSITY

by

Elizabeth Ndjendja

October 2018

ABSTRACT

The Namibian government has put processes in place to continuously improve its education system in line with educational development in the world. The education reform efforts are administered and coordinated by the Ministry of Education, Arts and Culture. At the centre of these reform efforts are curriculum policy documents and subject syllabuses with the intention of improving the teaching and learning process in the classrooms. These reform efforts appears to overlook the positive influence high stakes assessment has on the teaching and learning process. The study reported in this thesis was designed to investigate the feasibility of assessing separate elements of mathematical proficiency in the high stakes Mathematics assessment in Namibia.

The study was designed as a developmental, exploratory research that collected and analysed both qualitative and quantitative data in order to respond to issues raised by five specific research objectives. The data collected enabled the adaptation of some assessment tools in order to distinctly assess selected mathematical proficiency categories. The results further indicated that the envisaged proficiency assessment system could be used to characterise the examination question papers and revealed insights into the conceptualisation of the current assessment system. The results further indicated the visible distinguishability of different elements of proficiency through the developed tools and the learners' responses to the NSSCO examination. Finally, constrains and affordance which the original assessment system has in relation to the developed system were revealed and addressed.

In closing, the research suggested changes and possible adaptation of assessment tools to ensure the proper assessment of mathematical proficiency aspects through high stakes assessment. Immerging issues that needed further research were also highlighted.

PREFACE

1. Acknowledgements

Firstly, I would like to thank God almighty for his guidance throughout this study. It is because of His unyielding Grace that I was able to complete this study during the most trying period of my life.

Secondly, I would like to acknowledge all the people who helped me in anyway in the process of this study and in ensuring the successful completion of this research project. It pains me to know that I am not able to specifically mention every person who helped me, given this limited space. However, I would like to thank my colleagues at the Directorate of National Examinations and Assessment (DNEA) for their readiness to help at all times. Most specifically, I would like to acknowledge my colleague in the Mathematics sub-division Mr Abednego Ananias for his continuous support and assistance. The Namibia Senior Secondary Certificate Ordinary (NSSCO) level Mathematics examination markers, particularly the selected core group of markers with whom the research took place: without their readiness to help and to share their experience, this project would not have been possible.

I would also like to appreciate and thank my supervisor, Professor Bruce Brown, for his assistance and continuous guidance which made the completion of this study possible. Finally, I would also like to thank all staff of Rhodes University who assisted me in any way throughout this study.

2. Dedication

This work is dedicated to my parents: to my late mother, meme Wilhelmine Albugaste and tate David Philemon Ndjendja for raising me up to be the person I am today. This work is also dedicated to my children Mutua, Mutisya and Kioko who endured the stress as a result of my absence from time to time and tolerated the late meals as I tried to multi-task in order to finish this study.

Table of Contents

CHAPTER 1: INTRODUCTION 1
1.1. Introduction 1
1.2. The Context of the Research 1
1.3. The focus of the study 2
1.4. Personal motivation of the study 4
1.5. The site of the study 5
1.6. The statement of the problem 5
1.7. Research objectives 6
1.8. Outline of the Thesis 7
CHAPTER 2: Educational Assessment Context 8
2.1 Introduction 8
2.2 High stakes examination in Namibia 8
2.2.1 The NSSCO Mathematics examination system 9
2.2.2 Curriculum teaching and learning standards 11
2.2.3 Bloom's Taxonomy 12
2.2.4 Artefacts of the examination system 13
2.2.4.1 The NSSCO curriculum process standards 13
2.2.4.2 The examination question paper and marking scheme 14
2.2.4.3 Learners' answer scripts 15
2.3 Fundamental ideas in assessment 17
2.3.1 Formative assessment 17
2.3.2 Summative assessment 18
2.3.2. 1 Criterion-referenced assessment 19
2.3.2. 2 Proficiency complexity rubric 19
2.3.3 Validity in educational assessment 22
2.4 High stakes assessment 23
2.4.1 The influences of high stakes assessment on teaching and learning 23
2.4.1.1 Providing incentive and motivation to learn 24
2.4.1.2 Teaching to the test 24
2.4.2 Challenges to effective high stakes assessment 25
2.4.2.1. High stakes assessment adaptation challenges 26
2.4.2.2. Stability and constraints of systematic change of high stakes assessment 27
2.5 Conclusion 28

CHAPTER 3: The Review of Literature 29
3.1 Introduction 29
3.2 Theorizing assessment 29
3.2.1 Making exam takers' proficiency visible 29
3.2.1.1. Externalisation of thought 30
3.2.1.2. Vygotsky's culturally motivated activity model 30
3.2.1.3. Externalisation process and criterion based assessment 33
3.2.2 Cognitive complexity 33
3.2.2.1 Textual complexity 34
3.2.2.2 Response mode 35
3.2.2.3 Knowledge processing model 35
3.3 Assessing what is important to assess 36
3.4 Distinguishing elements of knowledge 37
3.4.1 The distinction between conceptual and procedural knowledge 37
3.5 Mathematical Proficiency 39
3.5.1. George Pólya's contribution 40
3.5.2. The National Research Council committee (the contribution of Kilpatrick et al.) 41
3.5.3. The contributions of James Milgram and Alan Schoenfeld 45
3.5.3.1 James Milgram's Contribution 46
3.5.3.2 Alan Schoenfeld's contribution 48
3.6 Including mathematical proficiency categories in high stakes assessment 52
3.7 Conclusion 53
CHAPTER 4: Research Methodology and Design 54
4.1. Introduction 54
4.2. Research Objectives 54
4.3. Philosophical Underpinnings 55
4.4. Research methods and methodologies 57
4.4.1 Design and development research methodology 58
4.4.2 Exploratory research design 59
4.4.3 Qualitative methods 61
4.4.4 Quantitative methods 61
4.5. Research sample 62
4.6. Research instruments 64
4.7. Data Collection methods 64

4.7.1. Adapting the curriculum process standards and generating assessment tools ... 64
4.7.1.1 Generating the adapted process standards ... 64
4.7.1.2 Generating the general proficiency-complexity rubric ... 66
4.7.2.1 Generating specific question paper rubrics ... 68
4.7.3.1 Remarking of the 2013 NSSCO sample of learners' answer scripts ... 69
4.7.3.2 Remarking of the 2014 NSSCO sample of learners' answer scripts ... 71
4.7.3.3 Remarking the remarked sample of 2013 and 2014 learners' answer scripts 71
4.8. Data Analysis ... 72
4.8.1.1 Literature Analysis ... 73
4.8.1.2 Textual Analysis ... 73
4.8.2 Distinguishing different elements of proficiency ... 74
4.8.3 Generating insight into the current examination system ... 76
4.9. Validity and reliability ... 78
4.10. Ethical consideration ... 79
4.11. Limitations of the study ... 80
CHAPTER 5: Research Findings and Analysis: Proficiency Categories and Assessment Tools ... 82
5.1 Introduction ... 82
5.2 Adapting the curriculum process standards and other assessment tools ... 82
5.2.1 Mathematical proficiency categorisation for the study ... 83
5.2.2 Generating the adapted process standards ... 86
5.2.3 Development of the proficiency-complexity rubric ... 91
5.2.3.1 Cognitive complexity scale for the study ... 96
5.3 Distinguishability of different elements of proficiency ... 100
5.3.1 Chi-square test comparison of the distinguishability of knowledge base and strategy categories ... 101
5.4 Conclusion ... 104
CHAPTER 6: Research Findings and Analysis: Investigate the principles and relationship of proficiency and Mathematics competence ... 105
6.1 Introduction ... 105
6.2 The possible distinction of proficiency in the learners' responses ... 105
6.2.1 Analysis of the full sample: First order effects ... 107
6.2.1.1. The whole sample means: comparison per paper ... 107
6.2.1.2. The full sample correlation and regression analysis ... 109
6.2.2 Analysis of key grade groups: Second order effects ... 109
6.2.2.1. Comparing the means of key grade groups ... 110

6.2.2.2. Correlation and regression analysis within key grade groups 110
6.2.3 Further analysis on the relationship between knowledge base and strategy in the learners' results 117
6.2.3.1. The difference between knowledge base and strategy in the learners' performance 117
6.2.3.2. The learners' performance in these categories at different levels of complexity 121
6.2.4 General observations of the categories distinguishability in the learners' results 125
6.3 Conclusion 127
CHAPTER 7: Research Findings and Analysis: Creating a better understanding of the current system 128
7.1. Introduction 128
7.2. Insight into the conceptualization of the current examinations 128
7.2.1 Percentage distribution comparison: Proficiency Balance 128
7.2.2 Percentage distribution comparison: Complexity Balance 131
7.3. Constraints of the current examination system and the possibilities of overcoming them 134
7.3.1 Constraints identified through markers' experience 135
7.3.2 Possibilities for overcoming these constraints 138
7.3.2.1. Question papers' construct 139
7.3.2.2. Question construct 139
7.4. Conclusion 141
CHAPTER 8: Discussion of the Research Findings 142
8.1 Introduction 142
8.2 The feasibility of deepening Mathematics high stakes assessment 142
8.2.1 Understanding the distinct aspects of Mathematics assessed by the NSSCO examination 143
8.2.2 Extending what Mathematics is assessed by the NSSCO Mathematics examination 145
8.3 The relationship between Mathematics capability and mathematical proficiency .146
8.3.1 The first order effect 146
8.3.2 The second order effect 147
8.3.3 Important contributions to assessment and Mathematics teaching 148
8.4 Insight into the NSSCO examination system 149
8.4.1 Mathematics Proficiency as an assessment framework base 150

8.4.2 Assessing valued aspects of mathematical proficiency not just what is easy to assess 152
8.5 Conclusion....154
CHAPTER 9: Research Conclusion....155
9.1 Introduction....155
9.2 The summary of the study's context and process....155
9.3 The research findings....155
9.3.1 Identifying assessable proficiency categories....156
9.3.2 Generating assessment tools....157
9.3.3 Distinction of proficiency aspects....157
9.3.4 Understanding distinction through learners' results....158
9.3.5 Understanding the conceptualization of current examinations....159
9.3.6 NSSCO constraints and adaptations....161
9.4 The research recommendations and the way forward....162
9.4.1 Minor adaptations to the NSSCO Mathematics examination questions....163
9.4.2 Areas for further research....164
9.5 Limitations of the study....165
REFERENCES....166
APPENDICES....175
APPENDIX A NSSCO Mathematics curriculum standards....175
APPENDIX B NSSCO Mathematics content standards....176
APPINDIX C Key grade descriptors....182
APPENDIX D Key grade descriptors – aligned with algebra....184
APENDIX E Standards translation table....185
APPENDIX F Expanded process standards....191
APPENDIX G General complexity scoring rubric....193
APPENDIX H Markers comment summary – round one....195
APPENDIX I Markers comment summary – round two....198
APPENDIX J Focus group members comment form....202
APPENDIX K Consent form....204
APPENDIX L Permission from the Ministry's Permanent Secretary....206

List of tables

Table 2.1 Brookhart & Nitko (2007) top-down approach (p.195).....20
Table 2.2 Brookhart & Nitko (2007) bottom-up approach (pp.195-196).....21
Table 2.3 Steps of the process followed to develop an analytic rubric (Mertler, 2001, p.8)...21
Table 5.1 Summary of the Emerging Mathematical Proficiency Standards.....86
Table 5.2 An extract of process standards translation working table 88
Table 5.3. Possible gaps between process standards and proficiency standards 91
Table 5.4. Process standards aligned to the content standards 93
Table 5.4 Examples of Mathematical Proficiency Standards and Examination Question Prototypes developed.....94
Table 5.6 Specific scoring rubric of paper 2 for 2013 for question 1 (a). 100
Table 5.7 Chi-squared test comparison for the two categories for the year 2013 102
Table 5.8 Chi-squared comparison for the two proficiency categories for the year 2014.....103
Table 6.1. The recording sheet for Paper 3, 2014.....106
Table 6.2. Mean marks for the two set of marks compared for both 2013 and 2014.....107
Table 6.3. Mean percentage marks for the two categories per paper 108
Table 6.4. Full sample correlation analysis 109
Table 6.5. Mean percentage marks for the year 2013.....110
Table 6.6. Key grade groups correlation analysis.....114
Table 6.7. Mean of the difference between knowledge base and strategy for 2013 and 2014 120
Table 7.1 Marks percentage per proficiency category for the four papers of 2013 and 2014. 129
Table 7.2 Percentages of scores testing different proficiency categories per examination level 130
Table 7.3 Percentages distribution of marks per proficiency categories within a complexity level for each paper.....132
Table 7.4 percentage distribution of marks per complexity level within each proficiency category for each paper.....133
Table 7.5 question 9, paper 4 2014.....136

List of figures

Figure 2.1. A general structure of the Namibian Mathematics examination system 10

Figure 3.1 Vygotsky's model of mediated act as adapted for this study (adapted from Bazerman (2012)..... 31

Figure 4.1. The adapted proficiency examination system 57

Figure 4.2. The developmental research process 58

Figure 4.3. The stages of the developmental phase 67

Figure 6.1(a). Scatter plots for all papers for the year 2013 111

Figure 6.2(a). Scatter plots of the difference between knowledge base and strategy for 2013 118

Figure 6.2(b). Scatter plots of the difference between knowledge base and strategy for 2014 118

Figure 6.3(a). Scatterplots of knowledge base and strategy scaled per level for paper 2, 2013 122

Figure 6.3(b). Scatterplots of knowledge base and strategy scaled per level for paper 2, 2014 122

Figure 6.4(a). Composite scatterplots scaled per level for paper 2, 2013 and 2014 123

Figure 6.4(b). Composite scatterplots scaled per level for paper 4, 2013 and 2014..... 123

Figure 6.5. Key grade scatterplots scaled per level for paper 2, 2013. 124

Figure 7.1. Marker Ma7's comment on testing high level competencies and skills..... 138

Figure 7.2. Marker Ma1's comment on the demands of the questions..... 140

Acronyms

DNEA	Directorate National Examinations and Assessment
CIE	Cambridge International Examinations
JSC	Junior Secondary Certificate
MEAC	Ministry of Education Arts and Culture
MEC	Ministry of Education and Culture
MoE	Ministry of Education
NIED	National Institute of Educational Development
NSSC	Namibia Senior Secondary Certificate
NSSCH	Namibia Senior Secondary Certificate Higher level
NSSCO	Namibia Senior Secondary Certificate Ordinary level
PIRLS	Progress in International Reading Study
PISA	Programme for International Student Assessment
SACMEQ	Southern Africa Consortium for Monitoring Educational Quality

CHAPTER 1: INTRODUCTION

1.1. Introduction

This chapter introduces the study reported in this thesis. It starts by describing the context and moves on to outlining the focus and the reasons for undertaking the study. Next, the research objectives of the study are identified. Finally, the chapter gives a brief overview of the other chapters.

1.2. The Context of the Research

High stakes examinations are seen in general as a tool for setting or reinforcing required curriculum standards. Hence, these examinations inevitably drive the teaching and learning process in classrooms (Leyendecker, 2008). According to Schoenfeld (2007a), in order to support teaching, high stakes mathematics examination tasks need to;

(i) "capture the kind of learner's mathematical thinking that is important,

(ii) act as levers for positive change, redirecting attention to Mathematics instruction and helping to raise the teaching standards" (Schoenfeld, 2007a, p. ix).

Assessment criteria used for high stakes examinations are generally based on cognitive demand, as described by Bloom's Taxonomy (Smith, Wood, Coupland, Stephenson, Crawford & Ball, 2010). Bloom proposed a hierarchy of different levels of cognitive complexity. In this hierarchy the lower levels of demand are assumed to have been subsumed by the higher levels (Krathwohl, 2002, and Smith et al., 2010). Bloom's taxonomy has proven to be effective in structuring assessment tasks: it ensures that the competencies and standards that are set are measurable (Smith et al., 2010) and caters for a varied ability of learners. However, this taxonomy is found to have some limitations when applied to the assessment of learners' mathematical capabilities (Smith et al., 2010). High stakes examinations for Mathematics, based on Bloom's taxonomy are claimed to be too general, in that:

- They do not allow learners to demonstrate that they are capable of thinking mathematically and to show their mathematical capabilities (Krathwohl, 2002; Yunus, Razak, Khairani, 2012).
- The way they influence Mathematics teaching does not necessarily result in the teaching envisioned to develop learners' Mathematics proficiency (Schoenfeld, 2007a).

The Namibia Senior Secondary Certificate Ordinary (NSSCO) level Mathematics examination, which is the focus of this study, is one of Namibia's high stakes examinations at the end of grade 12. Over the years, it has been observed that learners who completed grade 12 and graduated from their basic education schooling through the NSSCO level Mathematics examination, were not able to cope with the mathematical demands outside of the schooling system. In particular, when they join institutions of higher learning and also the job market, they seem to require additional bridging courses. They are said not to be sufficiently proficient in Mathematics (Iipinge, 2012).

This study was motivated by the concerns mentioned above as related to the Namibian high stakes Mathematics examination system. It investigatespd the possibility of adapting particular components of the Namibian Mathematics curriculum and assessment documents in order to assess the grade 12 learners' mathematical abilities and capacities more effectively through the high stakes examination system. Within this context, the study focused on gaining an understanding of how a more nuanced conception of mathematical proficiency could be incorporated into the high stakes NSSCO Mathematics examination.

According to Ottevanger, van den Akker & de Feiter (2007) assessment has proven to be a critical part of all education systems in the world. In light of this, Burkhard's (2007) highlighted that one of should be key assessment principles is to "measure what is important, not just what is easy to measure" (p.78). Thus, improved understanding on how to assess mathematical proficiency through high stakes examinations, may provide the chance for high stakes Mathematics examinations to benefit from such conceptions.

1.3. The focus of the study

The importance for every world citizen to learn Mathematics is now more apparent than ever before. The demand for every nation to drive an industrialized and knowledge based economy does not only increase the demand for the citizenry to have learned Mathematics, but to be proficient in it (Namibia. Ministry of Education [MoE], 2006). Many countries in the world have transformed their school curriculum to heed the call for reformed Mathematics teaching. Despite the fact that assessment drives the teaching process in the classroom, the reformulations of the curriculum standards have often put more emphasis on the teaching

process than the assessment, particularly the high stakes assessment processes (Kellaghan & Greaney, 2003).

The possible improvement on assessing distinct proficiency categories at varied complexity levels could help to inform the development of assessment tools that would assess what is important to assess, rather than what is easy to test. In addition, the outcome of this study would benefit the teaching of Mathematics, through the modification of high stakes assessment tools in ways that support the assessment of more complex mathematical proficiencies. The possible improvement in the visibility of learners' proficiency in their examination answers might help to improve teaching for proficiency since teachers may strive to teach what is expected in the examinations.

Measuring all aspects of mathematical proficiency is accepted in the assessment literature Burkhard (2007) as a difficult problem, that would require further research and deeper understanding conducted using more effective technical tools (Schoenfeld, 2007a). Firstly, this study focused on developing an understanding of how curriculum standards and other assessment tools could be adapted to incorporate proficiency insight. Secondly, the study investigated the interrelation between the adapted tools and the current question papers used in the NSSCO Mathematics examination system. The study then investigated what these adapted tools would offer in relation to the deepening insight into learners' mathematical proficiency through high stakes assessment. Thus, the feasibility of using the tools developed to make several aspects of mathematical proficiency more visible through high stakes examinations was investigated. Lastly, the study investigated constraints and possibilities for improving the assessment of distinct aspects of proficiency in the examination system. That is, the possible need for adaptation of the NSSCO Mathematics question papers to be in line with the developed tools to allow for a more effective assessment of all elements of mathematical proficiency was investigated.

A professional motive for focusing the study on the possibility of assessing mathematical proficiency through high stakes assessment, is the potential which a high stakes examination has of encouraging schools and teachers to teach to the test. In line with Burkhard's (2007) operative principle of What You Test Is What You Get (WYTIWYG), teaching to tests that are designed to test high quality mathematical skills and knowledge would encourage teaching high quality Mathematics. High stakes examinations have proven to be crucial to most

education systems especially in Africa, in that they have the potential to provide incentives for learning (Kellaghan & Greaney, 2003). The argument here is that, if the examination questions require learners to demonstrate mathematical proficiency and further test a deeper understanding of Mathematics and engagement in the mathematical process, then teaching to these tests would ensure the teaching for mathematical proficiency as promoted by the Mathematics education reform movement. Thus, the ultimate goal of the study was to investigate ways of enhancing Mathematics high stakes assessment in a way that would improve the teaching for mathematical proficiency.

1.4. Personal motivation of the study

As mentioned in earlier sections, the NSSCO examination is one of the Namibian high stakes assessments administered and conducted under the supervision of the Directorate National Examinations and Assessment (DNEA) within the Ministry of Education Arts and Culture. The researcher has been working for the DNEA for the past twelve years. As a Chief Education Officer for the Science and Mathematics sub-division her responsibilities are, amongst others, to ensure that the examination question papers for Mathematics and other Science subjects are ready for every examination session. This entails supervision and monitoring of the setting and moderation of the question papers. The administration and the supervision of the marking of the learners' answers scripts for the examination also falls under the researcher's duties at the DNEA. Although these activities are guided by the syllabus, the researcher is interested in possible development of what content is tested, how the questions are designed and how the learners' responses are marked in order to separately and appropriately assess the full range of the learners' mathematical proficiency.

Additionally, the researcher has been a member of the national Mathematics curriculum panel, at the National Institute for Educational Development (NIED) since 2004. The curriculum panel is a ministerial body formed under a Namibian Act of Parliament, Education Act number 16 of 2001 tasked with the development of school Mathematics syllabuses (Namibia. MoE, 2006). It also recommends changes to the syllabus every five years. The members of this body are committed to the on-going development of the Mathematics curriculum in a move to improve Mathematics education in the country.

Accordingly, the researcher's position in the Ministry has enabled her to have a clear understanding of what Mathematics is taught in Namibian schools and assessed through high stakes examinations. Her position in the Ministry has also given her a better understanding of the general strengths and weaknesses of the Namibian Mathematics high stakes examination. In particular, she is cognisant of the general strengths and weaknesses of all stages involved in conducting examinations, starting from the design of the tools to the implementation and administration of these examinations. Hence, she has a good understanding of the power high stakes examinations could have on informing the teaching and learning process through direct and indirect feedback systems as schools try to benchmark themselves against what is generally presumed to be the best standards. It is this understanding that motivated and drove the researcher to investigate the possibility of improving the testing of mathematical proficiency through NSSCO Mathematics high stakes examination in an attempt to better inform Mathematics classroom teaching.

1.5. The site of the study

The study was conducted within the NSSCO Mathematics examination operational structure. The adapted tools were; the process standards and the marking schemes. The other tools which were used were; the question papers and the learners' answer scripts formed part of the existing or current examination system which is conducted annually. Moreover, the participants in the study, the focus group, are current NSSCO Mathematics examination markers. The participants' meeting site was the DNEA office board room.

1.6. The statement of the problem

As mentioned earlier, there have been complaints and criticism that NSSCO Mathematics graduates are not sufficiently proficient to enter the job market, or to cope with Mathematics offered at the institutions of higher learning, in particular at the University of Namibia (UNAM) (Iipinge, 2012).

Through her position in the Ministry of Education, the researcher plays a major role in school Mathematics education in the country. All stakeholders in education in Namibia view national examinations as a tool for benchmarking the desired standards (Namibia. MoE, 2006). The Ministry therefore has a responsibility to ensure that the NSSCO Mathematics assessment activities continue to become better aligned with the development in education, particularly

with Mathematics education. The drive to ensure that school leavers become more proficient in Mathematics was one of the major moves on which the Namibian government embarked so as to ensure adequate development of the people for a knowledge-based society (Namibia. MoE, 2006). However, any changes to improve alignment have to be controlled and properly managed, resulting in the need to research the possibility and the process for such change.

The problem identified for consideration in this research, was to find ways to adapt the current high stake mathematics examinations Namibia to better evaluate the broad range of characteristics required for mathematical proficiency.

1.7. Research objectives

The study was designed with the general goal of gaining an understanding of how the assessment of mathematical proficiency may be improved in high stakes assessment, particularly in the case of the NSSCO Mathematics examination.

The study was a design and development project whose process was directed to achieve three aims, each of which is subdivided into a number of objectives. These aims and objectives are given as follows:

1. To investigate the possible design and adaptation of curriculum and assessment tools to incorporate conceptions of proficiency. This included the following objectives:

(a) To identify appropriate proficiency categories in the literature relating to mathematical proficiency.

(b) To investigate the feasibility of adapting the curriculum process standards and proficiency rubrics in order to more effectively assess the identified categories of mathematical proficiency in the NSSCO high stakes assessment.

(c) To investigate the distinguishability of the identified elements of proficiency, by using the generated assessment tools on the sample of recent examination question papers.

2. To use the designed tools to investigate the distinguishability of different proficiency elements in the learners' responses to the current question papers.

3. To use the designed tools to generate insight into the current examination system. This included the following objectives;

(a) To generate insight into the conceptualisation of the current examination, using the identified proficiency categories and developed tools.

(b) To investigate the constraints and affordances of the current examination system in relation to the identification and presentation of each of the identified proficiency categories, as well as the possibility of overcoming these constraints by adapting question papers.

1.8. Outline of the Thesis

The thesis consists of nine chapters. This chapter introduces the study. Chapter 2 places the study in the context of educational assessment. Chapter 3 is the review of literature on high stakes assessment and the literature on mathematical proficiency, in particular the literature that leads to a better understanding of how to assess mathematical proficiency through high stakes assessment. Chapter 4 outlines the methodology and research approach. It gives the grounds and justification for carrying out the research in the way it was done. Chapters 5, 6 and 7 present the research findings and their analysis. Chapter 8 discusses and reflects on the data, analysis and literature and draws salient issues and meanings from them. Chapter 9 gives a summary of the whole study and presents the conclusions, recommendations and areas for further research.

CHAPTER 2: Educational Assessment Context

2.1 Introduction

This chapter places the study in the context of educational assessment. It starts by giving a brief discussion of the Namibian educational system in relation to the Namibia high stakes assessment activities. Additionally, the chapter gives a brief overview of the fundamental concepts of assessment which provide the basic conceptual context for the research.

2.2 High stakes examination in Namibia

The study was conducted during the time when Namibia's education system was undergoing yet another major change in the curriculum process. As was usually done, the Ministry of Education Arts and Culture curriculum review activities were geared to ensure universal access to quality education for all Namibian children (Namibia. MoE, 2006). This was done as Namibia strives to ensure the achievement of its broad developmental goal of having a knowledge based economy. The Namibian broad curriculum for basic education and the subject syllabuses have been central to the review process. The broad curriculum was developed to provide a coherent framework which serves as the official curriculum policy for Formal Basic Education in Namibia. The framework guides schools in organising teaching, learning and assessment processes, in order to ensure that there is consistency in curriculum delivery throughout the country (Namibia. Ministry of Education [MoE], 2010a).

The broad curriculum is based on a learner-centred paradigm with criterion or standard-based assessment (Namibia. MoE, 2010a). Learners are seen by Katic, Hmelo-Silver & Weber (2009) as creating knowledge, based on their prior knowledge and experience and not just passively receiving information. Assessment activities form part of every review process: the major move for this was a change from norm referenced to criterion referenced assessment (Namibia. Ministry of Education and Culture, [MEC], 1993). The assessment activities which were of particular interest to this study were the national examinations at the end of the current Grade 12. These national examinations are high stakes examinations administered by the Ministry of Education Arts and Culture (MEAC) through the Directorate National Examinations and Assessment (DNEA).

High stakes examinations, according to Kellaghan & Greaney (2003) are generally called public or national or external examinations because they are conducted by external institutions at the national level. These institutions are often government departments or organisations independent from the government (Kellaghan & Greaney, 2003). The examinations are conducted and function through the development of artefacts. The artefacts include the curriculum teaching and learning standards which are comprised of process standards and the content reflected in the teaching syllabuses. The examination question papers, marking schemes and learners' answer scripts are also part of the artefacts produced for every examination.

Namibia has annual national examinations at two grades, the Junior Secondary Certificate (JSC) examination at the end of Grade 10 and the Namibian Senior Secondary Certificate (NSSC) examinations at the end of Grade 12 (Namibia. MEC, 1993). The NSSC curriculum is taught and examined at two levels: the Namibian Senior Secondary Certificate Higher level (NSSCH) and the Namibian Senior Secondary Certificate Ordinary level (NSSCO). The NSSC examinations at both levels are accredited by the University of Cambridge International Examination (CIE) department. The NSSC syllabi were developed as a result of the localisation of the Higher International General Certificate of Secondary Education (HIGCSE) and the International General Certificate of Secondary Education (IGCSE). The HIGCSE and IGCSE were fully administered by CIE from 1995 to 2006. As a result, the examinations for the NSSC were written for the first time, in most subjects, in Namibian schools in 2007. These examinations are conducted by the Ministry of Education, Arts and Culture (MEAC), under the direct supervision of the Directorate National Examinations and Assessment (DNEA). This study focuses on the NSSCO Mathematics examinations.

2.2.1 The NSSCO Mathematics examination system

In general, an examination is defined as a test given to the learners at the end of a course in order to examine what the learners have learned (Green, 2002; Earl, 2003; Sadler, 2005 and Namibia. MoE, 2010a). The Namibian Mathematics examination system is guided by a set of tools as given by the curriculum and subject syllabus, shown in Figure 2.1. The overarching tool of any examination system is the teaching and learning standards document, which defines the cognitive processes to be applied to the subject content standards, given in Figure 2.1 as subject content objectives. The examination is often given using one or a combination of examination question papers, depending on the requirement of the curriculum. The

examination question papers are designed and compiled annually based on the process standards and the content standards of the subject.

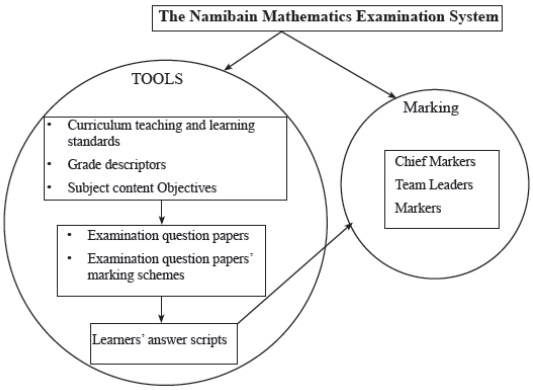


Figure 2.1. A general structure of the Namibian Mathematics examination system

The NSSCO Mathematics examination is written at two different examination levels, Core level and Extended level. Learners at each level take two written papers during the examination, depending on the examination level they chose to follow. Learners who register their examination at the Core level write Paper 1, a question paper with questions which demand short answers, and Paper 3, a question paper with questions which demand structured answers. Learners who have registered for the Extended level take Paper 2, a question paper requiring short to semi structured answers, and Paper 4, a question paper with structured questions (Namibia. Ministry of Education [MoE], 2010b). The sum of the marks for Paper 1 and Paper 3 make up the Core examination level marks and the sum of the marks for Paper 2 and Paper 4 make up the Extended examination level marks.

According to the NSSCO Mathematics syllabus, the short-answer questions fulfil an important function in ensuring syllabus coverage and allow for the testing of knowledge, understanding and manipulative skills (Namibia. MoE, 2010b). From experience, manipulative skills, in this regard, involve one or two step calculations, which are awarded one or two marks in the NSSCO system of mark allocation. Papers 3 and 4 are designed to contain questions that demand structured answers. In other words, learners have to clearly show the method they followed in order to work out their answers. A greater emphasis is placed on applications to the processes of problem solving in the structured answer papers (Namibia. Ministry of Education [MoE], 2010c). Just like every school examination, the NSSCO Mathematics examination is guided by the teaching and learning standards.

2.2.2 Curriculum teaching and learning standards

Seguin (1993) defines the curriculum as an educational scheme which defines the goals, aims and objectives of educational action. Simultaneously, the curriculum defines the way, means and activities employed in order to achieve these goals as well as the methods and instruments required to evaluate the success of the action. Curriculum has been defined by Beaco, Byram, Cavalli, Coste, Cuenat, Goullier and Panthier (2010) simply as a path for learning. The path which is paved with the help of assessment and also ends with assessment. Learners' assessments are focused on whether they have attained the objectives described in the curriculum standards (Beaco et al, 2010). The curriculum standards are generally developed to specify both the cognitive processes and the subject content with which these processes are to be aligned (Brookhart & Nitko, 2007). Subject content includes the fundamental subject concepts and topics and is aligned with the process described by the standards. The aim of these standards is to state what knowledge and skills a learner would demonstrate in an examination as the result of the teaching and learning process. Learners are required to demonstrate these standards in the context of the content and skills prescribed by the specific content objectives. The specific content objectives should be clustered according to the subject topics. For Mathematics, examples of the topics within which the content objectives are: algebra, geometry etc., while cognitive processes include the actions associated with using and applying the mathematical content to solve both routine and non-routine problems (Clarke, Goos & Morony, 2007).

According to the 2012 – 2017 Strategic Plan of the Namibian Ministry of Education, the verbs used to frame and direct standards statements, should be action verbs that describe an explicit behaviour that can be observed and measured in order to ensure that the competences achieved are made clear (Brookhart & Nitko, 2007; Namibia. Ministry of Education [MoE], 2012). The development of learning and assessment standards should therefore be done concurrently, resulting in a single product. Learners need to know what the assessments expect from them through classroom learning using the same standards for both teaching and assessment processes (Brookhart & Nitko, 2007; Namibia. MoE, 2012). The teachers' understanding of the learning targets set through the learning standards guides their teaching. This should therefore provide a criterion for evaluating or assessing learners' progress and achievement of the required knowledge and skills. It is for this reason that care should be taken when formulating learning and assessment standards. Brookhart & Nitko (2007), have laid out a process to be followed in the development of learning and assessment standards. The

development of learning and assessment standards should be informed by *general educational goals*, followed by developing what they term as *general learning targets* and then *specific learning targets* (Brookhart & Nitko, 2007, pp.49). According to Smith et al (2010) many high stakes examinations' standards in the world, just like the NSSCO examinations are largely based on cognitive demand, as described by Bloom's Taxonomy.

2.2.3 Bloom's Taxonomy

Bloom's Taxonomy has been used to frame many assessment models in the world since its publication in 1956 (Yunus, Razak & Khairani, 2012). One of the objectives of doing this, according to Yunus et al. (2012), is its usability in clarifying curriculum objectives and examination questions in order to differentiate between learners' abilities. In his work of 1956, Benjamin Bloom developed a classification of levels of rational behaviour which are important for learning. His categorisation has come to be known as Bloom's taxonomy and it is commonly used in education assessment in various countries including Namibia. Bloom's taxonomy has three domains of educational activities; cognitive (mental skills or knowledge), affective (attitudes) and psychomotor (physical skills) (Krathwohl, 2002). The cognitive domain is further divided into six different levels of complexity, starting with the lowest of recalling knowledge and facts followed by comprehension, application, analysis, synthesis and lastly evaluation (Krathwohl, 2002).

Bloom proposed a hierarchy in the different levels of cognitive demand and the lower levels of demand are assumed to have been subsumed by the higher levels (Krathwohl, 2002; Smith et al., 2010). This taxonomy has proven to be effective in structuring assessment tasks; because it ensures that the competencies and standards set are measurable (Smith et al., 2010). However, when used for assessment in Mathematics curricula, this taxonomy seems to have limitations (Smith et al., 2010). High stakes examinations for Mathematics, based on Blooms taxonomy, are claimed to be too general, in that:

- A notable weakness is the assumption that cognitive processes are ordered on a single dimension of simple-to-complex behaviour (Krathwohl, 2002). This indicates that the mastery of a more complex category requires prior mastery of all the categories below it. Such a hierarchy is inflexible, which may be a disadvantage, for example, it cannot adapt to the observation that certain demands for knowledge are more complex than certain demands for analysis or evaluation.

- They do not allow learners to demonstrate that they are capable of thinking mathematically and to show their proficiency in Mathematics (Krathwohl, 2002; Yunus et al., 2012).
- The way in which they influence Mathematics teaching does not necessarily result in the teaching envisioned by the changes in Mathematics education (Shannon, 1999).

In order to support teaching for mathematical proficiency, high stakes examination tasks need to be designed in such a way that they:

(i) "capture the kind of learner's mathematical thinking that is important,

(ii) act as levers for positive change, redirecting attention to Mathematics instruction and helping to raise the teaching standards" (Schoenfeld, 2007a, p. ix).

2.2.4 Artefacts of the examination system

Just as any other examination system, the NSSCO functions through the production of a series of artefacts. These are: the curriculum process standards, the examination question papers, the examination marking schemes, and the learners' examination scripts. Each of these artefacts and their relationships are briefly described in the following subsections.

2.2.4.1 The NSSCO curriculum process standards

The requirements of the examinations are described by the specification of standards, which according to NSSCO Mathematics syllabus include: assessment objectives, content objectives and key grade descriptors (Namibia. MoE, 2010b). The assessment objectives describe the process to be displayed by the learner when solving a mathematical problem. Thus, for this study they were termed process standards. The NSSCO process standards were based on cognitive demand as framed by Bloom's Taxonomy (Namibia. MoE, 2010a & b). The NSSCO Mathematics curriculum has summarised Bloom's taxonomy and developed a series of descriptive assessment objectives which are attached as Appendix A. The assessment objectives describe the desired cognitive process to be assessed in the NSSCO Mathematics examination, in the form of techniques with applications (Namibia. MoE, 2010b). The standards help to describe how the subject content objectives, contained in the syllabus, should be taught and learned in the classroom. They also describe the process to be demanded by the questions and displayed by the learners through assessment in the NSSCO national examination framework.

The key grade descriptors, on the other hand, describe the criteria to be used in the judgment of the key grades A, C, E and G and indicate the standards likely to be achieved by the learners awarded such a grade (Namibia. MoE, 2010 b). Grades A, C, E and G are the grades that are usually used by the institution of higher learning for selection purposes, hence they are termed key grades in the NSSCO syllabuses (Namibia. Ministry of Education [MoE], 2012). The key grade assessment standards provide the general scoring rubric for the NSSCO Mathematics syllabus (Namibia. MoE, 2010a). The descriptors in this rubric are holistic by nature in that they provide a general indication of the standards of achievement that are likely to have been shown by learners awarded particular grades (Namibia. MoE, 2010a). The descriptions are to be interpreted in relation to the subject content specified by the NSSCO Mathematics syllabus but are not designed to define that content (Namibia. MoE, 2010b). The scores reported using a holistic rubric, concentrate on the overall proficiency of the learner and are thus unidimensional (Mertler, 2001; Brookhart & Nitko, 2007). Due to their single dimensional nature they provide limited feedback to the process of teaching and learning. The more comprehensive the feedback provided by the high stakes assessment, the more the chances are that this assessment would influence the teaching and learning process. It would be advantageous if it could provide categorised, multi-dimensional feedback, bearing in mind the complexity of the mathematical proficiency that a learner is expected to achieve.

2.2.4.2 The examination question paper and marking scheme

The examination question paper is defined as a paper which contains examination questions printed on it which are set to test the knowledge, understanding and skills of an examination taker (Black, Suto & Bramely, 2011). An examination question paper is a booklet containing a series of questions to which the examination takers respond. The questions contained in the question papers are based on the process and content standards described in the subject syllabuses. The question papers are normally made up of different types of questions. The types of questions to be discussed in this study were grouped into two main groups.

- **a.** Questions which demand objective answers
- **b.** Questions which demand subjective answers

Questions which demand objective answers are questions which demand short answers while semi-structured to structured answers question types demand subjective answers, whose degree of subjectivity differs with the specific demand of the question (Namibia. MoE, 2010c). The questions which demand short answers are often questions which demand the learners'

knowledge of concepts and one- or two-step procedures, while the questions which demand structured answers are questions which demand multiple step answers.

The NSSCO Mathematics question papers consist of several questions which demand short answers to structured answers on various topics as shown in the syllabus. Some structured questions are broken down into several sub-questions Namibia. MoE (2010c) to allow learners better access to what is asked (Burkhardt, 2007). The number of marks to be awarded for a full correct answer, is given at the end of each question or sub-question and reflects the knowledge and skills demanded by the question.

The DNEA appoints examiners to set the examination question papers annually. The examiners are selected from the teachers who teach the subject at the NSSCO level (Namibia. MoE, 2010c). A teacher to be selected as an examiner needs to have at least five years teaching experience and three years national examination marking experience. The examiners use the syllabus standards to set the question papers and the corresponding marking schemes (Namibia. MoE, 2010c). The two documents (question paper and marking scheme) are then moderated (Namibia. MoE, 2010c) to ensure the validity of what is assessed.

The marking scheme is an instructional document which specifies how marks must be awarded. It is a document with the answers to the questions contained in the question paper. The marking scheme reflects the expected correct answer by indicating the ideal answer required as well as acceptable alternatives (Namibia. MoE, 2010c). Usually, the marking scheme is developed by the appointed examiner at the same time as the question papers. The ideal answers, which are contained in the marking schemes are formulated based on the competencies described by the process and subject standards. The purpose of the marking scheme is to define criteria that are used to score the learners' responses to the question paper. Thus, it is developed to ensure consistency of the marking process over a range of answers and by a number of markers. If learners do not present their answer as shown in the mark scheme, then partial credit can be given, guided by the criteria described in the marking scheme.

2.2.4.3 Learners' answer scripts

A learner's answer script refers to the paper where a learner has written his or her responses to the questions contained in the examination question paper (Black et al., 2011). In the case of

NSSCO Mathematics examinations the learners' responses are written in the spaces provided in the question papers. The learners' responses contain different features, some learners showing more features than the others. All NSSCO Mathematics question papers are designed with clear instructions on the cover page, demanding the learners to show their working towards a solution for each question. However, some learners tend to give only the final answer. The answer scripts are then marked to award each learner a score, which can be translated into a grade.

The NSSCO Mathematics learners' answers scripts are usually marked through a centralised marking system. The marking exercise is coordinated by the nation's examination body, the DNEA, which requires services from the subject teachers in the country.

The learners' responses to the questions are marked according to the criteria set in the marking schemes. Marking is the act of judging a learner's response and gives an indication with either a tick (✓) to indicate a response (or response part) that aligns well with the requirements set in the criteria, or a cross (✗) to indicate a response (or response part) that does not align with these requirements. This judgment makes it possible to award marks (numerical values) to learners' responses in order to create a scale for such an examination (Pollitt, Elliott & Ahamed, 2004). Criterion referenced marking was developed to address the concern that holistic judgement was unsafe in a time when there was little chance to control the quality of the judges (Pollitt et al., 2004). It grew with concerns for objectivity and transparency and an adherence to procedure to ensure reliability of the process rather than validity. As a result of the marking exercise, the learner's script is awarded a score or a mark which is usually a count of ticks, which may be compared to the expected total score (mark) to assign the learner a grade.

Grading refers to the process of determining the cut off marks, for each key grade on each subject's question paper. This is performed to appropriately define the achievement of learners awarded those marks as compared to the standards determined in previous years (Namibia. MoE, 2012). The lowest mark for each key grade on each paper and each level, is determined to define the learners' performance in a particular year which can be deemed equivalent to that achieved by learners at the lowest mark for the same grade the previous year, having made any necessary adjustments for evident disparities in the syllabus in the previous year. Thus, the general purpose of the grading process is to maintain the standards of the examination outcome year after year (Namibia. MoE, 2012). These comparisons are done using a combination of

examiners' professional judgement and the statistical evidence to ensure that the cut of marks decided upon is reasonably valid (Namibia. MoE, 2012). The examiners' professional judgements are guided by their experience in marking the previous examination in comparison with the current examination and the key grade descriptors as reflected in the subjects' syllabuses. The statistical evidence includes a variety of statistical information from the previous and the current examination (Namibia. MoE, 2012). At the end of the marking session the learners are awarded grades from A to G according to the criteria described above.

2.3 Fundamental ideas in assessment

Educational assessment is defined by Wragg (2001) and Brookhart & Nitko (2007) as processes of obtaining information about learners, curricula, educational programs and policies, that are used in making a variety of decisions by different education stakeholders (teachers, learners, parents, policy makers, etc.). Linn (1993) emphasised the importance of educational assessment by saying that often educational assessment is expected to serve the role of a neutral barometer of educational quality. The broad aim of assessment, according to Webb (1992), should be to provide a comprehensive account of a learners' functioning within a given program. It draws on a variety of evidence which could be qualitative and/or quantitative, depending on the purpose of the assessment activity (Webb, 1992). According to Brookhart & Nitko (2007) and Wragg (2001), assessment in education is carried out for several reasons, each of which defines the type of assessment to be used. The major types of assessment are formative, summative and diagnostic assessment. The focus of this research will be summative assessment and its impact on formative assessment. The discussion on diagnostic assessment was deliberately left out as its purposes were seen not to have direct benefit to the study.

2.3.1 Formative assessment

Formative assessment is described as assessment for teaching and learning (Earl, 2003). This is assessment that is designed to be part of or embedded into the classroom teaching and learning activities (Brookhart & Nitko, 2007). It includes all classroom assessment strategies that are aimed at enhancing learning, as opposed to merely monitoring learning (Brookhart & Nitko, 2007). Additionally, it refers to regular, interactive assessments of learners' progress and understanding. It may be done formally in order to identify learning needs during the teaching and learning process and then adjust the teaching accordingly. Teachers using formative assessment approaches and techniques are seen to be better prepared to meet diverse needs of the learners (Stiggins, 2001). Formative assessments are low-stakes by nature in that

the results do not have a direct impact on individual learner's future (Leyendecker, 2008). However, the great benefit of formative assessment is that it enables differentiation and adaptation of teaching to raise levels of learner achievement and to achieve a greater equity of learning outcomes (Stiggins, 2007). The research presented in this thesis is motivated by the need to develop summative assessment activities in such a way as to enable them to be useful for formative purposes as well.

2.3.2 Summative assessment

Summative assessment is generally described as assessment of teaching and learning (Earl, 2003). Although some classroom based assessment can be classified as summative, summative assessment is generally done at the end of learning cycle and it is high stakes by nature (Earl, 2003). National examinations are an example of summative assessment whose stakes are very high. These are often legally mandated by government institutions to evaluate education programmes and to certify and select examination takers after the programme (Brookhart & Nitko, 2007). They are used for making judgments or decisions for individual learners about learners' achievement and instructional achievement (Earl, 2003; Brookhart & Nitko, 2007). This is the type of assessment on which this study will focus.

There is a perceived tension between classroom based formative assessments, and high stakes summative examinations which are used to hold schools accountable for learners' achievements (Stiggins, 2007; Brookhart & Nitko, 2008). Inevitably, most teachers strive to use the summative assessment tasks for formative purposes. As envisioned by this study, appropriately set summative assessment tasks could enhance the teaching and learning process in the classroom. Therefore, to limit the apparent conflict, the tasks set for summative tests should also be designed to appropriately inform classroom teaching Harlen (2005).

Assessment is viewed by (Airasian, 1988; Baker & Stites, 1990; Resnick D & Resnick L, 1992 and Ross et al. 2002) as much more than a set of tests. They see assessment as a driver of many educational review processes. The results from national assessments and international assessments like the Programme for International Student Assessment (PISA), the Progress in International Reading Study (PIRLS) and the Southern Africa Consortium for Monitoring Educational Quality (SACMEQ) continue to draw policy makers' attention to the significance of assessment in education (Webb, 1992). As a result of this, the need to understand educational assessment strategies becomes more critical. The major change within educational assessment

throughout the world is the move from norm-referenced assessment to criterion-referenced assessment (Sadler, 2005). Details about criterion-referenced assessment strategies are discussed in the next section.

2.3.2. 1 Criterion-referenced assessment

The main goal of criterion-referenced assessment is to assess individual learner's performance based on pre-defined desirable qualities or dimensions of learners' performance across a range of levels (Shannon, 1999; Green, 2002; Sadler, 2005; Brookhart & Nitko, 2007). These qualities and dimensions are contained within pre-determined criteria and performance standards (Shannon, 1999; Green, 2002; Sadler, 2005). This strengthens the need to have curriculum process standards developed before developing the assessment tools and prior to using them in assessment and envisaged teaching and learning activities. The standards should be clearly defined to include all desired competencies and learners should be made aware of them, through the teaching process, before assessment takes place (Brookhart & Nitko, 2007).

Criterion-referenced assessment comprises tools which are generally developed from curriculum standards that describe the subject content in line with cognitive processes (Brookhart & Nitko, 2007). The desire to evaluate the learners' mathematical proficiency requires the use of an enhanced lens that incorporates the competences of mathematical proficiency. In the Namibian system, the description of cognitive processes described in curriculum standards, is guided by Bloom's taxonomy. However, as mentioned before, Bloom's taxonomy has been criticised for its weakness in assuming that cognitive processes are ordered on a single dimension of simple-to-complex behaviour (Krathwohl, 2002; Yunus, et al., 2012). That is, it is too general to allow effective access to the complex processes that comprise Mathematical proficiency (Krathwohl, 2002). Including mathematical proficiency traits in assessment criteria, would contribute towards deepening Mathematics assessment to access mathematical proficiency.

2.3.2. 2 Proficiency complexity rubric

A scoring rubric is defined by Mertler (2001); Nitko (2001) and Brookhart & Nitko (2007) as a rating scale used to evaluate the learners' performance in an assessment. This scoring tool clearly represents the performance expectations for an assessment task (Huba & Freed, 2000). The rubric reveals how the learners' responses to assessment tasks compare to the set process standards. In general, rubrics are viewed as improving the assessment and evaluation process.

Brookhart & Nitko (2007) underline the importance of a rubric by saying, "rubrics not only improve the scoring consistency, they also improve validity by clarifying the standards of achievement you will use to evaluate your students" (pp. 191).

According to Mertler (2001) and Brookhart & Nitko (2007) there are two types of scoring rubrics: holistic and analytic. Scoring rubrics according to Brookhart & Nitko (2007) can further be classified as either generic (general) or task specific. This section discusses analytic rubrics and thereafter discusses the literature that guided the process followed when developing an analytic rubric in this research. Analytic rubrics allow the assessment of different components of the task, providing separate scores for individual parts of the learners' response first and then allow the calculation of the total scores from the individual scores (Nitko, 2001 & Brookhart & Nitko, 2007). The mathematical proficiency intended to be assessed by the assessment framework developed in this research, consists of different inter-dependent proficiency categories. These all need to be assessed in order to provide a good understanding of learners' proficiency. Therefore, an analytical rubric was considered fit for a successful assessment of multi-dimensional mathematical proficiency (Brookhart & Nitko, 2007).

The literature reveals two separate descriptions of how a scoring rubric can be developed. Brookhart & Nitko (2007) suggest two different approaches in developing scoring rubrics that can be used for assessment. Their first is a five-step approach, what they call a *top-down approach* which begins by developing a conceptual framework which can be used to evaluate learners' performance. Their five-step top-down approach is summarised in Table 2.1 below.

Table 2.1 Brookhart & Nitko (2007) top-down approach (p.195).

Step	Description
1	Adapt or create a conceptual framework of achievement dimension that describes the content and performance that you should assess.
2	Develop a detailed outline that arranges the content and performance from step 1 in a way that identifies what you should include in the generic rubric
3	Write a general scoring rubric that conforms to this detailed outline and focus on the important aspects of content and process to be assessed across the tasks.
4	Write a specific scoring rubric for the specific performance task you are going to use.
5	Use the specific scoring rubric to assess the performance of several learners, use this experience to revise the rubric as necessary

The second approach of Brookhart & Nitko (2007) which also consists of five steps is called a *bottom-up approach*. The bottom-up approach begins with a sample of question responses

using the actual responses to create your own framework following these five well described steps (Brookhart & Nitko, 2007). Their five step bottom-up approach is summarised in Table 2.2 below.

Table 2.2 Brookhart & Nitko (2007) bottom-up approach (pp.195-196).

Step	Description
1	Obtain about 10 to 12 samples of the learners' actual responses to a performance item
2	Study these responses and sort them into three groups such as; high-quality responses, medium-quality responses and low-quality responses.
3	Study the grouped responses and write out specific reasons why and how the groups of responses are different. Specific comments should be written for each learner's response.
4	Identify the emerging dimensions from the comments made in step 3 in order to create your own framework. The dimensions identified could consist of different quality levels.
5	Separate for each quality levels of the achievement dimensions identified in step 4, write a specific learner-centered description of what the responses at that level are typically like.

The major difference in the two Brookhart & Nitkos' (2007) approaches is that the top-down approach first considers the content to be assessed in developing the framework, whereas the bottom-up approach uses the learners' responses to the specific task to develop a framework.

A second description is provided by Mertler (2001), in which he presents a seven-step procedure of designing scoring rubrics. Mertler's procedure appears to have combined Brookhart & Nitkos' (2007) two approaches, in that framework development starts from the content to be assessed but in the process also studies proto-typical examples of responses to identify observable attributes. Table 2.3 shows Mertler's (2001) seven-step development process for an analytic scoring rubric.

Table 2.3 Steps of the process followed to develop an analytic rubric (Mertler, 2001, p.8)

Step 1:	*Re-examine the learning objectives to be addressed by the task*. This allows you to match your scoring guide with your objectives and actual instruction.
Step 2:	*Identify specific observable attributes that you want to see (as well as those you don't want to see) your learners to demonstrate in their response*. Specify the characteristics, skills, or behaviors that you will be looking for, as well as common mistakes you do not want to see.
Step 3:	*Brainstorm characteristics that describe each attribute*. Identify ways to describe above average, average, and below average performance for each observable attribute identified in Step 2.
Step 4	*For analytic rubrics, write thorough narrative descriptions for excellent work and poor work for <u>each individual attribute</u>*. Describe the highest and lowest levels of performance using the descriptors for each attribute separately.
Step 5	*For analytic rubrics, complete the rubric by describing other levels on the continuum that ranges from excellent to poor work <u>for each attribute</u>*. Write descriptions for all intermediate levels of performance for each attribute separately.
Step 6:	*Collect samples of student work that exemplify each level*. These will help you score in the future by serving as benchmarks.
Step 7:	*Revise the rubric, as necessary*. Be prepared to reflect on the effectiveness of the rubric and revise it prior to its next implementation.

The development of the proficiency complexity rubric for this study followed the Mertler's (2001) rubric developmental process which also incorporated the combination Brookhart & Nitko's (2007) approach. The adapted process standards were used to define that content as a way of re-examining the learning objectives to be addressed by the task Mertler's (2001) and an initial stage of framework development. The main objective for Mertler's (2001) step 5 is to describe the levels of complexity of the expected response and to determine a scale of scores to be awarded. In general, the levels described at this step were the levels that represent the cognitive complexity of the required responses as set out by the process standards.

2.3.3 Validity in educational assessment

Validity in educational assessment refers to the soundness of the interpretations made on the available evidence, by everyone involved in assessment for any particular reason (Moss, Girard & Haniford, 2006; Kane, 2006). Validity in educational assessment is described by Messick (1995) as "an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment" (p. 741).Validity was discussed here in order to understand its importance to the development of examination question papers. When examination questions developed using unclear standards they would also be unclear to the examination takers. According to Moss et al., 2006 and Namibia. MoE (2010c) unclear questions lead to unintended interpretations of the expected tested competence particularly in standardised tests such as those in national high stakes examinations. Questions set to have a common interpretation by the learners or the evaluators/markers are said to be valid questions (Moss et al., 2006; Namibia. MoE, 2010c).

The types of validity that concern this study directly were content and construct validity. Moss et al. (2006) defines assessment content validity as how well the examination samples the subject content matter about which inferences are to be drawn. It refers to the representativeness of the content components included in the questions. Construct validity on the other hand is viewed by Messick (1995) as based on an integration of any evidence that bears on the interpretation or meaning of the test scores. It includes the clarity of subject content and criterion-related evidence presented in the question. According to Messick (1995) construct validity subsumes content validity. While Moss et al. (2006) define assessment construct validity as that involving a combination of logical and empirical occurrence to

investigate what qualities an examination measures. In relation to this research study, the distinction of different aspects of proficiency would ensure the clarity in the representation of these aspects in the question papers.

2.4 High stakes assessment

High stakes assessment in education refers to any assessment activity that is used to make important decisions about the people taking them: teachers, schools, or an education administration structure (Kellaghan & Greaney, 2003; Schoenfeld, 2007a, Leyendecker, 2008,). These high stakes assessment tasks outcomes are commonly used for the purpose of selection and certification of the learners and for the accountability of educational institutions (Wragg, 2001). In many African countries high stakes assessment is generally called public or national or external examinations (Kellaghan & Greaney, 2003). They are public examinations because they are written by all learners in a particular country at a particular grade level. They are developed and administered by examination bodies that are outside the school setup, hence the name external examination (Kellaghan & Greaney, 2003). Obtaining high scores in these examinations has important benefits to the examination takers. Acceptance to institutions of higher learning, or securing scholarships or jobs are some of the important benefits. These examinations are often summative by nature and are administered in a form of paper-and-pencil tests (Leyendecker, 2008). According to Leyendecker (2008) summative examinations "hold a great significance as a rite of passage for the learners" (pp. 73). The next section discusses the influences of high stake assessment on teaching and learning.

2.4.1 The influences of high stakes assessment on teaching and learning

Towards the end of the 20 century and the beginning of the 21st century, more emphasis has been placed on enhancing strategies to support criterion-referenced assessment through formative and diagnostic testing (Shannon, 1999, Green, 2002, Sadler, 2005). The reason could be that the two are seen as having a direct impact on the teaching and learning process as opposed to summative or high stakes assessment (Brookhart & Nitko, 2007). Yet, high stakes examinations have a considerable influence on the teaching and learning process in school classrooms (Lee, Harrison & Black, 2004; Black, Burkhardt, Daro, Jones, Lappan, Pead & Stephens, 2012). Some of these influences appear to be positive and some appear to be negative. In the next sub-sections, the three major influences of high stakes assessment on the classroom teaching and learning process will be discussed. These influences are;

- Providing incentive and motivation to learn
- Teaching to the test
- Challenges to effective assessment for learning

In the discussion, the literature on these influences is used to highlight their positive contribution when activities designed for high stake assessment are used to inform classroom teaching.

2.4.1.1 Providing incentive and motivation to learn

According to Leyendecker (2008), high stakes examinations provide incentives and motivation for the learners to learn. Leyendecker (2008) suggested that high stakes examination encourages learners to work harder in school and gives them a reason to study for them in order to achieve better grades. This is a positive influence of high stakes assessment, which is significant for this study. Together with Burkhardt's (2007) What You Test Is What You Get (WYTIWYG) principle, the study was designed to take advantage of this influence. The study investigated ways to enable high stakes Mathematics examination questions to require learners to show deeper engagement with more aspects of proficiency and thus motivate learners to develop greater proficiency in Mathematics.

2.4.1.2 Teaching to the test

High stakes assessment is said to influence both the content taught and how this content is taught (Wilson, 2007; Black et al., 2012). Teachers tend to give priority to classroom activities similar to those see in the examinations. Experience shows that many teachers in Namibia, particularly Mathematics teachers, work hard to ensure that their learners are well acquainted with what past examinations demanded. This is what Black et al. (2012) termed 'teaching to the test'. Schoenfeld (2007a), and Wilson (2007), argue that the act of teaching to the high stakes examinations is unavoidable given the important consequences that its outcome has for the learners. However, teaching to the high stakes examinations in order to maximize the score might encourage rote learning without long term retention of the content learned (William et al., 2004; Wilson, 2007). Concentrating in an isolated fashion on particular competencies needed in order for learners to pass the examination fails to reflect the curriculum goals adequately. According to Kellaghan & Greaney (2003), this causes curriculum wash back and threatens the validity of future examinations. High stakes examinations do not necessarily cover every single standard which was required to be taught in classrooms, as a result of this

teaching might only be focused on the frequently assessed content. This is what is called curriculum wash back (Kellaghan & Greaney, 2003).

Despite a negative undertone attached to this influence, Black et al. (2012) indicated that it could be used to influence teaching positively when the tools are set to support teaching. High stakes assessment provides learners with better information about their own knowledge and skills (Kellaghan & Greaney, 2003). It gives a chance for the learners to learn from each other, particularly the new cohort learning from the old cohort of learners who may give an indication about what competencies are demanded and what externalized knowledge is rewarded (Kellaghan & Greaney, 2003). In this regard, the use of high quality examinations might positively influence the teaching and learning process in that teaching by using the old examination questions might help education systems to educate their learners better (Black et al., 2012). Examinations that are designed to demand high level thinking and deeper understanding of a subject, would promote learning for high level thinking and deeper understanding of that subject. Therefore, in the view of the above and again as coined by Burkhardt's (2007) principle of the project WYTIWYG, high stakes assessment should be designed to have high quality questions and promote teaching for mathematical proficiency. This study investigated the possible ways of adapting the Mathematics examinations to demand greater mathematical proficiency of the learners, as required by the syllabus.

2.4.2 Challenges to effective high stakes assessment

The pressure of high stakes examinations might also create a challenge to effective implementation of formative assessment strategies (William et. al., 2004). In order to ensure successful assessment for learning, tasks should be designed purposely to support the learning process (Brookhart & Nitko, 2007). Tasks designed for high stakes assessment, whose design does not support classroom learning for mathematical proficiency, might undermine a teacher's efforts to teach for proficiency (Kellaghan & Greaney, 2003; Brookhart & Nitko, 2007). Therefore, for high stake assessment tasks to support Mathematics formative assessment it is important that these tasks should be designed to support proficiency at all complexity levels (Burkhardt, 2007; Brookhart & Nitko, 2007; Schoenfeld, 2007a). They should include the assessment of all the aspects of mathematical proficiency which we desire our learners to acquire through teaching.

High stakes examination tasks should test all important processes and skills defined by the curriculum standards and not just what is easy to assess (Burkhardt, 2007). These should include all aspects of what defines a mathematically proficient learner. However, some aspects of mathematical proficiency seem to be more subtle to assess, especially in current high stakes assessments (Schoenfeld, 2007b), which is done mainly through paper and pencil. Conceptual understanding for example, needs to be clearly defined to ensure its assessability (Burkhardt, 2007) and its visibility in the learners' responses. The main goal of this study was to gain an understanding of how the assessment of mathematical proficiency may be improved in high stakes assessment. One of the specific objectives was to investigate the feasibility of adapting the curriculum process standards and proficiency rubrics in order to more effectively assess the identified categories of mathematical proficiency. In this light, examination questions should be written in such a way to allow the distinct elements of mathematical proficiency to be made visible. Thus, the expected responses to the questions, given in the marking schemes should show these distinct elements in order to enable the elements to be observe in the learners' responses to these questions.

2.4.2.1. High stakes assessment adaptation challenges

Numerous attempts have been made in many countries to review and improve the quality of examinations over the years (Kellaghan & Greaney, 2003). One of the common reviews that have been carried out to high stakes assessment is the shifting from a norm-referenced to a criterion-based approach. The criterion-referenced approach is designed to measure learners' success or failure in relation to criteria which represent competence rather than with reference to the performance of other learners. The improvements were often technical, designed to improve efficiency (Kellaghan & Greaney, 2003; Brookhart & Nitko, 2007). As the numbers of learners and the problem of processing very substantial numbers of examination scripts grew, many examination bodies adopted a multiple-choice questioning format (Bude & Lewin, 1997; Lewin & Dunne, 2000). Multiple-choice questioning format was later abandoned in some examination bodies and short to semi-structured answer format of questions were favoured (Bude & Lewin, 1997). Changes to the content areas examined have also positively resulted in a shift in the content to which learners are exposed in class. However, there is only limited empirical evidence that changes in examinations would result in an improvement in the overall level of learner achievements or in their cognitive processing skills (Bude & Lewin, 1997; Lewin & Dunne, 2000).

As mentioned earlier, the fact that high stakes are attached to examination performance can have a number of negative, unintended, consequences for school practice, whatever the quality of the examinations (Lewin & Dunne, 2000). These include narrowing of the implemented curriculum, neglecting what is not examined, emphasising learning strategies that are superficial or short-term (such as memorising, rehearsing, and rote learning), devoting a significant amount of time to test-preparation activities, and a heavy reliance on extrinsic rather than intrinsic motivation in student learning (Kellaghan & Greaney, 2003). Of particular significance in the context of Education For All (MoE, 2006; Iipinge & Likando, 2012) is the fact that teachers might focus their efforts on learners who are most likely to succeed, because their reputations depend on how well their learners perform in examinations (Ottevanger et al., 2007). This, in turn, might be associated with high rates of grade repetition and early dropout. In relation to this study, if assessment concentrates on including all important aspects of proficiency it would stand a better chance promote teaching all learners for proficiency and moderate the adverse effects of high stakes assessment.

2.4.2.2. Stability and constraints of systematic change of high stakes assessment

Avenstrup (1993) defined the general systematic change to an assessment system as a process of moving from one operational system to another new or adapted system (Avenstrup, 1993). He gave the changes made to the Namibian examination system after Namibia independence in 1990 as an example of a systematic change. Namibia changed from the South African Senior Secondary Education Certificate examination to the (H)IGCSE. In this regard, the changes include a complete change in the curriculum standards which would result in a change in assessment scheme (Avenstrup, 1993).

Generally, before major changes to curriculum standards are implemented through assessment, the curriculum standards are implemented in the classroom first through teaching (van der Merwe, 1993). The system changes create uncertainties of how comparable the new assessment outcomes would be to what a nation is already used to (Grewel, 1993 and van der Merwe, 1993). In other words, people become anxious of the quality of the graduates from the new examination system as compared to the graduates from the old systems. Stability is usually assured through the quality assurance measures followed during the changing process of any examination systems (Iipinge & Likando, 2012). For example, during the change of Namibian education and examination system from the South African Senior Secondary Education to CIE (H) IGCSE in 1994, curriculum standards were changed by a selected group of teachers and

went through different review processes to ensure quality outcome (Avenstrup, 1993 and van der Merwe, 1993). The artefacts adaptation in the study reported here could result in a similar systematic change to ensure their successful implementation. This is because one of the specific objectives of this study was to investigate the feasibility of adapting the curriculum process standards and proficiency rubrics in order to more effectively assess the identified categories of mathematical.

2.5 Conclusion

The context within which this study was conducted was discussed in this chapter. The chapter started off by unpacking the Namibian Mathematics examination system and continued to discuss the general assessment principles which helped to characterise the context of the study. The environment within which the study was conducted and the artefacts used during the study were described in this chapter. The next chapter, chapter 3, presents a review of the literature used to inform the study.

CHAPTER 3: The Review of Literature

3.1 Introduction

This chapter provides a review of literature relevant to enhancing the assessment of mathematical proficiency through high stakes examinations. The ongoing discussion on what is involved in a learner's ability to do Mathematics will inform this assessment. Given the fact that the study was designed to investigate the possible high stakes assessment of distinct elements of mathematical proficiency, the chapter starts by unpacking elements of knowledge in general and mathematical knowledge in particular. Different conceptions of mathematical proficiency will then be discussed.

Finally, it is generally understood that learners respond to high stakes Mathematics assessment in the form of paper and pencil only. They communicate the concepts and skills in their thoughts through what they write down. The chapter therefore highlights the literature that frames the study within Vygotsky's socio-cultural theory in terms of the internalisation and externalisation of thought.

3.2 Theorizing assessment

Learning and assessment are mutually inclusive processes of any educational program (Agaton, 2016). Thus assessment practices are directly dependent on learning theories. These provide coherence and understanding of how assessment practices are employed.

3.2.1 Making exam takers' proficiency visible

High stakes or summative assessments are grounded on the assumption that, in learning, the learner develops some capacity that is possible to assess (Bazerman, 2012). In order to be assessed, learners are required to make visible, or externalise this capacity through writing down their answers to the examination questions. Any form of writing involves some form of externalisation of prior internalised knowledge or thoughts (Le Cornu, 2009; Bazerman, 2012). Bazerman (2012) refers to new formulation, or writing words, phrases or sentences not seen before as deeper externalisation. In the context of this study, the visibility of a proficiency element is defined as the extent to which the examination artefacts or tools enable that proficiency element to be shown by the learners and noticed by the markers in the learners' responses. Visibility then depends on the process of externalising the internalised thought that could have been transformed into individualised tacit knowledge (Le Cornu, 2009). The

thought could be externalised in the same form as it was internalised or the internalised thought could grow and develop as it was externalised (Lawrence & Valsiner, 2003).

3.2.1.1. Externalisation of thought

The theory of internalisation and externalisation which guided the discussion in this section is developed from the work of Soviet psychologist, Lev Vygotsky. Vygotsky's approach to thought and its relation to the way in which it is externalised in an expression, are summarized by Bazerman, (2012) using the following points:

- An individual's thought is not apparent to others until it is externalised and concretised materially in gestures, speech and/or images. In this process the thought is also developed and differentiated.
- An individual's thought is produced through different semiotic resources.
- An individual's thought is a dialectical unit that is externalised in communicative action, and any single expression is incomplete, one-sided, and in constitutive relation to the whole.

Any form of externalisation of internalised knowledge or thought is prompted by the situation (Pape & Tchoshanov, 200; Bazerman, 2012). Therefore, the representation of internal mental images takes place within a social interaction (Pape & Tchoshanov, 2001). This theory has its base in Vygotsky's idea that knowledge is developed under the influence of a situation through relationships and interaction with others (Hardman, 2008). Vygotsky views knowledge as generated through culturally motivated activities and signs in which a situation can be modelled as a mediated act. Tools mediate the achievement of the object by the subject (Hardman, 2008). The externalisation process is complex, particularly for learners during an examination, because it is influenced by learners' learning prior to the examination and by the situation the learners are in during the examination (Lawrence & Valsiner, 2003). The learners' beliefs about examinations, their experience of the subject and the properties of the tools or artefacts that facilitate this process are some of the factors that influence this externalisation process.

3.2.1.2. Vygotsky's culturally motivated activity model

In general, an examination is viewed as a culturally motivated activity, which in this case consists of two subjects (the learners and the teachers). The tools of this activity are the curriculum standards, the examination papers, marking schemes and the learners' answer scripts. The tools are used in the examination system as an activity in two distinct but related

ways; first the examination papers are used by the learners (the first subject), to mediate the object of the activity, which is to demonstrate their mathematical proficiency through their responses to examination questions. Secondly, the marking schemes and the learners' answer scripts are used by teachers or markers (the second subject), to mediate the object of the activity, which is to identify the mathematical competencies in the learners' responses. The study focused on the latter in that, the tools adapted for this study were the curriculum standards and the marking schemes. The marking schemes were utilised to mediate the process of making visible the mathematical proficiency externalised by the learners in the form of language as suggested by (Le Cornu, 2009; Bazerman, 2012). The forms of language here could be through speaking, writing in a form of letters or mathematical symbols and any other illustrations as a result of their internalised Mathematics knowledge. The identification of learners' mathematical competencies were to be made visible by the adapted tools and noticed by the teachers marking the scripts. Figure 3.1 below shows Vygotsky's model of mediated action as specialised for this study showing the triadic relationship between tools or artefacts, subject and object. One of the specific objectives of the study was to investigate possible adaptation of these artefacts to improve the visibility of identified elements of mathematical proficiency when the learner interacts with the artefacts in the examination situation.

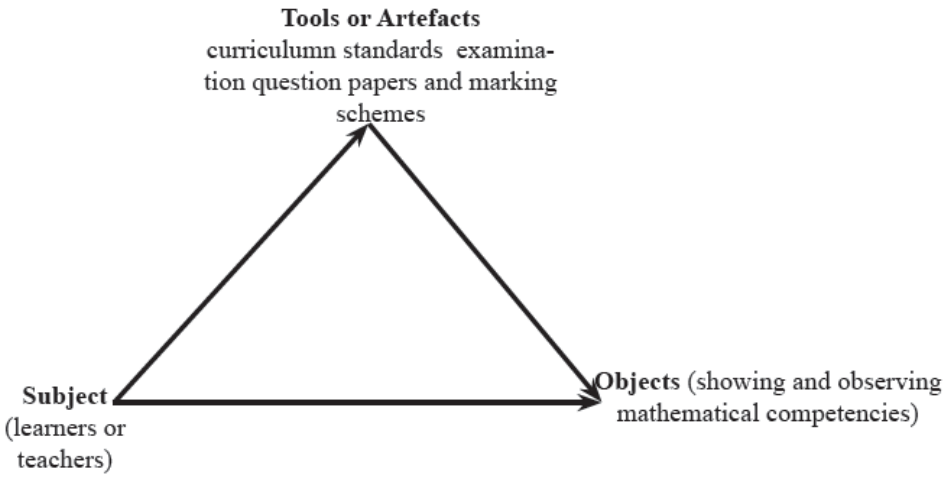


Figure 3.1 Vygotsky's model of mediated act as adapted for this study (adapted from Bazerman (2012).

The artefacts should enable the identified aspects of mathematical performance to be distinctively noticed by the markers in the learners' responses. The learners' externalised responses are presented through different formulations of words, phrases and mathematical

expressions. However, the meaning (to the learner and to others) of the learners' externalised words, phrases and mathematical expressions depends on their previous interactions (Bazerman, 2012). This formulation according to Le Cornu (2009) and Bazerman (2012) undergoes public discipline, requiring that the externalised concept or ideas be intelligible to others. The writer, who in this case is the learner, has to work to evoke the presence, meaning and value of the expressed ideas (Ericsson & Simon, 1998; Bazerman, 2012).

An individual's expression either in speech or written form reflects aspects of their thoughts and they typically verbalise or make visible only a fraction of these thoughts (Ericsson & Simon, 1998; Daniels, 2001). That is, this expression does not fully represent the individual's mental activities, and some of these are actively suppressed. This study looks at the development of a system of standards that would encourage learners to move from the current practice of merely supplying calculated answers to one which encourages them to show more of their mathematical thinking. The current system is seen not to be encouraging an individual's externalisation of thoughts.

The main goal of this study was to gain an understanding of how the assessment of mathematical proficiency may be improved in high stakes assessment. To do this, possible changes to a practice which could allow greater visibility of the learners' internalised thoughts that they apply when producing the calculated responses are studied. In particular, possible ways of adapting the examination tools or artefacts within a criterion-referenced assessment as discussed in the paragraphs above were investigated. The adaptation of the standards that define the criteria, aims to ensure that the rubrics developed for the examination question papers would yield an improved visibility of the learners' externalised thoughts, to be identified by the markers. The general curriculum standards developed and adapted for this study were fundamental artefacts which provided much of the basis for adapting the other tools and also formalized the adapted conceptual orientation to mathematics. Following the curriculum standards, the general proficiency-complexity rubric was developed. This followed the development of the marking schemes, and specific question paper proficiency-complexity rubrics, which were the central part of the research and were used in the collection of the empirical data. The marking schemes or specific question paper proficiency-complexity rubrics were developed to enable the assessment of separate elements of mathematical proficiency.

3.2.1.3. Externalisation process and criterion based assessment

The degree of externalisation of the learners' thoughts also depends on the 'quality' of the artefacts mediating the objects. Criterion referenced assessment, as discussed in section 2.3.3 was employed through the development of the primary artefact, the process standards. These were used to define the criteria that guide both teaching and assessment. In order for criterion-referenced assessment to work, it was necessary to be clear about what the learner should be learning in terms of process standards or performance criteria that define the expected competencies. The process standards would guide the design of assessment tasks which would allow learners to show how well they have met the criteria through their externalised response (Biggs, 2003).

The basic principle of Vygotsky's theory is that artefacts occupy a mediating role in human reaction and interaction with the world (Biggs, 2003; Bazerman, 2012). For an examination system the artefacts would be; the process standards, the question papers and the marking schemes or rubrics for the question papers. These are social items with certain modes of operation developed socially by the teachers and perform the mediating role because they correspond to the objectives of a practical action (Verenikina, 2010). In this case, the practical actions include the process of setting question papers by the teachers, learners' engagements in responding to the questions in order to demonstrate their proficiency and, of particular interest to this study, the marking of the learners' responses to the question papers. Moreover, in the context of criterion-based assessment, the practical actions are spelled out by the criteria given by the process standards. These are then observed through the visible features of the learners' response to the examination questions by the teachers marking or judging these responses.

3.2.2 Cognitive complexity

Teaching and assessment standards do not only consider the learners' knowledge and skills in the subject topic, but also require different types of cognitive processes while processing that knowledge (Mertler, 2001). Such cognitive processes differ in complexity. Cognitive complexity refers to the cognitive demand of tasks associated with the standards (Webb, 1999). It is also referred to by Niebling (2013) as a lens for understanding the expectations of the teaching and learning standards. Over the past years, attempts have been made to develop models for understanding cognitive complexity as it relates to designing instruction and assessment tasks (Niebling, 2013). The cognitive complexity of an examination question is

determined by the assessed content area, cognitive demand and linguistic demand required by the learner in order to understand, process, and respond successfully to that question (Ferrara, Svetina, Skucha, & Murphy, 2011).

Ferrara, et al. (2011) further highlighted that there are three main sources of question complexity; the command of textual evidence, response mode and the processing demand. The above presents the general understanding of cognitive taxonomy as currently applied to educational assessment. The cognitive complexity is usually represented through different categorisation such as Bloom's taxonomy as used in the NSSCO and others such that of Norman Webb's Depth of Knowledge (DOK) (Webb, 1999). In order to define the proficiency categories arrangement and the direction of the scale of the categories' complexity which is viewed appropriate for this study, the main sources of question complexity suited for this study are first discussed. Therefore, the next subsections discuss these three main causes of complexity in relation to this study.

3.2.2.1 Textual complexity

Textual complexity is regarded as the command of textual evidence (Ferrara, et al., 2011). The command of textual evidence, in relation to an examination question, refers to the amount of the textual data or information that a learner is required to process before attempting to respond to a question (Ferrara, et al., 2011). Examination takers should develop special skills and strategies for reading text in the Mathematics content area (Shanahan, 2013). According to Hiebert (2011) textual complexity is defined through three factors. The first factor is the **quantitative indices** which refers to the information on the number of specific words or textual representation and length of sentences. For Mathematics specific words could be formulae used to solve particular problems or any other mathematical terminologies and representations (Hiebert, 2011). The second factor is the **qualitative analysis**, which refers to the levels of meaning, structure, and language conformism and clarity, and knowledge demands (Hiebert, 2011). For qualitative analysis, the readers need to make inferences and understand the reason for using the characters in the text. The third and last factor is the **reader and task variables**, which refers to the readers' background knowledge of a topic to which the text used relates. It also includes the ways in which situations influence readers' interactions with a text (Hiebert, 2011; Shanahan, 2013). In relation to this study, this relates to both the learners, as the examination taker and the markers who reads and judges the learners' responses to the questions.

3.2.2.2 Response mode

The response mode refers to how a leaner would be required to respond to an item. In general, selecting a response is less demanding than constructing a response (Ferrara et al., 2011). Furthermore, written responses, which are the ones required in the Mathematics examinations researched here, is said to yield fewer idea units than the spoken response (Kellogg, 2007). The reason could be that writing required the activation of additional grapheme codes and took relatively longer to come out than speaking (Putnam & Roediger III, 2012). Additionally Putnam & Roediger III (2012) highlighted that responding through writing used more working memory capacity than speaking, meaning that the learner had fewer resources available for recall when giving their answers in a written form.

3.2.2.3 Knowledge processing model

Knowledge processing model discussion in this section is done in order to understand the study of how people encode, structure, store, retrieve, use or otherwise acquire knowledge (Lutz & Huitt, 2003). The way in which knowledge is believed to have been learned and stored (internalised) affects the way in which knowledge is believed to have been retrieved and used (externalised) (Bazerman, 2012). The hierarchical Bloom's taxonomy discussed in section 2.2.3 suggests a dependency of the higher order knowledge processing categories on the lower ones. According to Lutz & Huitt, (2003) the network mode on the other hand suggest the knowledge categories could be altered by the individual's experience. However, they argue that its hierarchical system suggests a very linear development (Lutz & Huitt, 2003). In the debate against the network model Smith, Shoben, & Rips (1974) claimed that information should be stored and externalised as sets of defining characteristics, instead of being organized in a hierarchal system, as given in the network model.

This study was designed to align with Smith et al.'s (1974) argument that knowledge is not necessarily processed and externalised in a defined hierarchical manner. Thus, an individual's knowledge processing model should support his or her natural preferred way of externalising the learned and stored information Smith, et al. (1974). In relation to this study, the knowledge elements in a form of the proficiency categories could be externalised parallel to each other, simultaneously and even from any proficiency category an individual is disposed to progress (Lutz & Huitt, 2003). Thus, despite the fact that the complexity of processing the knowledge from one category to another might increase, the order of processing such information need not to follow a systematic linear progression. In other words, a learner may also be able to make

inferences as a way to solve a problem without necessarily understanding the concept. They might first encode what they learned last (the procedure) instead of displaying the knowledge of the content to which the procedures are applied. This is what is related to as a top-down approach by (Lutz & Huitt, 2003). The next section discusses the cognitive complexity scale for the study.

3.3 Assessing what is important to assess

High stakes assessment in the world is dominantly conducted using examination question papers that consist of highly objective questions. These are questions whose responses can be objectively marked (Kellaghan & Greaney, 2003). According to Burkhardt (2007) and Schoenfeld (2007b) examination bodies around the world have resorted to measuring only what seems easy to measure and easy to assure reliability in order to ensure the reliability of high stakes examination outcomes. Burkhardt (2007) argues that for effective mathematical assessment to be conducted, it is essential, to assess all aspects of Mathematics, particularly those that are important, not just what is easy to assess. Also the key assessment design principle should be to "measure what is important, not just what is easy to measure" (Burkhardt, 2007 p. 78).

Objectively marked examinations are seen to provide a "legally bullet-proof" (Schoenfeld, 2007a, p.8) assurance for the reliability of examination outcomes (Schoenfeld, 2007a and Burkhardt, 2007). The more objective the learners' responses, the more valid and reliable the outcome of the marking process, because their reliability is not affected by the markers judgement (Namibia. MoE, 2012). Studies by Hill (1973); James (1974); Murphy, (1978) as reviewed by Meadows and Billington (2005) have shown that more closely defined questions, which demand definite answers, are regarded to have higher examination reliability as opposed to more subjective responses. Subjective responses would result from testing aspects of mathematical knowledge using more openly defined questions (Meadows and Billington, 2005). In particular, aspects that require learners to give a structured and subjective response include the demand of description, definition and explanation. These formats of answering are regarded as more subjective than multiple choice and objective routine calculations (Namibia. MoE, 2010c). However, a further argument is made by Meadows and Billington (2005), that objective tests achieve high reliability at the expense of validity. To ensure valid assessment of all important aspects of Mathematics, through high stakes assessment, it is therefore important for this study to understand the distinction between the aspects to be tested.

3.4 Distinguishing elements of knowledge

The contrast between viewing knowledge as a unitary whole and viewing knowledge as involving separately identifiable elements, has been debated in general and in relation to different subjects for decades (Skemp, 1979; Star & Stylianider, 2013). Star & Stylianider (2013) highlight that the origins of the debate on the distinction of knowledge are hard to identify precisely. However, the distinction between the types of knowledge became widely known in mathematics education after the publication of the book edited by James Hiebert in 1986, entitled 'Conceptual and Procedural Knowledge in Mathematics: An Introductory Analysis' (Star & Stylianider, 2013). Historically, in mathematics education, this debate seemed to have started from the work of Richard Skemp in 1979, who discussed knowledge elements such as 'know what and know how' and spelled out the distinction between what he termed as relational and instrumental knowledge in relation to Mathematics teaching (Skemp, 1979).

The concept of distinguishable proficiency categories that is fundamental for the discussion in this study, is anchored in and based on this debate. Theories of learning and cognition, such as those of Piaget and Vygotsky viewed human behaviour as shaped by at least two different kinds of knowledge (Schneider & Stern, 2010b). Piaget saw knowledge as having three types: physical, logical-mathematical and social knowledge, while Vygotsky viewed knowledge as structural or instrumental (Schneider & Stern, 2010b). According to Haapasalo (2003) conceptual knowledge "provides an abstract understanding of the principles and relations between pieces of knowledge in a certain domain, and another one enabling us to quickly and efficiently solve problems" (p.1). Rittle-Johnson, Siegler, and Alibali (2001) suggested that, in many subjects learners should learn both fundamental concepts and procedures for solving problems. However, they warn that the developmental relationship between conceptual and procedural knowledge is not well understood.

3.4.1 The distinction between conceptual and procedural knowledge

According to Schneider & Stern, (2010) the interaction between conceptual and procedural knowledge influences the development of Mathematical competencies. They went on to say that "many developmental psychologists have found it useful to treat knowledge not as a unitary contrast but as differentiated into at least two kinds of knowledge" (p. 178). Star &

Stylianider (2013) consider the distinction between conceptual and procedural knowledge to be useful in Mathematics learning and state that the clearer we can be about their distinction the better. Conceptual knowledge refers to the knowledge of concepts, facts and principles, including the knowledge of problem representation (Haapasalo & Kadijevich, 2000). It is viewed as general, abstract knowledge of the fundamental principles and interrelations in a particular subject (Schneider & Stern, 2010a and Schneider & Stern, 2010b). On the other hand, procedural knowledge refers to a dynamic process and the action of carrying out the procedures shown through relevant representation (Haapasalo & Kadijevich, 2000). A procedure is a series of steps, or actions, done to accomplish a goal, the knowledge of these steps is what is called procedural knowledge (Haapasalo & Kadijevich, 2000).

Star & Stylianider (2013) warn that the distinction should not be used as a classification scheme into which all knowledge can or should be sorted. This is because not all knowledge can be usefully described as either conceptual or procedural. Some knowledge seems to be a little of both and some seems to be neither. In line with this, the distinction was not necessarily used in this study as a way to classify knowledge into different units but rather to characterise different interactive aspects of one unit. Therefore, the classification of knowledge should be used to view knowledge from different related perspectives, each of which gives some insight into the others, resulting in a web that forms unitary contrast (Rittle-Johnson & Alibali, 1999). This insight should be used to create an understanding of how each type of knowledge is measured (Rittle-Johnson, Siegler & Alibali, 2001). The distinction is critical for interpreting evidence on the relationships between conceptual and procedural knowledge. Conceptual knowledge has been assessed in a large variety of ways, whereas there is much less variability in how procedural knowledge is measured. Rittle-Johnson, Siegler & Alibali (2001) suggest that the distinction in the assessment of the knowledge elements should be based on the novelty of the question asked. Contrary to this, Schneider & Stern (2010b) suggest that the distinction between assessments of conceptual and procedural knowledge should be based on the distinguishable features of the learners' response to the question asked. This is because different views of knowledge may complement each other as a learner produced the required response to a question (Schneider & Stern, 2010a).

The study used the distinction as a way of taking multiple views of knowledge (one thing) that provide complementary insight into knowledge that becomes much deeper once the insight is synthesised. Unlike the study conducted by Rittle-Johnson, Siegler & Alibali (2001), this study distinguished between assessments of conceptual and procedural knowledge on the basis of distinguishable features of the learners' responses, as guided by how the questions are asked. The study used the distinction provided through the literature on mathematical proficiency to characterise different features of the learners' responses to the examination questions. Mathematical proficiency was used in this study to define distinguishable elements of mathematical knowledge with an effort to use them in defining the learners' proficiency through high stakes assessment.

3.5 Mathematical Proficiency

Theories of mathematical proficiency provide general formulations of how learners engage successfully with Mathematics. These discussions can be dated back to the early 20th century work of George Polya on mathematical problem solving to more recent work of Kilpatrick, Swafford and Findel (2001), Schoenfeld (2007a) and Burkhardt (2007). The subject was stimulated and grew thicker roots as a result of the call made at the United Nations Educational Scientific and Cultural Organisation (UNESCO) conference on Science and Technology Education in 1984 to ensure a full development of all learners' mathematical abilities (UNESCO, 1984). The discussion was motivated by a need to move from the traditional way of teaching Mathematics to a way that discourages rote learning and encourages learning for understanding (UNESCO, 1984; Lester, 1994) in order to encourage the development of knowledge based economies in the world. The ongoing discussions on the need to improve how learners learn and do Mathematics, call for learners to be able to use mathematical knowledge and skills to make appropriate connections (Schoenfeld, 2007a). The debates on the matter encourage learners to engage successfully in the mathematical process and demonstrate a mastery of Mathematics (Kilpatrick, et al. 2001 and Schoenfeld, 2007a). The study reported here used this motivation to take advantage of the apparent power high stakes assessment has on the teaching and learning process. The following sub-sections discuss the literature arranged

into four related but separate contributions that attempted to establish a clearer characterisation of what may be defined as a learner's mathematical capabilities.

3.5.1. George Pólya's contribution

The first formulation that can be used to understand learners' mathematical capabilities can be found in the work of George Pólya in his 1957 book 'How to Solve It'. His work opened up the study of problem solving strategies. In a stage he called 'looking back', Pólya encouraged problem solvers to use their strategic thinking skills and reflect on what they did when solving a mathematical problem (Pólya, 1957). He stressed that learners should use the mathematical knowledge they had internalised to solve mathematical problems. Mathematical problem solving, as discussed by George Polya is characterised by a four step approach: *understanding the problem*, *devising a plan*, *carrying out the plan* and *looking back* (Pólya, 1957). Pólya's first step of 'understanding the problem' encourages learners that are tasked with solving a mathematical problem to develop an understanding of the problem they are to solve. They should know what the question is all about and what the question demands. For example; what is the unknown and what are the other conditions they need to use to get to the solution. Therefore, he indirectly challenges Mathematical question setters to give enough information to enable learners to find a solution to the problem (Pólya, 1957). In the second step of 'devising a plan' he encourages the problem solver, to apply their strategic thinking skills to come up with a plan that could be carried out to solve the problem. He also encourages the problem solver to reflect on the process of solving the problem through his fourth and final step of looking back (Pólya, 1957). Pólya's four steps could be seen as indicating three distinct functional elements which a mathematical problem solver could use to successfully solve a mathematical problem.

Relating Pólya's four steps contribution to the distinction of knowledge into conceptual and procedural knowledge shows that Pólya provides the starting impetus to the consideration of distinguishable elements in doing mathematics. However, his strong focus on problem solving hinders the distinguishable knowledge elements from coming out explicitly. He opened up the awareness of the field to the fact that problem solving is not the same as automatic, skill based work, but rather involves a number of processes that require meaning, flexibility and deliberate decision making. He requires the problem solver to have the procedural capacity to carry out the chosen mathematical procedures appropriately. Although conceptual knowledge is necessary, this conscious, flexible decision making is not focused on mathematical concepts

and meanings, but rather focuses on the process of solving the problem. His contribution provides a parallel strand, more or less independent of the procedural/conceptual distinction needed to understand the distinct elements of mathematical proficiency. The importance of a learner having the knowledge of concepts, rules, algorithms and procedures (Haapasalo & Kadijevich, 2000) seems not to have been made apparent. Pólya's discussion of problem solving was instrumental in the development of mathematics education and it also seem to have inspired the discussions in the two contributions that are discussed in the next two sections.

3.5.2. The National Research Council committee (the contribution of Kilpatrick et al.)

The second formulation is the work of the National Research Council committee, chaired by Jeremy Kilpatrick and described in their report Kilpatrick, et al. (2001). They reviewed and studied a broad range of research studies related to the learning and teaching of Mathematics from pre-kindergarten through to the eighth grade. The objective of this review was to respond to several concerns on "inadequate teaching, poorly designed Mathematics curricula and low test scores which fuelled fears that learners might be deficient in the mathematical skills demanded by society" (Kilpatrick, et al., 2001, p.xiii). How well school children were learning mathematics, was a growing general fear around the world including the United States. The concerns might have been fuelled by the observation that the earlier classification of mathematical knowledge does not provide enough rigour for all of the knowledge components (Star, 1999). She argued that treatment of procedural knowledge from Hiebert's conceptual/procedural knowledge framework is inadequate. In order to address these concerns, Kilpatrick, et al. (2001) analysed research on the Mathematics to be learned, research in cognitive psychology and research in Mathematics education to reach their recommendations. The review provided "research-based recommendations for teaching, teacher education, and curriculum, for improving learners' learning and to identify areas where research is needed and to give advice and guidance to educators, researchers, publishers, policy makers, and parents" (p.3).

As a result of these review meetings, they adopted a more comprehensive, composite view of successful Mathematics learning. They chose the term 'mathematical proficiency' to denote this view (Kilpatrick, et al., 2001). Mathematical proficiency in this case included learners' capability, competence, knowledge and skills in Mathematics. They characterised mathematical proficiency as a multi-dimensional trait, which can be broken down into the following five intertwined and interdependent strands:

- conceptual understanding
- procedural fluency
- strategic competence
- adaptive reasoning
- productive disposition

Conceptual understanding refers to a learner's comprehension of mathematical concepts, operations, and relations. Learners with conceptual understanding have their knowledge organised into a coherent whole, "which enables them to learn new ideas by connecting those ideas to what they already know" (Kilpatrick, et al. 2001, p. 118). When concepts, facts and procedures are learned with understanding they can be used to make connections with newer ones and can easily be reconstructed when forgotten (Kilpatrick, et al. 2001). Although Kilpatrick et al., (2001) argue that conceptual understanding need not be explicit, the learners' ability to verbalise connections between concepts and make their understanding visible through different representations was important for this study. This appears to be in line with Hiebert's conceptual knowledge in that the knowledge of concepts can be assessed verbally and through a variety of tasks which would bring out different representation (Star, 1999).

Procedural fluency, the second strand, refers to a learner's skills in carrying out procedures flexibly, accurately, efficiently, and appropriately. The learner needs to have the "knowledge of procedures, knowledge of when and how to use them appropriately, and skill in performing them flexibly, accurately, and efficiently" (Kilpatrick, et al. 2001, p. 121). The learner's skills of carrying out mathematical procedures flexibly, accurately, efficiently, and appropriately in all Mathematics subject topics also enable their understanding of concepts and procedures to be made evident particularly through assessment (Kilpatrick, et al. 2001). To highlight the fact that procedural fluency and conceptual understanding are interwoven, Kilpatrick, et al. 2001 further stated that "a certain level of skill is required to learn many mathematical concepts with understanding, and using procedures can help strengthen and develop that understanding" (p. 122). In contrast with Hiebert's procedural knowledge, Kilpatrick, et al.'s procedural fluency is observed through more than just uni-dimensional and non-verbally automated performance of skills (Star, 1999). However, both conceptual understanding and procedural fluency appear to be related to Polya's stage of carrying out the plan to solve the Mathematics problem.

Strategic competence is defined as a learner's ability to formulate, represent, and solve mathematical problems (Kilpatrick, et al. 2001). Learners need to be able to figure out exactly what the problem is all about and restate it in their own words. "They should know a variety of solution strategies as well as which strategies might be useful for solving a specific problem" (Kilpatrick, et al. 2001, p. 124). Strategic competence is closely related to Polya's 'devising a plan' and carrying out the plan to solve the problem, involving a number of processes that require meaning, flexibility and deliberate decision making. However, it does not link directly with the knowledge distinction described by Hiebert (1986) where procedural knowledge seemed not to be explicitly visible and was only observed through one-dimensional and non-verbally automated performance of skills. A learner with strategic competence should be able to represent a given formulated problem as the first step in solving it. For the sake of successful assessment, a mathematical representation of some kind whether numerical, symbolic, verbal, or graphical should be made visible.

Adaptive reasoning refers to a learner's capacity for logical thought, reflection, explanation, and reasoning, it includes the knowledge of how to justify a conclusion (Kilpatrick et al., 2001). Kilpatrick et al. (2001) further stated that adaptive reasoning includes not only "informal explanation and justification but also intuitive and inductive reasoning based on pattern, analogy, and metaphor" (p. 129). As part of the reflection process, adaptive reasoning relates and includes Polya's 'looking back'. However Hiebert's distinction of knowledge does not explicitly relate to adaptive reasoning.

Productive disposition refers to a learner's habitual inclination to see Mathematics as sensible, useful, and valuable, "coupled with a belief in diligence and one's own efficacy" (Kilpatrick et al., 2001). If learners are to be proficient in Mathematics they need to have the understanding of the concepts and be fluent in the procedure of solving Mathematical problems. They also need to be competent in carrying out the strategies and be able to reason adaptively, they must also believe that Mathematics is understandable and not arbitrary (Kilpatrick et al., 2001). A learner with a productive disposition believes that with hard work and perseverance Mathematics can be learned and used; and they themselves are capable of doing it. Either Polya or Hiebert have talked directly about the elements of productive disposition in their mathematical knowledge distinction.

One important property of Kilpatrick et al's. (2001) formulation of mathematical proficiency is the fact that the five strands are interdependent. Learners would not be able to shows skills that reflect their procedural fluency if they are not equipped with conceptual understanding. The learner equipped with conceptual understanding and automated procedures would be enabled to think about other aspects of a mathematical problem and would be able to tackle new kinds of problems, which leads to new understanding (Kilpatrick, et al. 2001). These learners would be able to apply their thinking logically, reflect on what they are doing and even realise the value of Mathematics (Kilpatrick, et al. 2001). Hence, learners who are proficient in Mathematics "understand basic concepts, are fluent in performing basic operations, exercise a repertoire of strategic knowledge, reason clearly and flexibly, and maintain a positive outlook toward mathematics" (Kilpatrick et al., 2001, p. 409). According to Ball (2003), the interdependence of these strands does not only have serious implications for how learners acquire mathematical proficiency, but also for how the knowledge they have acquired should be evaluated. Despite that, these strands need to be interactively engaged during problem solving (Ball, 2003), evaluating these strands would only be valid if the competencies associated with these strands are made explicitly visible through the learners' written work. One of these implications in relation to this study was that they could be treated as a unitary entity, posing a challenge on evaluating them separately. However, Kilpatrick et al. (2001) clarify the interdependence by stating that the strands represent different aspects of a complex whole, which can be distinctly observed. In other words, mathematical proficiency should not be viewed as a one-dimensional trait, but a mathematical proficient learner should be equipped with skills that represents all strands. This is because mathematical proficiency cannot be achieved by focusing on just one or two of these strands but on all of them (Kilpatrick et al., 2001).

Additionally, Kilpatrick et al. (2001) acknowledged that the interwoven property also has an implication on how teachers develop their learners' proficiency. However, if teachers recognise the individual contribution of these aspects to a learner's mathematical proficiency, the coherent teaching process would benefit from such individual insight. In relation to teaching, the interdependence encourages teaching for mathematical proficiency so as to have a clear interaction between the teachers and learners as they interact with the mathematical knowledge and skills expected to be learnt by the learners (Kilpatrick et al., 2001). A teacher should be able to relate to what the learners can or cannot do through their interaction either spoken or written (Kilpatrick, et al. 2001; Ball, 2003). This is what Kilpatrick et al. (2001) called "a

mutual and interdependent interaction of the three elements; mathematical content, teacher, learners" (p. 9). Although, this interconnection is explained here in the teaching context it could be extended to teachers evaluating learners' interaction with the content through their written work, which is the focus of this study. In line with this study, interdependence should be interpreted as a network of different aspects of knowledge creating one observable unit. It should be understood as a discourse through which distinct but related aspects of proficiency should be understood and used by both teachers and learners in order to ensure the development of learners' proficiency. Thus, in order to recognise individual insight, each aspects of mathematical proficiency which contributes to the interwoven web of Mathematics knowledge, needs to be distinctly assessed. The above highlights a need to understand the characteristics of individual elements of proficiency to ensure their separate and valid identification through assessment. This could help to create an understanding of mathematical proficiency as a unitary construct and be able to characterise the learner's achievement of mathematical proficiency.

The mathematical proficiency discussed under Kilpatrick, et al.'s (2001) contribution was designed to describe what is meant for anyone to learn Mathematics successfully (Kilpatrick, et al., 2001). It brought out a number of valuable distinct traits and competencies that can be used to describe a mathematically proficient person and relevant for this study. However, its elements do not seem to bring out the clarity needed in this study to understand the success and failures of a mathematically proficient person during the mathematical problem solving process (Schoenfeld & Kilpatric, 2008). The distinction of the aspects of proficiency required for this study was guided by the question, how the learners' proficiency in Mathematics can be measured, through high stakes assessment (Milgram, 2007; Schoenfeld, 2007b). High stakes testing is constrained by several factors; construct and predictive validity and reliability which greatly inhibits the testing of the full complement of Kilpatrick, et al.'s (2001) five strands (Groves, 2012). To measure Mathematics through high stakes assessment, the emphasis is put on the fact that the assessment framework needs to be precise in order to ensure that the assessor demands precise and specific answers. The following sections engage more deeply with this concern.

3.5.3. The contributions of James Milgram and Alan Schoenfeld

The next contribution that was chosen to discuss the concept of mathematical proficiency was the work of James Milgram and Alan Schoenfeld in their separate contributions which are included in a book titled "Assessing Mathematical Proficiency" edited by Schoenfeld (2007).

In contrast to the Kilpatrick, et al.'s (2001) contribution, James Milgram and Alan Schoenfelds' contributions were designed to characterise a mathematically proficient person as reflected by what he or she demonstrates through the process of solving mathematical problems. Therefore, the views discussed below encourage the examination question setter to design well-posed problems where all the terms need to be precisely defined and refer to a single universe where mathematics can be done (Milgram, 2007; Groves, 2012).

Two of the contributors to this book; James Milgram and Alan Schoenfeld, presented their distinct but related interpretations of what it means to be mathematically proficient. Milgram (2007) and Schoenfeld (2007a) define mathematical proficiency as the ability to do more than just being able to reproduce standardised knowledge on demand. A mathematically proficient person should be able to extend the Mathematics they know, find new solutions and apply mathematical knowledge in different and new contexts (Schoenfeld, 2007a; Milgram, 2007 and Lester, 1994). Their contribution seemed to have been informed by the Kilpatrick, et al. (2001) formulation discussed in section 3.5.2 above, however their formulation seemed to be less divided and more concise than the Kilpatrick, et al.'s (2001) formulation. Their individual formulations are presented in the next paragraphs and were related to the earlier contribution of Kilpatrick et al. (2001). The discussion in these sections concluded by highlighting the separate aspects of proficiency selected for this study.

3.5.3.1 James Milgram's Contribution

In an attempt to articulate his understanding of mathematical proficiency, Milgram (2007) presents Mathematics through what he views as very significant features. According to Milgram (2007) these features are: precision and having a well-posed problem whose demands are clear to enable the solver to solve them. He refers to Mathematics as "having precise definitions of all terms, operations and properties of these operations" (p. 33). He thus underscores the importance of definitions by saying "the first key component of mathematical proficiency is the ability to understand, use and as necessary, create definitions" (p. 44). This relates to Kilpatrick et al.'s conceptual understanding in that, such a definition will help to enhance and most importantly make visible the precise understanding of a concept. Representing a concept through procedural calculation alone does not guarantee the understanding of the concept (Milgram, 2007). Such procedures could have been memorised and regurgitated on demand during an examination (Milgram, 2007). The true understanding

of the facts, concept or operations, according to Milgram (2007) is demonstrated mainly by stating in one's own words what that fact or concept is. In other words, the learners' use of language while defining the concepts or procedure makes their knowledge of the concepts and procedure more visible. Having well posed problems on the other hand, facilitates the understanding of a mathematical problem to be solved and enables the solver to give what is demanded. This therefore encourages the examiners to demand precise information using exact and accurate representations.

Furthermore, Milgram (2007) views Mathematics as having three basic components. The first component is that Mathematics operates *on stages*. Mathematical concepts, such as numbers, sets, linear inequalities etc., form the stages where Mathematics operates. Concepts provide a mathematically proficient person with sufficient structure to apply procedures when doing Mathematics (Milgram, 2007). The second of Milgram's components needed for proficiency is *precision*, which is defined as how much information is conveyed by a representation. It refers to how exact and accurate the representation of the information is and enables the understanding and utilisation of mathematical concepts in problem solving (Milgram, 2007). The last one of Milgram's three components is *problem solving* (Milgram, 2007). He discusses four key steps that should be taken in problem solving and these are identical to Polya's steps: *understanding the problem*, *developing a plan to solve the problem*, *carrying out the plan* and *looking bac*k. Milgram (2007) also recognises that there are verbal and non-verbal aspects to problem solving. For example, learners are required to practise with numbers until the operations become automatic to evoke the non-verbal aspects of problem solving (Milgram, 2007).

Milgram's (2007) characterization of a mathematically proficient learner, just like Polya's appears to have limited clarity in distinguishing elements of proficiency and appears to be describing simplified competencies that can be subsumed in the other categorisations. For instance, the knowledge of Mathematics concepts needed to incorporate the knowledge of other mathematical facts and procedures, which in Milgram's categorisation could be assumed in his third component, problem solving and could challenge the construct validly needed in high stakes examinations (Groves, 2012; Hoover, 2014). Another example is his emphasis on precision, which could be incorporated into the competencies that are needed with procedural fluency, strategic competence (Hoover, 2014). Like Polya, Milgram's (2007) provides the starting motivation to the consideration of distinguishable elements in doing mathematics. He

therefore also provides a strong focus on problem solving. However his emphasis on precision creates a strong alliance with Kilpatrick, et al.'s (2001) conceptual understanding, by encouraging precise definitions of concepts, facts and procedures. Moreover, precision indirectly links to Kilpatrick, et al.'s (2001) procedural fluency and strategic competency in that exact and accurate representation of any solution using any procedure of solving problem allows for valid evaluation.

3.5.3.2 Alan Schoenfeld's contribution

The final classification of mathematical proficiency that was considered for this study was the four strand categorisation of Schoenfeld (2007b). These categories are: Knowledge base, Strategy, Metacognition and Beliefs and disposition.

(a) Knowledge base

Knowledge base refers to the comprehension and understanding of mathematical concepts, and operations in order to be able to produce facts and definitions, and execute procedures on command (Schoenfeld, 2007b). It relates to the understanding of the concepts or operations, the representation of these concepts or operations and relationships between these concepts or operations Schoenfeld (2007b). It includes Milgram's (2007) idea of having precise definitions and descriptions of all terms, operations and properties of the operations. In relation to Kilpatrick's description, knowledge base involves both *understanding concepts* and *understanding procedural skills*. He goes on to state that "a person who has limited knowledge of facts, definition, concepts and the knowledge of procedures is significantly handicapped in Mathematics" (p. 60). Therefore, although knowledge base is directly linked to Hiebert's conceptual knowledge, Polya's understanding a problem and Kilpatrick, et al.'s (2001) conceptual understanding it goes a little further to incorporate the knowledge of procedures.

(b) Strategies

Strategy refers to the understanding of the mathematical problem and the process of solving the problem. It incorporates understanding the problem, as well as representations and computational methods used to solve such problems (Schoenfeld, 2007b). It combines Kilpatrick et al.'s (2001) strategic competence, with elements of

adaptive reasoning including a learner's capacity for logical thought, reflection, explanation, and reasoning (Kilpatrick's et al., 2001). The importance of these categories is also further supported by the arguments provided by Harel (2008) through his repeated reasoning principle which states that "learners must practise reasoning in order to internalise desirable ways of understanding and ways of thinking" (p. 902). Furthermore, the strategy categorisation is also linked to Polya's devising and carrying out the problem solving plan and to Milgram's precision when it is related to the need for the representation of accurate and exact solutions.

(c) Metacognition

Metacognition is defined as using what you know effectively and monitoring one's own thinking when involved in processes of reasoning, communication and making connections. Schoenfeld (2007b) and Livingston (1997) defined metacognition as 'thinking about thinking'. An awareness of when one does and does not understand, knowledge of how to use the available information to achieve the goal and the ability to judge the cognitive demands of a particular task are some of the characteristics of metacognition. The characteristics of metacognition include the knowledge of what strategies to use for what purposes and when the exam takers are able to assess their progress both during and after the performance (Livingston, 1997). This characteristic therefore links metacognition to the strategy categories as it is made visible through how the strategy is used to obtain the solution. A learner who is able to apply the knowledge learned in the classroom to other contexts is also displaying metacognitive skills (Schoenfeld, 2007b). This is because, applying learned knowledge in a novel situation or a different context requires a learner to make connections between that knowledge and the context. Pugalee (2001) emphasised that, "the monitoring of one's mental activities, is essential to employing the appropriate information and strategies during problem solving. Learners should be aware of their mental processes such as; planning, monitoring and evaluating" (p. 237). Metacognition does not relate directly to any of Hiebert's distinct elements of knowledge, but it does relate to Polya's looking back. It also relates to a partial section of Kilpatrick, et al.'s (2001) adaptive reasoning in relation to it demanding the learners' logical thought and reflection during problem solving.

(d) Beliefs and disposition

Schoenfeld's fourth strand is beliefs and disposition. In line with Kilpatrick et al. (2001), Schoenfeld (2007b) maintains that a "mathematically proficient person should be able to see Mathematics as sensible, useful and worthwhile, and have a belief in diligence and one's own efficacy" (Schoenfeld, 2007b, p 68). In other words if a learner believes that Mathematics does not make sense, and that solving Mathematics problems involves worthless operations or symbols, the learner's responses will be senseless. Therefore, beliefs are important and the love of Mathematics matters (Schoenfeld, 2007b). Beliefs and disposition do not seem to have a direct link to Polya, Hiebert and Milgram's knowledge categorisation, but they are directly linked to Kilpatrick, et al.'s (2001) productive disposition.

Although Schoenfeld (2007a&b) views a mathematical knowledge base as an important component of proficiency, he emphasises that good mathematical problem solvers are also resourced with a number of strategies; and they are flexible in formulating, representing and solving mathematical problems using what they know. In other words, he has also identified strategy and metacognition categories as important and distinguishable. According to him a mathematically proficient learner should be able "to devise alternative approaches when they get stuck or hit road blocks; they are efficient with and make use of what they know" (Schoenfeld, 2007b, p. 60). He further stated that understanding the concept should not just mean being able to compute the answer to a Mathematics question which involves that concept. There is a need to gain a better understanding of the cognitive and metacognitive processes involved in problem solving, and to obtain more specific information about the major problem solving attributes that have been reported to have an influence on the problem solving process. In summary, he and others authors involved in this debate, proposes that proficiency in Mathematics be described as the ability of a person to comprehend, calculate, reason, justify, reflect and show the love and usefulness of Mathematics (Kilpatrick et al., 2001, Schoenfeld, 2007b and Burkhardt, 2007).

Schoenfeld (2007b)'s categorisation appeared to be concise, broader or more general than those of both Hiebert, Polya, Milgram and Kilpatrick's et al., 2001 (Schoenfeld & Kilpatrick, 2008). His first category of knowledge base incorporates conceptual understanding and knowledge of other mathematical facts including the knowledge of procedures and other representations. His second category, which is strategies, integrates all skills that indicate that the learner can carry out the mathematical procedures to solve the mathematical problem, it incorporates

Kilpatrick's strategic competency. His elements also include reflection under the category of metacognition as discussed earlier. However, Schoenfeld (2007b)'s categorisation seems not to have directly incorporated the skills of reasoning and argumentation, which are generally seen as important competencies needed by a mathematically proficient learner (Kilpatrick et al., 2001, Schoenfeld & Kilpatrick, 2008). For these reasons, Schoenfeld's (2007b) categorisation was chosen and clarified for the study, in order to improve what Groves (2012) construct validity for assessment of mathematical proficiency strands.

The study's objective was to investigate the possibility of assessing multiple elements that characterise knowledge (mathematical proficiency) with an attempt to understand the learners' level of achievement in mathematical proficiency. In this case, to enable valid assessment, through high stakes examinations, of a construct like, mathematical proficiency, whose elements could be distinct, it was vital to have the elements clearly characterised (Moss et al, 2006). Additionally, this distinction appears to have been incorporated in some way into most proficiency systems informing this study. Therefore, the distinction between conceptual and procedural knowledge in Mathematics was used here to argue for the importance and the possibility of making distinct proficiency categories visible in the high stakes examination artefacts as a possible way of enhancing Mathematics examinations. The view of separate aspects of knowledge taken by this study was used to provide a lens for enabling individual assessment of multiple aspects or characteristics Schneider & Stern, (2010b) of mathematical proficiency. Understanding the assessment of separate aspects of mathematical proficiency could provide insight that could be used to form a coherent understanding of how this single construct of mathematical proficiency could be assessed.

The mastery of Mathematics could be demonstrated through assessment activities. Thus Mathematics assessment, particularly high stakes assessment, should be used to model the mathematical knowledge and skills that would encourage the learners to demonstrate that they are proficient in Mathematics. This study first investigated the feasibility of adapting the curriculum process standards in order to assess mathematical proficiency more effectively through high stakes assessment, particularly in the case of the NSSCO Mathematics examination. To do this, appropriate and distinguishable aspects of Mathematics that could be used to define a mathematically proficient learner needed to be identified. The distinguishable aspects of Mathematics would facilitate the generation of the curriculum and assessment tools in order to develop the assessment framework that would more effectively assess the learners'

proficiency in Mathematics. As a result, a single formulation of proficiency was constituted as per discussion above. In order to identify an appropriate mathematical proficiency categorisation, there was a need to study the mathematical proficiency and Mathematics problem solving literature relating to how learners learn Mathematics successfully. The selected categorisation testing should be able to ensure that the high stakes assessments tasks were designed to cover all of what is required for the learners to learn (Burkhardt, 2007).

3.6 Including mathematical proficiency categories in high stakes assessment

The goals of assessment and its mechanisms are complex, particularly those of high stakes assessment (Wragg, 2001, Schoenfeld, 2007a; Groves, 2012). Different educational stakeholders have very different needs for the results of high stakes assessments (Wragg, 2001, Burkhardt, 2007, Schoenfeld, 2007a). This demands the need for these assessments to produce outcomes that can be interpreted in different ways. Mathematical proficiency has been presented as a multi-stranded construct by various authors such as (Kilpatrick et al., 2001; Schoenfeld, 2007a) as discussed above. Mathematical proficiency strands are said to be interconnected (Schoenfeld, 2007a). Accordingly, its assessment calls for assessment strategies that focus on a broad range of skills (Schoenfeld, 2007a). However, due to the fact that there are different purposes of high stakes assessment, there is a need to present the outcome of the high stakes Mathematics assessment in different forms in order to better address the various purposes.

Traditionally, parents use high stakes assessment results to know their child's aggregated mark reported in the form of a grade and in terms of class position in relation to the other learners in a class (Wragg, 2001, Schoenfeld, 2007a). These parents would also want to know how that relates to the previous year's group of learners. The same comparison would also be demanded by politicians who would want to know how the schools in their region performed. Hence, what they are interested in is one single mark that can easily be related to previous examination outcomes. On the other hand, teachers have an additional interest to the above, "they also want the assessment results to help both the learners and teacher to understand what the learner knows, and to identify areas in which the learners need improvement" (Schoenfeld, 2007a, p. 9). While, all these needs were seen to be important, the study was designed to strengthen the impact of the second one. This study endeavoured to find ways to distinguish between different proficiency elements through assessment of learners in particular groupings. This would help teachers to develop a better understanding of different learners' mathematical understanding,

as can be revealed by appropriate feedback given to them, based on the previous cohorts' performance in these elements. It would also help improve their teaching in order to prepare the learners for examinations but at the same time the study valued and preserved the importance of reporting a single mark.

Additionally, the proficiency system investigated in this study includes aspects that require learners to respond to some questions in structured ways which could include description, definition and explanation. These formats of answering are regarded as more subjective than objective routine calculations (MoE, 2010c). Objective questions are said to repeatedly produce more reliable results than open-ended answers or essay type questions (Meadows and Billington, 2005). Learners perceive the examinations as representing what they are expected to know and these are used to set standards of learning for the future examinations. In view of Burkhardt (2007) "What You Test Is What You Get" (WYTIWYG) principle, the researcher believes that extending the proficiency aspects tested by the NSSCO Mathematics examination to include all aspects of mathematical proficiency demanded would promote the teaching and learning for all mathematical proficiency aspects rather than only demanding memorised algorithms in order to rank learners against each other (Khairani & Nordin, 2011).

3.7 Conclusion

In summary, this chapter presented the literature reviewed in order to inform the conceptualisation of this study and to extract the competencies that define different elements of mathematical proficiency. This review included the literature on mathematical capabilities of Mathematics learners and Mathematics assessment which was specifically studied for the developmental process. The research methodology used to adapt the NSSCO Mathematics process standards in order to reflect the identified proficiency traits is explained in the next chapter, chapter 4.

CHAPTER 4: Research Methodology and Design

4.1. Introduction

The research reported in this thesis follows an interpretative exploratory research approach (Richey & Klein, 2005) aligned with a developmental research process which collected and analysed data through a mixture of qualitative and quantitative methods. The research data is gathered from different sources by using different procedures.

This chapter presents the methodology adopted for this research. It describes the steps that were taken to address and answer the research questions. Any research study is based on some underlying philosophical assumptions which position the study in a worldview and within a social context. The chapter starts by outlining the philosophical underpinnings (the underlying ontological and epistemological assumptions) of this research and then links them to social research methods within a developmental and an applied exploratory research process. The research methods used and the design are also presented.

This chapter begins by presenting the research objectives and questions in section 4.2, followed by the philosophical underpinnings of the research in section 4.3. In section 4.4 the research methods and methodologies used and their relevance to the study are discussed. Section 4.5 describes the sample used for the study followed by the justification for the selection method of the research design. Section 4.6 discusses the research instruments used. The data collection methods are discussed in section 4.7. Section 4.8 discusses the procedure followed to analyse the data. The validity and the reliability of the study are justified in section 4.9, and finally section 4.10 and 4.11 discuss the ethical considerations and the limitations of the study.

4.2. Research Objectives

The study was designed with the general goal of gaining an understanding of how the assessment of mathematical proficiency may be improved in high stakes assessment, particularly in the case of the NSSCO Mathematics examination.

The study was a design and development project whose process was directed to achieve three aims, each of which is subdivided into a number of objective(s). These aims and objectives are given as follows:

1. To investigate the possible design and adaptation of curriculum and assessment tools to incorporate conceptions of proficiency. This included the following objectives:

(a) To identify appropriate proficiency categories in the literature relating to mathematical proficiency.

(b) To investigate the feasibility of adapting the curriculum process standards and proficiency rubrics in order to more effectively assess the identified categories of mathematical proficiency in the NSSCO high stakes assessment.

(c) To investigate the distinguishability of the identified elements of proficiency, by using the generated assessment tools on the sample of recent examination question papers.

2. Use the designed tools to investigate the distinguishability of different proficiency elements in the learners' responses to the current question papers.

3. Use the designed tools to generate insight into the current examination system This included the following objectives;

(a) To generate insight into the conceptualisation of the current examinations, using the identified proficiency categories and developed tools.

(b) To investigate the constraints and affordances of the current examination system in relation to the identification and presentation of each of the identified proficiency categories, as well as the possibility of overcoming these constraints by adapting the question papers.

4.3. Philosophical Underpinnings

Research is characterised by Kothari (2004) "as a process of searching for knowledge, a process of collecting, analysing and interpreting data in order to find the solution to a particular problem" (p. 18). The main purpose of research is to discover answers to research questions through the application of scientific methods. The study was designed with the general goal of gaining an understanding of how the assessment of mathematical proficiency may be improved in high stakes assessment, particularly in the case of the NSSCO Mathematics examination. Accordingly, the possibility of differentiated assessment of different elements of mathematical proficiency in high stakes examinations of Mathematics was investigated. The methodology and methods followed in searching for knowledge in this study were directed by both the researcher's view of reality (ontological assumptions) and the assumptions of how people learn (epistemological assumptions). Ontological and epistemological aspects concern what is commonly referred to as a person's worldview, which has a significant influence on how the person perceives the relative importance of the aspects of reality (Wand and Weber, 2008).

With regard to the study of social phenomena, Grix (2004) defines ontology as the study of assertions and conventions that are made about the nature of reality. She goes on to say ontology is the claim about what exists as reality, what reality looks like and how elements that build it up interact with each other and form social relations (Grix, 2004). Ritchie, Lewis, Nicholls and Ormston (2007) view the key ontological questions for social science, as concerned with whether or not there is a social reality that exists independently of human conceptions and interpretations or whether there is a shared social reality or only multiple, context-specific ones. The ontological position of the researcher is that of the existence of a shared social reality, a realistic ontological stance (MacDonald, 2001).

The way of understanding and explaining how we learn and what we learn, according to (Crotty, 2003) is the researcher's epistemological position. According to Cohen, Manion, & Morrison, (2007), epistemological assumptions are concerned with how knowledge is created acquired and communicated, what they call "what it means to know" (Cohen et al., 2007, p7). The epistemological position of the researcher is that knowledge is relative to an individual context. The construction of knowledge is motivated by the need for it to be used and communicated. In this case, knowledge is constructed before assessment of this knowledge takes place and assessment questions allows the knowledge an individual have to be made visible through their response to these questions.

The ontological and epistemological assumptions made in this research may be aligned with the philosophy of social constructivism. Knowledge and social reality are constructed as a result of interaction between humans and their environment (Scotland, 2012). The development and transmission of this knowledge takes place in a social context. This research study seeks to understand and explore the assessment of mathematical proficiency as related to high stakes assessment in a quest to enhance teaching and learning for mathematical proficiency. In relation to this study learners externalise their constructed knowledge through their responses to the examination questions as a result of their interaction with the question papers. Thus, the ultimate goal is to ensure that the examination artefacts create an appropriate environment to enable the learners to show the knowledge they have. In this regard, the examination questions need to be design in such a way to allow the learners to construct a response that reflects their proficiency in mathematics.

4.4. Research methods and methodologies

During this study, the possibility of enhancing the high stakes assessment of mathematical proficiency was investigated, through the development and adaptation of some of the tools of the Namibian examination system, as shown in Figure 4.1. The research followed three different stages closely related to the research objectives.

Stage one dealt with the adaptation of the curriculum process standards and the development of the general proficiency-complexity rubric (see Figure 4.1). The curriculum process standards that describe measurable competencies in Mathematics were adapted in order to incorporate mathematical proficiency thinking. Subsequently, the general proficiency-complexity rubric that would guide the assessment of the adapted standards were also developed. The general proficiency-complexity rubric were used in the development of the specific question papers rubrics for the 2013 and 2014 NSSCO Mathematics examinations. The distribution of the mark allocation was then analysed in order to investigate the distinguishability of the two constructs in these marking rubrics.

In stage two, a sample of learners' scripts was marked using these adapted rubrics. These results were then analysed to investigate the distinguishability of the different elements of proficiency in the learners' responses. In the final stage, the tools and results generated in the first two stages were used to assist in creating a better understanding of the current NSSCO system. In addition, the possibility for improving the visibility of mathematical proficiency in the learners' responses through adapting the question papers was investigated in this stage.

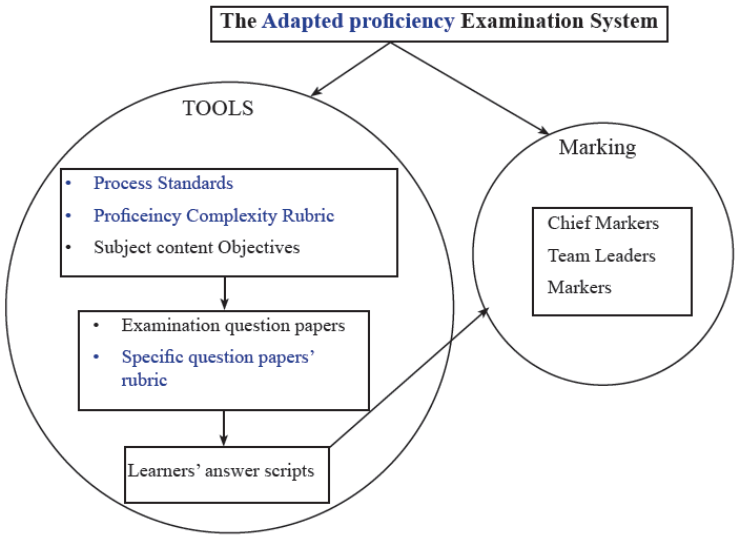


Figure 4.1. The adapted proficiency examination system

The study followed a design and development research methodology aligned with an exploratory research design. The research made use of both qualitative and quantitative research methods in two different ways. During stage one of the research, qualitative methods were predominantly used to gather and analyse the data. Stage two and three both followed quantitative and qualitative approaches, which were intentionally integrated as components of the research.

4.4.1 Design and development research methodology

Design and development research is defined by Hasan, (2003) as a methodical investigation conducted in order to develop a product or a program in context for the purpose of improving either the product or the program being developed, or the person(s) developing it. Educational development research is a systematic process which involves evaluating and designing educational programs, processes and products that should meet the criteria for internal consistency and effectiveness (Richey, Klein, & Nelson, 2004; Plomp, 2013). According to Plomp (2013) educational developmental research is done "with a dual aim of generating research-based solutions for complex problems in educational practice and advancing our knowledge about the characteristics of these interventions" (p.16). Richey & Klein, (2005) added that "in developmental research new procedures, techniques, and tools are established upon a systematic analysis of specific cases" (p. 24). They argue that developmental research in instructional design, aims to improve the process of instructional design, development and evaluation based on either solving a context-specific problem or a generalised inquiry of procedures (Richey & Klein, 2005). Hence, the general objective for design and development research is to establish an experimental base for the creation of instructional and non-instructional products (Richey & Klein, 2007). To do this, educational developmental research follows a systematic design process. This process consists of four stages of development as shown in figure 4.2. Figure 4.2 illustrates the adapted version of Amiel & Reeves (2008) approach to developmental research.

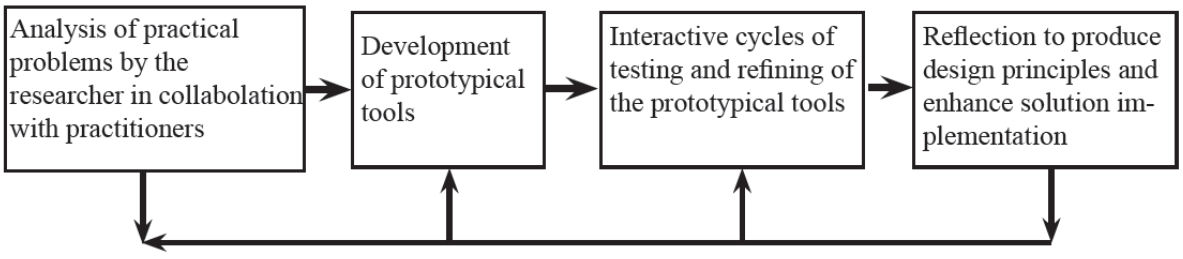


Figure 4.2. The developmental research process

According to Richey et al. (2004) design and development research is divided into two categories: Type 1 and Type 2. The difference in the two types of design and development research is brought about by whether the conclusions drawn from the research are generalizable or context specific (Richey & Klein, 2005; Plomp, 2013). Type 1 development studies focuses on a specific instructional product, program design, development process or tool. According to Richey & Klein (2005), Type 1 development studies do not only address the product design and development but also the evaluation of such a product. Furthermore, products of Type 1 developmental research projects are lessons learned from developing specific products and analysing conditions that facilitate their use. The conclusions drawn from Type 1 studies are context specific (Richey et al., 2004). On the other hand, Type 2 developmental research focuses on a given design, development or evaluation tool, model or process. The products from Type 2 developmental studies are new designs, development and evaluation procedures and/or models and conditions that facilitate their use (Richey et al., 2004). In general, the conclusions drawn from Type 2 studies can be generalised.

The study reported here is classified as a Type 1 development study because of the following reasons. It focuses on the design and development of an examination system that could better assess various aspects of mathematical proficiency which is based on the NSSCO Mathematics examinations. The evaluation of the developed products was done using answers scripts of past learners for the NSSCO Mathematics examinations and no attempt was made to generate general design principles as a general characteristic of type 2 research. However, despite the fact that the study was designed based on the NSSCO Mathematics examinations, its conclusions may be generalized to other suit of Grade 12 high stakes Mathematics examinations including the Namibian Senior Secondary Certificate Higher level (NSSCH). The study could thus inform the general high stakes assessment of mathematical proficiency, an area that has not had a great deal of development. As a result, the study was positioned within an exploratory research design.

4.4.2 Exploratory research design

Exploratory research is undertaken to explore or to provide a better understanding of a situation where little is known (Kumar, 2011). The exploration of new phenomena in this way might help the researchers to better understand the phenomena. It might also test the feasibility of a more extensive study, or determine the best methods to be used in a subsequent study (Dawson, 2002; Richey & Klein, 2005). For these reasons, exploratory research is broad in focus,

providing insight and generating hypotheses, but it is not designed to provide conclusive validation of specific research hypotheses. The exploration of new phenomena in this study addresses the researcher's interest in better understanding the assessment of mathematical proficiency. This research study is classified as exploratory for the reason that a new phenomenon of assessment of mathematical proficiency in high stakes assessment, is being investigated. Little is known about testing mathematical proficiency through high stakes examinations (Burkhardt, 2007).

When a study is defined as exploratory research, Richey & Klein, (2005) noted that the research needs to be guided by well-defined research objectives or clearly stated research questions. They further cautioned that a researcher should not just call the study "exploratory" for the lack of a definition. Exploratory studies are also known as "formulative research, because they are designed to formulate or develop working hypothesis from an operational point of view, which can be used in further investigations" (Kothari, 2004, p.35). The objective of exploratory research is to identify key issues and key variables within the explored phenomena (Kothari, 2004). For example, for this particular study the outcome was a Mathematics examination system that could better distinguish and assess different elements of mathematical proficiency. The working hypothesis and key issues would therefore include hypotheses about the distinction and interaction of the proficiency categories, as well as the insights into proficiency, developed through this research.

According to Kothari (2004) a researcher conducting exploratory research might use three different methods as follows:

(a) Conducting a literature study or what he terms a survey of concerned literature search.

(b) Conducting focus group interviews or what Kothari (2004) calls experience survey and

(c) Analysing 'insight-stimulating' examples.

This study is conducted using all three methods followed by experimental testing in order to evaluate the developed tools. The researcher studied the literature on mathematical proficiency following the literature study method, where related literature was studied to extract key issues and concepts that might define proficiency. The NSSCO Mathematics examination past question papers and their marking schemes for the years 2010, 2011 and 2012 were studied, thus analysing 'insight-stimulating' examples. Finally, this was followed by the use of a focus group of expert examination markers who assisted in the design of the research examination tools and evaluated the designed tools.

4.4.3 Qualitative methods

Qualitative research is used to explore attitudes, opinions, behaviours and experiences through several methods (Dawson, 2002). This is done through collecting, analysing, and interpreting the data of what people do and say (Patton, 1990; Dawson, 2002; Cohen et al., 2007). Qualitative data refers to the meanings, concepts, definitions, characteristics, metaphors, symbols, and descriptions of things (Cohen et al., 2007). Qualitative research is much more subjective than quantitative research and uses very different methods of collecting information, mainly individual, in-depth interviews and focus groups discussions. In line with qualitative research methods, during the third phase of this research a relatively small focus group was conducted. Participants were asked to give their views and the group moderator probed and explored their responses to identify and define their perceptions, opinions and feelings about the topic being discussed as recommended by (Cohen et.al, 2007). Cohen et. al (2007) further state that the quality of the findings from qualitative research is directly dependent upon the skills, experience and sensitivity of the interviewer or group moderator. For this study, qualitative research methodologies were used to collect and analyse the data that responded to the research objective **1 (a):** to identify appropriate proficiency categories in the literature relating to mathematical proficiency and to investigate the feasibility of adapting the curriculum process standards and proficiency rubrics in order to more effectively assess the identified categories of mathematical proficiency in the NSSCO high stakes assessment. The data that responded to objective **3(b):** to investigate the constraints and affordances of the current examination system in relation to the identification and presentation of each of the identified proficiency categories, as well as the possibilities for overcoming these constraints by adapting question papers, were also qualitative by nature.

4.4.4 Quantitative methods

Quantitative research, according to Kothari (2004) is defined as research based on the measurement of quantities or amounts. It refers to counts and measures of things and it is used to explain a phenomenon by collecting numerical data that are analysed with methods that are based on Mathematics (Aliaga & Gunderson, 2000; Dowsan, 2002; Kothari, 2004). The measurement used in quantitative research must be objective, measurable and statistically valid. In other words, quantitative research is about numbers and hard objective data (Cohen et al., 2007). Additionally, Cohen et al. (2007) defined quantitative social research as social research that employs empirical methods and empirical statements. The data that was obtained in order to respond to the research objectives **1(b), 2** and **3(a)** was quantitative data constituted

by allocated marks and the learners' marks. This quantitative data was used to investigate the distinguishability of the selected proficiency elements and generate insight into the structure of the examination papers.

4.5. Research sample

The first part of the study which dealt with adapting the curriculum process standards was carried out through a document study, while the remaining part of the study used a focus group as a tool to gather research data (Vaughn, Schumm, & Sinagub, 1996). The sample for the first part of the study, the adapting of the curriculum process standards, was a set of documents that were classified into two different groups. The first set consisted of several documents on the literature that could help to define a mathematically proficient learner and the assessment of several Mathematical proficiencies. The second set consisted of documents which formed part of the NSSCO Mathematics examination framework.

In order to narrow down the literature used, from the first set of documents, the researcher categorised the literature documents studied into two groups. The first group consisted of documents containing a general description of mathematical proficiency that did not formulate their own proficiency framework. The second group contained documents that defined a mathematical proficiency framework with distinguishable characteristics. These appeared to have contributed to the discussions contained in the documents of the first group. This second set of the reviewed literature, was narrowed down to two books:

(a) Adding It Up: Helping Children to Learn Mathematics the 2001 edition by the National Research Council, edited by Jeremy Kilpatrick, Jane Swafford and Bradford Findell.

(b) Assessing Mathematical Proficiency, the 2007 edition, edited Alan H. Schoenfeld.

The second set of documents analysed for this purpose consisted of selected artefacts and tools from the NSSCO Mathematics examination framework, these were:

(a) The curriculum aims and objectives which describes both teaching and assessment standards in the NSSCO Mathematics syllabus. For the sake of clarity these are here termed process standards.

(b) NSSCO past question papers for the 3 years, 2010, 2011 and 2012 together with their marking schemes. These three years were selected because they form a complete examination cycle. In the context of the NSSCO examinations, an examination cycle is

a period of time during which the examiners should ensure a complete coverage of the subject syllabuses in the examination question paper they developed.

The data that was collected during stage two of the study, were all obtained using the focus group of makers. Part of this process involved a document sample that could also be seen as consisting of two parts:

(a) The NSSCO Mathematics question papers and their marking schemes for the years 2013 and 2014. Two years were selected to increase the reliability of the outcome. The two years 2013 and 2014 happened to be the years immediately before the data gathering process started, therefore the learners' answer scripts were still available in the DNEA archives.

(b) A sample of the NSSCO Mathematics learners' answers scripts for 2013 and 2014 examinations. The sample consisted of 568 learners in 2013 and 618 learners in 2014. The size of the sample was determined by the number of scripts the markers could manage to mark within the course of four days per session.

Learners' answer scripts were selected through purposive sampling, by selecting learners within different key grade groups. The key grade groups were as determined by the NSSCO grading system, which are; A, C, E and G. In general a grade indicates a learners' overall level of achievement on a set of examination papers as required by a particular syllabus (Namibia. MoE, 2012). The key grades are those grades that are used by the institutions of higher learning as a selection criterion for their intake.

The focus group sample was selected as follows. A group of eight expert examiners and markers were selected, consisting only of senior markers in the NSSCO Mathematics examination marking groups. All focus group participants for this research were experienced markers in that, they all have a minimum of 5 years of marking in at least a team leader position in the NSSCO Mathematics examinations. They all had at least 10 years of experience in teaching Mathematics at NSSCO level and some of them were national examiners for the NSSCO Mathematics examinations. The gender of the members was not part of the selection process; however the selection came up with 4 male and 4 female participants of the group. The data for the adaptation phase was obtained from focus group comments and observations on the application of the developed system during the evaluation phase.

4.6. Research instruments

According to Cohen et.al (2007) research instruments are tools used for data collection in a research. The instruments used in stage one of the research were documents as pointed out in section 4.5 above. While in stages two and three the tool used to collect the research data was the focus group of markers. The collection of the data by focus group was facilitated through guided discussion. During the development of the specific question paper rubrics, the markers allocated marks to the expected responses. They then used these rubrics to award marks to the learners' responses as per the selected proficiency categories. The allocation of marks was recorded in each specific question paper rubric and the marks awarded to learners were recorded in a form designed to record marks per learner for various categories scored. The markers also gave their comments on their observations about constrains and affordance of the current system. These were recorded in a specially designed form attached as Appendix J.

4.7. Data Collection methods

This section describes the methods followed to collect data for the research study. For easy reference the discussion is arranged according to the research objectives.

4.7.1. Adapting the curriculum process standards and generating assessment tools

This section discusses the data collection methods followed through the first phase of the study. The data was collected in four different stages. The data collected here aimed to respond to the first research objective; 'to investigate the feasibility of adapting the curriculum process standards in order to assess mathematical proficiency through high stakes assessment' more effectively. This stage used the selected literature to extract mathematical proficiency traits that would be use in the adaptation of the process standards. The second stage was used to translate the current NSSCO standards in relation to the traits isolated in stage one. Stages three and four collected data that was aimed at developing the proficiency complexity rubric. Therefore, the process followed in this section produced two documents: the adapted process standards and the developed general proficiency rubric. The four stages are discussed in the following subsections.

4.7.1.1 Generating the adapted process standards

The data under this section was collected through document analysis method by analysis different documents and the literature that contribute to the topic. Document analysis is

described by Bowen (2009) as a form of qualitative research in which documents are interpreted by the researcher to extract meaning about the topic being studied. Kothari, (2004) called the analysis of documents with related literature as literature analysis. In order to obtain data for this section, several documents were studied and analysed. The process here followed the first two stages mentioned above. The details of these stages are discussed below;

(a) During the first stage of document analysis, several books and documents by various authors were reviewed in order to identify and extract different properties that described the competencies of a mathematically proficient learner that could be tested through high stakes examinations. The documents were carefully read so as to recognise patterns and highlight the emerging categories for analysis. Other literature related to the successful learning of Mathematics and the literature on mathematical assessment was also analysed to identify the emerging skills and competencies of a learner who could be proficient in Mathematics. The process started broadly with several documents of different authors but finally did a thorough analysis of the relevant sections of the two books mentioned in section 4.7 above.

(b) In the second stage of document analysis, the researcher conducted a textual analysis of NSSCO Mathematics examination artefacts such as the syllabus, particularly the teaching, learning and assessment standards and question papers with their corresponding marking schemes for the years 2010, 2011 and 2012. The analysis was directed by the mathematical proficiency competencies extracted from the literature in (a) above. The analysis in this section was driven by two main objectives:

(i) *The first objective was to identify overlaps*

The aim of this process was to identify the competencies and standards that are common to the NSSCO Mathematics examinations standards and the extracted mathematical proficiency competencies. In some cases, standards were reformulated to make the link with the mathematical proficiency competencies more noticeable. The reformulation process was in order to clarify the links that existed between the NSSCO standards and the extracted mathematical proficiency standards. The identified competencies and standards that were common between the NSSCO Mathematics examinations standards and the extracted mathematical proficiency competencies were termed overlaps.

(ii) *To identify gaps or mismatches.*

The aim of this section was to identify the extracted mathematical proficiency competencies that are not part of the NSSCO Mathematics process standards and thus are not tested through the NSSCO Mathematics examination. These were termed gaps or mismatches. The gaps and mismatches were used to formulate new standards that could be related to the NSSCO syllabus content.

The result of this analysis was a new set of standards which were termed expanded process standards for the sake of clarity. This is attached as Appendix F.

4.7.1.2 Generating the general proficiency-complexity rubric

The process of generating the general proficiency complexity rubric followed stages three and four mentioned in 4.9.1 above and they are discussed in detail below:

(a) During the third stage, examination question prototypes were developed in line with the expanded process standards. To do this, the NSSCO past question papers with their corresponding marking schemes, for 2010, 2011 and 2012, were used to extract exemplar questions per key grade. Thereafter, the exemplar questions were extended by either rephrasing the question or adding the features of the expected response to ensure a closer link to the expanded process standards. Some new questions were also developed to cover the competencies that were developed as a result of gaps or mismatches. This process helped to further clarify the overlaps and gaps or mismatches identified in 4.9.1.1 (b). The prototypes were aimed at specifying observable attributes that were expected to be demonstrated by learners in their responses to questions in the examination papers. The prototypes specified the characteristics, skills, or behaviours that examination markers will be looking for (Mertler, 2001).

(b) Finally, during the fourth stage the general complexity scoring rubric was designed and developed. The rubric was designed to enable the scoring of the separate mathematical proficiency categories demonstrated in each of the prototype responses (Schoenfeld, 2007a). The proficiency-complexity scoring rubric was developed to provide a scheme of assessment and the criteria for evaluating and scoring learners' performance as stated

by (Nitko, 2001). It defined the general process required per proficiency category in the hierarchy of the competency complexity (Nitko, 2001 and Brookhart & Nitko, 2007).

The separate proficiency category scales of the proficiency-complexity rubric were to be applied concurrently in each question. For example, one learner's solution can be awarded a knowledge base score showing the complexity of the knowledge applied, a strategy score reflecting the complexity of the strategy used and a metacognition score showing the complexity of the learner's own monitoring and regulation when solving the problem.

The product of stage three and stage four of the process were the general proficiency-complexity rubrics which were based on the expanded process standards developed. Figure 4.3 summarises the stages of the developmental phase.

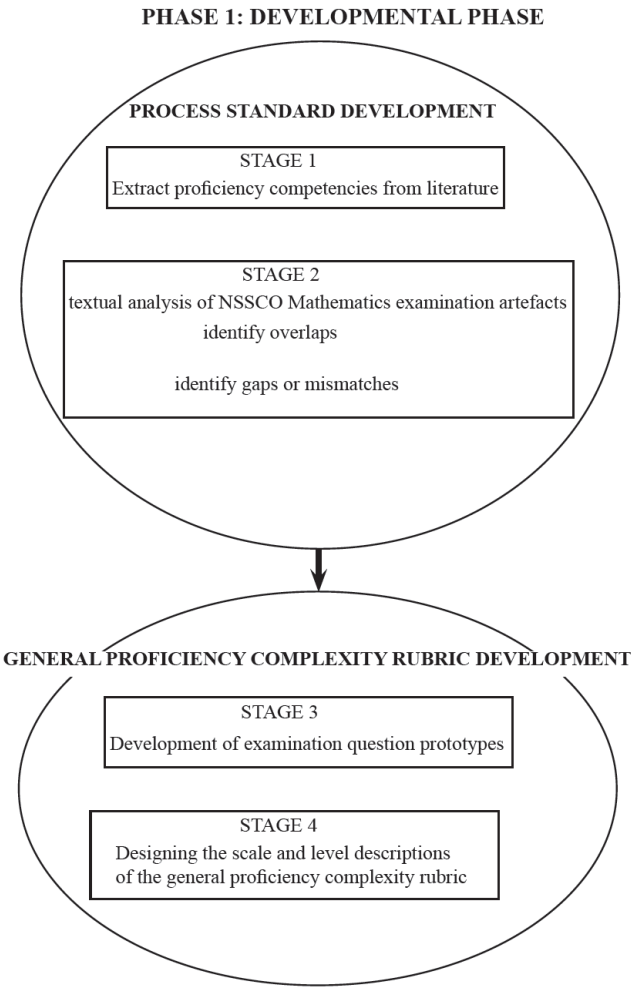


Figure 4.3. The stages of the developmental phase

The first outcome of the developmental phase of this research was a set of adapted and expanded process standards, see Appendix E. The second product is a descriptive proficiency-complexity scoring rubric that could enable the scoring of the aspects of mathematical proficiency that are made visible by the learners in their response to high stakes examinations questions (Dawson, 2002). The proficiency-complexity scoring rubric is attached as Appendix G.

4.7.2. Distinguishing the proficiency elements and their impact on the structure of question papers

The main purpose for collecting this data, was to investigate the distinguishability of different elements of proficiency, by using the generated assessment tools. The tools generated earlier as per section 4.7.1 above were the expanded process standards and the general proficiency complexity rubric. The data collected in this phase was analysed in two steps. The first step in this process was to develop specific question paper rubrics developed for the 2013 and 2014 NSSCO Mathematics examination papers in line with the expanded process standards and the general proficiency complexity rubric which was done by the focus group of markers.

The focus group, a group of eight senior markers, was brought together to evaluate the developed tools, using the learners' examination answer scripts for the 2013 and 2014 NSSCO Mathematics examinations. The group worked in subgroups of two markers per paper, i.e., two markers working with Paper1, two with Papers 2, two with Paper 3 and two with Paper 4. As mentioned earlier, a total of four papers were written by the learners who were enrol for NSSCO Mathematics examination. The NSSCO Mathematics examination is divided into two levels; the Core Level where the learners write two papers: Paper 1 and Paper 3 and the Extended Level where the learners write two different papers: Paper 2 and Paper 4. The data collected here by the focus group was used to investigate the distinction between the different proficiency elements, and generate insight into the structure of the examination papers. The process followed the following stages;

4.7.2.1 Generating specific question paper rubrics

During the first stage the group developed the specific question paper rubrics. The first meeting started with the researcher presenting the developed documents to the focus group participants. The documents presented in this meeting were the developed expanded process standards and the general proficiency-complexity rubric to be used in scoring the learners' responses. The

aim of this meeting was to develop the specific proficiency-complexity rubric per question paper for the 2013 NSSCO Mathematics examinations. This entailed applying the general proficiency-complexity rubric to all the NSSCO question papers and the marking schemes.

Using the developed proficiency-complexity rubric as a guide, the markers wrote the competencies reflected by the question responses in accordance with the selected separate proficiency categories. To allocate the marks per question, the markers first formulated expected responses to each question and then described these question responses in terms of separate proficiency categories and levels of complexity, as described in the general rubric. Based on this description, appropriate marks for the responses were assigned, to form a specific question paper scoring rubric. The number of marks allocated to each category depended on the complexity of the competencies showed in the learners' responses. That is, 1 mark was allocated for complexity level 1, 2 marks for complexity level 2 and 3 marks for complexity level 3, with a possible additional mark for the number of elements required in the response, see the complexity rubric Appendix G. The specific question paper rubrics developed were discussed by the whole group of markers to ensure a common understanding and consistent application of the general rubric. During this three day meeting the developed documents were also refined. Thus, the expanded process standards and the general proficiency-complexity rubric were finalised with the input of the focus group. Although the participants worked in groups of two markers, the discussion included the whole group with the researcher as a moderator.

4.7.3. Remarking of learners' answers scripts

The sample of learners' answer scripts used for the evaluation process had already been marked in red through the original NSSCO examination marking system. As a result of this, original scores had already been assigned to every learner in the sample. The evaluation process discussed in the next section explains the process of remarking using the newly developed marking criteria given by the specific question paper rubric.

4.7.3.1 Remarking of the 2013 NSSCO sample of learners' answer scripts

The marking of the 2013 sample of learners' answers scripts took place a week after the activity in 4.7.2.1 was finalised. In order to refresh the participants' minds, the general proficiency-complexity rubric was first presented and explained to the focus group of 8 senior expert markers. Thereafter the group worked in groups of two markers per paper (Paper 1, 2, 3 or 4)

as was done in 4.7.2.1 above. Several issues were discussed and shared across the room by the whole group. This process followed the following steps.

(a) The group reviewed the specific question paper rubrics, for the 2013 examination question papers, developed in 4.7.2.1 above. The discussion was aimed at refining the specific question paper rubrics before applying them to the 2013 learners' examination answer scripts.

(b) Using a blue pen, they remarked the 2013 learners' answer scripts, using the following steps

(i) They first both marked five scripts together, one script at a time, to ensure a common understanding on how to apply the specific question paper rubric and enhance the reliability of each marker's work. Some discussions at this stage were also shared in the whole group to ensure a common understanding. The first five answer scripts were selected from the three different key grades for a particular paper, to make sure a common understanding was obtained within varied performances of the learners.

(ii) After the five common scripts were remarked, the markers were encouraged to proceed with the marking individually but still sitting in the same group to encourage discussions where necessary.

(iii) The markers recorded the learners' marks per question and per proficiency category and thereafter recorded the sum of the individual category marks on the same recording form. That is, for each question they indicated how many marks of the available score were awarded per separate category, i.e for Knowledge Base (KB), Strategy (S) and Metacognition (M) and their sum as the evaluation marks (EM).

(iv) They were also requested to record the learners' original marks (OM) on the same recording form as the new set of marks.

(c) The markers were requested to only use a blue pen to do the remarking because the answer scripts had already been marked with a red pen, a green pen if they were moderated and a black pen if they were moderated for a second time during the original marking session.

(d) A total of 568 answer scripts were remarked for the 2013 examination across the papers.

(e) The markers met to discuss their observations of the first round and give their general feeling on how the questions could possibly be adapted to better suit the expectations of the developed proficiency examination system.

(f) The markers were asked to summarise their comments and observations in their own handwriting on a specially designed form, attached as Appendix J.

(g) After the discussion the group was asked to work in the same pairs to develop the specific rubric on the 2014 question papers in preparation for the second round of remarking of the 2014 sample of the learners' answer scripts.

4.7.3.2 Remarking of the 2014 NSSCO sample of learners' answer scripts

The meeting for remarking the 2014 sample of learners' answer scripts was conducted a week after the 2013 sample scripts were marked. The gap between the two rounds was kept as short as possible to ensure that the group members did not forget the application and jeopardise consistency in the application of the rubric.

(a) This round followed exactly the same process as the marking of the 2013 scripts which was elaborated in 4.7.3.1 above.

(b) The marking of the 2014 sample seemed to have been a bit quicker because the markers seemed to have been more comfortable and perhaps were much clearer on what to do. In this round a total of 618 answer scripts for the year 2014 across the four different papers were used in this evaluation.

4.7.3.3 Remarking the remarked sample of 2013 and 2014 learners' answer scripts

After the first level of analysis, the specific question paper rubrics were reworked to review some of the complexity level marks allocation. The same sample of learners' answer scripts were remarked again to ensure that the marking also considered the level of complexity of the learners' performance. After this round the recording of learners' remarking marks included the complexity level marks.

4.7.4. Generating insight into the current examination system

The mark allocation shown in the specific question paper rubrics developed in section 4.7.2.1 were used in this section to investigate the conceptualisation of the four question papers in line with the criteria described by the adapted process standards. This made up the first kind of data collected for this section. Insight into the current examination was also obtained from the markers' input.

At the end of each round of remarking, each focus group member was requested to write an individual detailed qualitative report. The markers were requested to comment after every round of the remarking process. The comments included their general observation on different

aspects as a way of comparing the two systems of marking and the usability of the proficiency marking system in marking separate categories of proficiency. This data was used to investigate the constraints of the current examination system in relation to the identification and presentation of each of the identified proficiency categories, as well as the possibility of overcoming these constraints by adapting question papers. The markers also highlighted constraints and suggested possible adaptation of the NSSCO Mathematics examination system, for both the design of the questions and the overall construction of the question papers. The data obtained under this section was also enriched by the markers' informal comments made during the marking, which were recorded by the researcher as a moderator.

4.8. Data Analysis

This section discusses the processes followed in analysing the data collected for this research project. The first part of the study gathered data through literature review and document analysis of examination materials in order to develop the process standards and the complexity rubric. The second phase generated quantitative data from marking scripts and this data was analysed statistically. The third phase involved collating and analysing the empirical qualitative data collected through the contribution of the focus group.

4.8.1 Adapting the curriculum process standards and generating assessment tools

As mentioned before this was the first phase of the study. The data that was used to investigate of the feasibility of adapting the curriculum process standards in order to more effectively assess mathematical proficiency through high stakes assessment was obtained through two related methods of document analysis. The first method was the analysis of the literature reviewed, which was done to extract core ideas and competencies that could define a mathematically proficient learner. The result lead to the identification of distinct mathematical proficiency categories. The literature was studied in order to identify the desirable categorisation of mathematical proficiency suitable for assessment through high stakes testing. The second method was a textual analysis of the NSSCO Mathematics artefacts. This was performed in order to bring out and clarify the core traits tested by the NSSCO Mathematics teaching and assessment framework and resulted in the generation of the process standards and the general proficiency-complexity rubric that characterised the adapted assessment system.

4.8.1.1 Literature Analysis

Literature analysis is defined by Lin (2009) as a qualitative analytical method where a researcher reads through, analyses and sorts the appropriate literature in order to identify the essential part of what is written. In the context of this study, literature analysis was used to identify one mathematical proficiency categorisation that can be easily accessed through high stakes assessment. The analysis of literature was done in a way to be able to cluster and separate traits and competencies of a mathematically proficient learner into different categories.

The literature analysis here involved identifying key categories from different authors on mathematical proficiency and translating them into one another. The term 'translating' in this context and as described by Thomas & Harden (2009) is referring to the process of taking concepts from the work of one author and recognising the same concepts in another author's work, though they may not be expressed using identical words.

This literature analysis on mathematical proficiency directed the researcher to decide to use part of Alan Schoenfeld's (2007a) proficiency formulation as the basis for suitable categories for describing proficiency standards that are separable and assessable through high stakes examinations.

4.8.1.2 Textual Analysis

The second part of the document analysis was a textual analysis of the existing NSSCO Mathematics syllabus and the NSSCO Mathematics examination papers for three years of examinations, 2010, 2011 and 2012. This was carried out to clearly identify the competencies demanded by the current question papers and to help to develop the envisaged examination question prototypes. The textual analysis of the process standards was performed by rewriting them using simple and descriptive (command) words. In some cases, complex sentences of the process standards were first broken down into simpler sentences in order to clarify them.

The question papers, their marking schemes and their examiner's grid for the three years were analysed in order to classify the questions into groups of their target grades (levels of performance). In this context, if the question was set with grade C as a target grade, it was expected that the learners who would achieve grade C or better should be able to obtain full marks. The number of marks per question was also reflected in this analysis. This process

guided the development of the examination question prototypes, in that some of the questions were used as a starting point and only extended to ensure the visibility of separate proficiency categories. The examination question prototypes thus developed resulted in the development of the proficiency-complexity rubric which was used during the evaluation phase.

4.8.2 Distinguishing different elements of proficiency

As the second phase of the study, the data analysed in this section was directed to respond to two objective as follows; Objectives 1(b), was to investigate the distinguishability of the identified elements of proficiency, by the generated assessment tools when applied to the sample of recent examination question papers. The second objective targeted to investigate the distinguishability of different proficiency elements in the learners' responses to the current question papers.

The data analysed in this phase was mainly quantitative data, resulting from the new allocation of marks based on the proficiency marking system for each specific question paper scoring rubric. The first part of the analysis in this section was done to investigate the distinction of different elements of proficiency, by the generated assessment tools using the resulting marks allocation. The investigation followed the analysis of the allocated marks in a two-dimensional way; marks allocated according to the proficiency category and according to the complexity level. The distinction between the selected elements of proficiency was investigated through a statistical analysis of the mark allocation according to the generated specific question paper rubric. The relationship between the mark distributions in the separate categories was observed using the inferential statistical approach of chi-square tests (Hogg, McKean & Craig, 2013). The tests were done in order to ascertain or test the significance of association between the two selected proficiency categories (Kotari, 2004) tested in all the papers for the two years' examinations.

In order to run the chi-square tests for the mark distribution in different papers, frequency counts were worked out in accordance with the proficiency categories. Different possible combinations of marks per proficiency categories formed a number of mutually exclusive groups. The count of frequencies for each group in relation to the distribution of marks in the questions and sub-questions within a particular question paper was determined. The chi-square tests were run on frequency counts. The frequencies that were needed for the chi-square test

were the frequency counts of the observations according to each group. Each test was seen as a random experiment which gave information about the distribution of proficiency/complexity for a specific paper. These frequency distributions for the different categories were then compared per pair of papers. The comparisons were done under a null hypothesis for each paper that the probability of each proficiency/complexity category was the same. Rejecting this hypothesis allowed a conclusion to be drawn at the given significance level that the proficiency/complexity distributions are different for the paper being considered.

The second part of the analysis in this section dealt with the analysis of the data used to investigate the distinguishability of different proficiency elements in the learners' responses to the current question papers. The main objective for this section is to describe the analysis performed on the data resulting from this process. In addition, the data analysis in this section was intended to investigate the visibility of the possible distinction between the selected proficiency categories in the learners' responses. Initially the learners' marks obtained during the remarking process, here termed evaluation marks (EM), were compared to the learners' original marks (OM). The first observation showed a big difference between the two sets of marks.

To begin with, the learners' marks in the different categories were normalised by converting them to a percentage of the available total marks. Then the set of data was summarised in terms of the mean marks of separate category marks per paper using the normalised marks. The mean marks of the learners' marks in knowledge base and strategy categories were calculated in the combined group of learners and subsequently in the learners' groups according to the key grades per category and per paper.

In order to observe a possible distinction of the two proficiency categories through the learners' performance, the data was statistically analysed by first evaluating the correlation between the learners' performance in the two different sets of marks. The Pearson product-moment correlation coefficient was computed for all sets of data to determine how the two variables related to each other and also how each of the two categories related to the original marks (Urdan, 2010). The strength of Pearson product-moment correlation coefficients ranges from –1.00 to +1.00. A correlation coefficient of 0.00 indicates that there is no relationship between the two variables being studied (Urdan, 2010). Whereas, a correlation of +1 shows a perfect positive correlation and reveals that for every member of the sample or population, a higher

score on one variable is related to a higher score on the other variable. On the other hand a value of –1 shows a perfect negative correlation which indicates that for every member of the sample or population, a higher score on one variable is related to a lower score on the other variable (Urdan, 2010). The data was also presented in scatterplots or scatter grams to allow for the visualization of the correlation strength between the groups of compared marks (Urdan, 2010).

Additionally, as extension to the correlation analysis, a simple linear regression analysis was done to examine the nature of strength of the relationship between knowledge base and strategy in terms of percentage variance (Higgins, 2005; Urdan, 2010). In a linear regression analysis each time there is a designated predictor variable and a designated dependent variable. However in this case, the regression analysis began with the assumption that the two variables are linearly related; that is to assume the learners' performance in the knowledge base category was the same as the learners' performance in the strategy category (Urdan, 2010). Thus, either one of the variables could be seen as either a designated predictor or dependent variable of the other variable. In other words, if the two variables were actually related to each other, we assume that every time there was an increase or decrease in the size of one variable, there was a corresponding increase or decrease in the value of the other valuable depending on their correlation as discussed by (Urdan, 2010). Therefore, either one of the categories could be predicted from the other one.

Further analysis was done on the learners' results to establish the relationship between their performance in the knowledge base category and the strategy category. To do this, differences between the learners' marks in knowledge base and strategy categories were calculated and analysed.

4.8.3 Generating insight into the current examination system

The data analysed in this section was targeted to generate insight into the current examination system using the developed examination tools. The first part of the analysis was to generate insight into the conceptualisation of the current examinations, using the identified proficiency categories and developed tools. Thus, the analysis in this section was aimed at investigating whether the identified categories and the developed tools generated any insight into the conceptualisation of the current examinations. This was done by observing and explaining the

new proficiency complexity balance and how this varied from question paper to question paper. In other words, the weight contributions of the marks as allocated to the selected proficiency categories and their complexity level per paper were analysed. The process started off by presenting and summarising the marks per category and per complexity level in simple contingency tables. The summarising process included the normalisation of data to avoid biased observation in the process of seeking relationships and probing for insight.

The second set of data analysed for this section aimed to investigate the constraints and affordances of the current examination system in relation to the identification and presentation of each of the identified proficiency categories, as well as the possibility of overcoming these constraints by adapting question papers. The data analysed here was made up of comments made by the markers on their observation as they developed the specific question paper rubric and remarked the learners' answers scripts. A content analysis was performed on this data. The data was first read while annotating different characteristics and highlighting comments that related to different issues in the discussion. According to Dey (2005) the data can be categorised using the categories that immerged from the analysis process. Potential categories were explored and identified during the reading of the data. The data was then broken down into the identified categories as follows;

(a) General observations

(b) Critique on old question paper design

(c) Possible changes or adaptations

(d) Anything else not belonging to any of the above groups

The data collected was analysed with an effort to respond to the question: what are the possible ways of adapting the question papers in order to improve visibility of distinction of the selected proficiency categories? To respond to this question, the analysis brought out two particular concepts: the markers' experience of the distinguishability of the selected categories and the markers' suggestions on the way the questions and question papers would need to be adapted. Therefore, different points and issues raised by different markers in the different categories helped to create profiles of these categories. That is, a summary of the information collected in each category was produced and they are attached as Appendices H and I.

4.9. Validity and reliability

According to Cohen et al. (2007) validity is concerned with the accuracy and truthfulness of the scientific findings, whether what is tested is what was expected to be tested. Validity born from positivist tradition, culminates from various empirical formations, which as discussed by Winter (2000) are; universal laws, evidence, objectivity, truth, actuality, deduction, reason, fact and mathematical data. The validity of quantitative research is described by Winter (2000) as a construct validity. A construct is the initial research concept which is framed through the research objectives or hypothesis and questions that determine which data is gathered and how this data is gathered (Golafshani, 2003). While in qualitative research, Winter (2000) describes validity as a conditional construct, which is unavoidably grounded in the processes and intentions of a particular research methodology.

The content validity for the development phase follows from the validity of the documents being analysed. The competencies of proficiency were extracted from several published documents on mathematical proficiency which included the core book used, the 2007 edition of *Assessing Mathematical Proficiency*, edited by Alan H. Schoenfeld. This was considered as the valid content. Moreover, the process standards from the NSSCO Mathematics syllabus and standardised NSSCO question papers which were used during this phase were also regarded as valid from their original developmental process. Finally, the adaptation and the changes effected on these documents were validated by using a group of expert examiners and markers in the focus group who discussed the adopted changes in detail.

On the other hand, reliability is concerned with the consistency, stability and repeatability of the testing or assessment method being applied, whether the results obtained would be the same if the research was repeated as discussed by (Cohen et al., 2007). An important consideration that would contribute to establishing the reliability and the validity of the methods employed in the study was to follow accepted processes and use established tools as they were designed to be used (Ellis & Levy, 2009). In this regard, the process of discussing the specific question paper rubric within the group of markers and ensuring a common understanding of their application in marking also ensured enhanced reliability of the outcomes.

The research was conducted in a social setting where it was natural for humans to react not only to what affects them, but also to what was happening to others around them. Therefore,

the researcher was well aware that it is not possible to be completely objective. Thus, the reliability of the outcome of the evaluation phase was ensured through the use of a focus group of experts in the field. As mentioned earlier, one of the methods followed by the group of markers to improve reliability was to ensure that a common understanding was reached on the application of the specific rubric developed before and during the remarking process. The question paper specific rubrics that were developed were thoroughly discussed by the pairs of the markers who marked them through the process of standardisation. Furthermore, the use of the samples from the two years (2013 and 2014) has confirmed a reasonable consistency of the evaluation outcome. Thus, the reliability of the evaluation phase was addressed through using samples from two different years.

4.10. Ethical consideration

Ethical issues in educational research refers to the moral principles which guide the conducts of the researcher. These ethical issues according to Govil (2013) and Cohen et al. (2007) are related to the methods of collecting valid and reliable data. The researcher has an ethical responsibility towards all research participants, directly or indirectly involved in the study (Govil, 2013). In this regard, permission to make use of a sample of learners' NSSCO Mathematics examination answer scripts from the year 2013 and 2014 was sought through the office of the Permanent Secretary of the Ministry of Education in Namibia. The permission was granted with an assurance that the identity of the learners whose scripts were going to be part of the sample analysed will not be revealed in the research outputs. The learners in this case were treated as indirect participants.

Furthermore, the consent of the participants of the focus group, who were direct participants (Govil, 2013) in the evaluation and adaptation phases, was requested through a specially designed form, attached as Appendix K. The consent form was designed to have four sections containing the following information:

(a) The general information section; in this section the researcher gave a brief explanation of what the research was all about and what the participants were expected to do. The nature of the research which made full disclosure of the research goals possible was also explained in this section.

(b) The study participants' information section; in this section the description of the research participants in terms of their position and their experience in the marking of NSSCO

Mathematics examinations was given. The section further highlighted that, although their participation might benefit them professionally, it was on a voluntary basis and they were free to leave the group any time during the study. It also gave the approximate time of the study in hours, which was not necessarily continuous.

(c) The third section assured the participants that their input would remain confidential and their names were not going to be mentioned during the reporting of the research outcomes as recommend by Dawson (2002) as an appropriate ethical practice. The participants were also assured that issues of confidentiality, as suggested by Cohen et al. (2007) would be addressed with the whole group together to ensure that all participants understood the importance of keeping the information given during the discussion confidential.

(d) The final section provided blank writing spaces to be signed by the participants to indicate their consented to participate in the group. The section also gave the contact information of the researcher in case the participants needed to contact the researcher before and during the course of the research.

4.11. Limitations of the study

Cohen et al., 2007 describe limitations as matters and occurrences that arise in a study which are out of the researcher's control. They are unavoidable. The study was conducted in a natural working environment, in line with the researcher's work. The members of the focus group were part of the NSSCO examination marking cohort. They have worked for a number of years under the researcher's supervision. Therefore, there might be an element of conformity in their performance of the work and they could give comments which they might think the researcher wanted to hear. To minimize this risk, the researcher clearly assured the members of the group that their honest opinion was required and would not be held against them in any way. In addition, the researcher worked to respond to all comments made by the participants in a positive and non-judgemental fashion.

The study was conducted in a broader field of the NSSCO Mathematics examination, which is an ongoing and operational examinations system in Namibia. This limited the evaluation part of the study to only adapt what could be adapted to the marking schemes in line with the questions already asked. In terms of validity of examination activities, adapting questions in line with some of the expanded process standards would require a sample of the examination takers to be used, to be allowed time to be taught the new expanded process standards before they could write the exams.

In summary, this chapter presented the research methods and methodologies used to collect and analyse the research data. The data collected to address the issues raised by the research objectives is presented and analysed in the next three chapters, 5, 6 and 7.

CHAPTER 5: Research Findings and Analysis: Proficiency Categories and Assessment Tools

5.1 Introduction

This section presents the data and the analysis used to select proficiency categories and use them to generate assessment tools. In this way it responds to the first research aim: To investigate the possible design and adaptation of curriculum and assessment tools to incorporate conceptions of proficiency. This included the following objectives:

(a) To identify appropriate proficiency categories in the literature relating to mathematical proficiency.

(b) To investigate the feasibility of adapting the curriculum process standards and proficiency rubrics in order to assess the identified categories of mathematical proficiency in the NSSCO high stakes assessment more effectively.

(c) To investigate the distinguishability of the identified elements of proficiency, by the generated assessment tools when applied to the sample of recent examination question papers.

5.2 Adapting the curriculum process standards and other assessment tools

The data presented and analysed in this section was collected through the process of document analysis. The document analysis was carried out in two parts. The first part involved a literature review of documents and papers written on mathematical proficiency, with a particular interest in identifying and extracting proficiency categories and competencies which describe a mathematically proficiency learner, that may be assessable through high stakes examinations. This literature review is detailed in chapter 3 section 3.4 and the result was a formulation of mathematical proficiency standards which were categorised in appropriate proficiency categories according to the categorization of (Schoenfeld, 2007b).

Schoenfeld (2007b)'s four mathematical proficiency categories were found to be broader than the five strands formulated by Kilpatrick et al. (2001) discussed in section 3.5.2. For example, the knowledge base category is more than just conceptual understanding, it includes the knowledge of procedures and their representations (Schoenfeld, 2007b). The broader categorisation is seen by Groves (2012) as having better and clearer edges between each other which could better enable their valid assessment. The researcher's view in this regard was that the areas of intersection between some of Kilpatrick et al.'s (2001) proficiency categories are

too big compared to the area of intersection between all four of Schoenfeld's (2007b) proficiency categories. Of particular concern, was that the area of intersection between procedural fluency and strategic competency was too big to allow for valid assessment of each category. For example, the skills of 'carrying out procedures flexibly and accurately' and that of formulating, representing, and solving a mathematical problem' (Kilpatrick, et al. 2001)' seem to be made visible by the same features in the expected learner's response to a high stakes examination question only given in a written form. Thus, it would be a challenge to separately make a skill such as 'carrying out procedures flexibly' visible through writing (Groves, 2012). As a result of this observation the researcher chose and used Schoenfeld's (2007b) three categories as a base of what best describe the proficiency needed to be assessed through the NSSCO Mathematics examinations.

5.2.1 Mathematical proficiency categorisation for the study

Burkhardt (2007) sustains that assessing all aspects of mathematical proficiency is important in order to evaluate learners' proficiency in Mathematics. The assessment of all aspects of proficiency is seen to be a difficult task (Schoenfeld, 2007b). Categories like strategy, metacognition and disposition are believed to be more subtle and difficult to assess (Schoenfeld, 2007b). The adaptation of the criterion of standards was done with the aim to improve clarity and measurability of learning and assessment standards. The process was done by including the desired proficiency aspects in an effort to improve their distinguishable visibility in learners' responses to high stakes examination questions. In line with this, Schoenfeld (2007a) stated that the use of different terminologies in the language used to define intended competencies in learning standards might make a great deal of difference in assessment. Therefore, the study adapted the process standards to ensure that the appropriate terminologies in the language are used to define the desired proficiency competencies in the standards.

The proficiency categorisation selected for this study, Schoenfeld's (2007b) categorisation, formed the basis for the adaptation process. Schoenfeld's (2007b) first three categories that were selected for this study as they are discussed in section 3.5.3.3 are;

(a) Knowledge base (Kb),
(b) Strategy (S) and
(c) Metacognition (M).

The competencies that guided the adaptation of process standards are discussed here and the summarised version is presented in Table 5.1 below.

(a) Defining knowledge base competency

Mathematically proficient learners who are well versed in the knowledge base category are able to define and use mathematical concepts, facts, and relationships (Schoenfeld, 2007a; Milgram, 2007). They should be able to identify and apply basic computational facts and at the same time they should be able to identify the relationship between mathematical procedures through a description or any other representation (Schoenfeld, 2007a; Milgram, 2007). Schoenfeld (2007a) added that a learner who is equipped with knowledge base competencies should be able to describe and justify a procedure and visualise the Mathematics behind any procedure. For example; $2 \times 3 = 6$, as two groups of three items, resulting in six items all together. These learners should also be able to explain why a solution to a question is one and not the other (Voskoglou, 2008). Finally, Carlson & Bloom (2005) stated that a knowledge base competent learner should be able to explain how he or she reached a certain solution and find other ways to obtain the same solution.

(b) Defining strategy competency

A learner competent in the strategy category traits is describe by Schoenfeld (2007a) as a learner who is able to carry out a mathematical computation in order to apply a given method to solve a mathematical problem. He or she is also able to restate a given mathematical problem in a different way and come up with new statements and ideas that would help in solving the problem (Schoenfeld, 2007a & Voskoglou, 2008). The learner competent in the strategy category should also be able to devise a strategy to solve any mathematical problem (Voskoglou, 2008). Voskoglou (2008) gave the examples of these strategies as; observing symmetries; working backwards; making use of analogy; making a generalisation; re-stating or re-formulating a problem; using graphs and tables; looking for counter examples and altering the given example so it is easier to understand and solve (Voskoglou, 2008). Finally, a learner competent in the strategy category should also be able to; generate an argument in the process of coming to a mathematical solution (Kilpatrick et al. 2001 and Harel, 2008).

(c) Defining metacognition competency

Mathematically proficient learners who are competent in metacognition are said to be able to reflect on their own progress while engaged in problem solving and act accordingly

(Livingston, 1997; Pugalee, 2001; Schoenfeld, 2007a). Their process of reflecting on a method to be used could result in changing the strategy of solving the problem when they reach obstacles while solving it (Schoenfeld, 2007a). They should be able to talk themselves through solving a problem and always be aware of what strategies to use for what purposes (Pugalee, 2001; Schoenfeld, 2007a). Finally, these learners should be able to assess their own progress both during and after the process of solving the problem (Gourgey, 1998).

As mentioned earlier, some aspects of Mathematical proficiency are difficult to assess, particularly through high stakes examinations (Burkhardt, 2007). In this light, the fourth of Schoenfeld (2007b)'s categories, beliefs and disposition, was not included in this study. The category of beliefs and disposition, was discussed in section 3.5.3.2 (d). The reason for leaving beliefs and disposition out was because it was seen as posing a challenge for assessment particularly through high stakes assessment using paper and pencil method as is the case in the NSSCO Mathematics examination.

The important competencies which characterise the chosen categories are summarised in Table 5.1 below. These competencies were used in this study in the process of adapting and enhancing the NSSCO process standards during the standard development process. The standards were developed to ensure that the traits and competences they described can be validly assessed through high stakes examinations.

Table 5.1 Summary of the Emerging Mathematical Proficiency Standards

Knowledge base (KB)	**Strategy (S)**	**Metacognition (M)**
A proficient learner should be able to	**A proficient learner should be able to**	**A proficient learner should be able to**
○ Define mathematical concepts and facts ○ Identify and explain basic computational facts ○ Explain why a solution to a question is one and not the other ○ Identify the relationship between mathematical procedures ○ Explain how he or she reached a certain solution and find other ways that can get to the same solution ○ Visualize the Mathematics behind any procedure, example, 2 x 4 is two groups of four objects, giving eight all together	○ Decide/generate a strategy and use or apply it. ○ Generate a relevant and appropriate easier related problem, and use its solution to solve the related problem ○ Make use of analogy ○ Make generalisations ○ Restate or reformulate problems ○ Exploit solutions to similar problems ○ Exploit symmetry ○ Work the problem backwards ○ Present or represent solutions in a tabular form, graphs and other diagrams	○ Reflect on what is being done and change the strategies in the course of solving the problem ○ Evaluate the quality of a solution

The second part of the document analysis studied the policy and practice documents that form part of the current NSSCO mathematical assessment framework. The NSSCO Mathematics documents used to extract the baseline for this analysis were as follows:

- The NSSCO process standards attached as Appendix A,
- Curriculum/Content Standards for algebra as a selected topic attached as Appendix B
- The Key Grade descriptors (as the current holistic Rubric) attached as Appendix C.
- The past examination question papers for three consecutive years (2010, 2011 and 2012) were also analysed to provide examples of examination questions.

The two processes resulted in the generation of two documents which could allow Mathematics examination takers to make their proficiency more visible. The documents generated are:

(a) The adapted process standards

(b) The general complexity-proficiency rubric

5.2.2 Generating the adapted process standards

As mentioned in chapter 1, national examinations within any education system are guided broadly by the national teaching and assessment standards. These are translated into specific subject content standards within the subject syllabi (Brookhart & Nitko, 2007). It is to be noted that these standards drive both classroom teaching and assessment activities, including the high stakes national examinations. As a core objective of this study, this section aims to present and analyse the data that was used to enhance the visibility of mathematical proficiency in the

NSSCO standards with the aim of improving visibility of separate proficiency elements through high stakes Mathematics examinations. The current process standards, which are attached as Appendix A and the current curriculum/content standards based on algebra, attached as Appendix B, were used as a basis of the developmental process that generated the expanded process standards, attached as Appendix F. The process followed two essential steps as follows;

(a) Aligning the current process standards to the competencies aligned with each of the identified mathematical proficiency categories, as summarised in the proficiency standards in Table 5.1 above.

(b) Identifying gaps between the current process standards and the mathematical proficiency standards.

This section explains how this data was analysed and used in the development of the adapted process standards.

(a) The alignment of NSSCO standards with proficiency standards

Initially, a qualitative analysis of the NSSCO Mathematics process standards was conducted. A qualitative summative content/textual analysis method (Dey, 2005) was performed on the NSSCO process standards following a translational process. The NSSCO process standards were first broken down into shorter and more precise standards. For example, the statement '*Organise, interpret and present information accurately in written, tabular, graphical and diagrammatic forms*' was broken down in order to separate the verbs: '*organise'*, '*interpret'* and '*present*'. As part of this analysis process, key descriptive verbs for each NSSCO process standard were identified and in some cases translated in order to clarify their meaning. This was done in order to enable recognition of links to the extracted proficiency standards. In some cases, the descriptive verb was replaced by a completely new verb that was viewed as the one making the competence more visible. Table 5.2 shows some examples of how the process standards were rewritten in order to link to the proficiency standards. The complete translation working table is attached as Appendix E**.** Column 4 of Table 5.2 shows the resulting expanded standards which are attached as Appendix F.

Table 5.2 An extract of process standards translation working table

No.	Current Standard (Process)	Interpretation of the current Process standards: Discussion with possible examples	Link to proficiency standards (Schoenfeld) *Knowledge based, strategies and metacognition*	Resulting process standard (Expanded Process Standard)
1(a)	*Organise*, information accurately in written, tabular, graphical and diagrammatic forms; Production – for thinking? (1) is about representation	*Given statistical information or a list of numbers - rearrange or position them in a systematic order* either in a written, tabular / graphical or diagrammatic form. Example given a list of numbers, order them by magnitude or statistical information	*Use specified (representation or ordering) procedures to solve a problem – systematizing / ordering /→**Strategy +knowledge based***	• Present or represent solutions in a tabular form, graphs, diagrams and any other organised form and write down conclusions drawn from the information in such a form.
1 (b)	*interpret* information accurately in written, tabular, graphical and diagrammatic forms; Interpretation?	*Assign a meaning to or make sense of* information in written tabular, graphical and diagrammatic forms. Example annual rainfall distribution by month.	*Use knowledge of Mathematical concepts to come up with mathematical representation (tables, graphs and other diagrams like Venn diagrams etc.* *use knowledge of representations to come up with concepts →**Strategy +knowledge based***	
1 (c)	*present* information accurately in written, tabular, graphical and diagrammatic forms; Production – for communicating?	*show or display* of information accurately in written, tabular, graphical and diagrammatic forms	*Present or represent solutions in a tabular form, graphs and other diagrams* ***Strategies***	
2	*perform* calculations by suitable methods	use a suitable method *to do or to carry out* calculations	*Decide/generate a strategy and use it* *Use specified calculation procedures to solve a problem →calculation **strategies***	• Identify and use basic computational facts • Use a given strategy to solve a mathematical problem. • *Generate a strategy and use it to carry out calculations in solving a problem* • *Describe a strategy use and justify its use in solving a mathematical problem*

				• *Reflecting on the solution or method used to solve a problem and suggest other ways of solving it.*
3	*understand* systems of measurement in everyday use and make use of them in the solution **or solving** problems;	*Know and comprehend the nature or meaning of* systems of measurement in everyday use and make use of them in the solution of problems **(how is it measurable? Through description of how the instruments are used and be able to explain the meaning)**	*Use knowledge of Mathematical concepts to comprehend and to apply semantic knowledge in solving problems in context* →***Knowledge based***	• Describe systems of measurement in everyday use and make use of them in the solution or solving a of problems

The development of the proficiency standards therefore, followed the following steps;

- taking the NSSCO process standard,
- translating the key or descriptive verbs used to describe the competencies were reformulated to ensure that they were assessable
- redefining and aligning the process standards in line with the proficiency standards extracted from literature,
- the process resulted in the expanded proficiency standards given in column 5 and attached as Appendix E.

For example, in standard 1(c) in Table 5.2 the key word '*present*' was translated to the words '*show or display*', reflecting other verbs that could be used to describe this process. In this case, the link was easier to create since the standard itself directly matched one strategy standard from the extracted proficiency standards. Standard 2 in Table 5.2 presented a greater challenge, the verb '*perform*' was translated to '*to do or to carry out*' and the process standard itself was then reformulated as '*Generate a strategy and use it to carry out calculations in solving problems'* making it clearer and more general while at the same time highlighting that the calculations are done in the process of problem solving rather than just performing a calculation. This is important in order to encourage learning with understanding and discourage performing memorised procedures. Finally, the ideas that were originally together were combined again where necessary to form the redefined general process standards. For example 1(a), (b) and (c) were recombined with very minor changes to the original as; 'Present or

represent solutions in a tabular form, graphs, diagrams and any other organised form' and write down conclusions drawn from the information in such a form. In the process of aligning the two set of standards gaps or mismatches were identified.

(b) Identifying gaps and mismatches

The textual analysis also allowed for the identification of possible gaps and mismatches between the extracted proficiency standards and the existing NSSCO Mathematics process standards. The analysis showed that some extracted proficiency standards were already part of the competencies described by the current NSSCO standard. These NSSCO process standards were only reformulated to ensure closer alignment to the proficiency standards. However, the analysis also showed that some extracted proficiency competencies were clearly not part of the NSSCO process standards. The standards that were not part of the NSSCO Mathematics standards were identified as gaps or mismatches.

For each identified gap, a new process standard needed to be formulated to ensure that these standards spanned the corresponding proficiency category. An example of a standard that did not form part of the NSSCO standards is being able to "explain how he or she reached a certain solution and find other ways that can get you the same solution" (see KB5, Table 5.1). Also, it was noted that no standards categorised under the category Metacognition (M), Table 5.1, were included as part of the NSSCO set of general standards.

As a result of the process followed in Table 5.2, gaps between the NSSCO Mathematics standards and the extracted proficiency standards were identified and these are shown in Table 5.3 below. A description of each gap was also given, as shown in the second column of Table 5.3. The standards that resulted from the identified gaps were then added to the *redefined process standards* table. This formed the final *expanded process standards* table attached as Appendix F.

Table 5.3. Possible gaps between process standards and proficiency standards

Possible gap	Description of the gap	Proficiency standard driven
Pose questions	The ability of the candidate to formulate or reformulate a problem.	restating or reformulating problems ***– as Knowledge base and strategy***
Identify irrelevancies	The ability to identify irrelevant issues either in the problem or in a given solution to a problem	Generate a relevant and appropriate easier related problem, and use its solution to solve the related problem **– *as a strategy***
Describe multiple strategies	The ability to describe various strategies to be used in solving a specified mathematical problem	Describe a strategy used and justify its use in solving a mathematical problem And Explain how a certain solution was reached and find other ways that can be used to get the same solution. **– *as a strategy***
Model the problem	Present a simplified description of a problem in order to assist the calculation or in solving the problem.	Restate the problem in simplified words or context in order to solve the problem **– *as a strategy + metacognition***
Identify obstacles	The ability to identify obstacles and giving suggestions on how to overcome them	Given the working to a solution of a specific problem that might have reached a dead end – reflect on what has been done wrong and suggest ways on how to fix it **– *as a strategy***
Explain or Justify solutions or strategies used and produce alternative strategies	Make a fact or a concept or strategy used comprehensible OR communicate plainly the strategy used OR provide adequate grounds for the solution obtained.	Explain how he or she reached a certain solution - **Knowledge base** and find other ways that can get you the same solution - **strategy.** The whole process can be seen as (Reflect on what is being done and change the strategies in the course of solving the problem) **- metacognition** Explain how he or she reached a certain solution and find other ways that can get you the same solution **- metacognition**
Solve backwards	Given the final steps of the solution be able to work back to the initial steps. (*could be there but not standing out)*	Working backwards as a strategy to reformulate, represent and solve mathematical problems - **strategy**
Evaluate quality of solution	Assess or appraise the valuable features of the solution given – to be given in a form of description or explanation	Explain or describe the valuable features of certain solution **–knowledge base and metacognition**

5.2.3 Development of the proficiency-complexity rubric

This section describes the development of the general proficiency-complexity scoring rubrics, in order to improve the visibility of the identified mathematical proficiency elements in the NSSCO Mathematics examination. This was done by;

(a) Aligning the expanded process standards with the content standards of the algebra section of the NSSCO Mathematics syllabus.

(b) Developing a range of examination question prototypes with their prototypical responses based on the expanded process standards.

(c) Using the developed prototypes to develop the complexity rubric.

(d) Developing the specific question paper rubrics.

The documents that informed this analysis were;

- the NSSCO Mathematics content standards for algebra,
- the NSSCO key grade descriptors, with particular interest on section of algebra and
- the past question papers and marking schemes for the 3 consecutive years (2010, 2011 and 2012).

(a) Aligning the expanded process standards with the Mathematics content standards

Initially, the past question papers were used to extract exemplar NSSCO examination questions for each key grade descriptor for the topic of algebra. These were adapted in order to align them to the expanded process standards. Here the topic of algebra was used as a pilot to exemplify prototypical questions at different levels of complexity within the same topic. Additionally, algebra was seen to bring out a majority of competencies shown in other topics. However, the expanded process standards and the rest of the artefacts developed considered all topics in the NSSCO Mathematics. The process followed to align and match the more specific content standards and key grade descriptors with the expanded process standards, is described in the following paragraphs.

The key descriptive verbs of the content standards and those of the key grade descriptors were highlighted and translated where necessary. In some cases, new verbs were introduced to ensure that the verbs in the content standards that were seen as being too broad were made more specific, to ensure closer alignment to the process standards. For example, the knowledge of concepts on which most examination questions are based is often assumed, when a leaner is required to solve a mathematical problem. In other words, when a learner presents his/her calculation as an answer to a mathematical problem, it is assumed that he/she has the knowledge of the content on which the question was based. Making the verbs more specific ensures that the knowledge of such concepts is made explicit in the question, not just assumed in its application. For example, for a process standard 'identify different representations of mathematical concepts'. A clear command word '*define*' was used instead of *identify*, because before a learner can identify, he/she needs to know what it is (Brookhart & Nitko, 2007). Defining the concept was viewed to be making the knowledge of the learner more visible than

just identifying a feature. Table 5.4 shows the examples of the expanded process standards aligned with the content standards.

Table 5.4. Process standards aligned to the content standards

	Expanded process standard	**Aligned content standard**
1	Demonstrate an understanding of mathematical statements expressed in words or symbols (through transformation, representation or explanation)	Simplify algebraic fractions and explain the method used
		Transform novice/context arithmetic problems into algebraic equations by presenting variable values as letters in order to solve them. Or Generate algebraic equations or inequalities from context problems in order to solve them.
2	*Working backwards as a strategy to reformulate, represent and solve mathematical problems*	Show that a certain expression is a solution to an algebraic problem.

The above process followed two stages; firstly, examples of examination questions per key grade and per topic were identified from the past question papers and tabulated, see Appendix E. The tabulating process considered the number of marks originally assigned to a question and the key grade for which it was designed. In the second stage each specific key grade descriptor with the assigned example questions from the past question papers were matched to the developed process standards. This was done as an attempt to establish or verify the link between the existing process standards and the questions set on them. This created a foundation for the development of the examination question prototypes based on the expanded standards.

(b) Developing examination question prototypes

The development of examination question prototypes and their prototypical answers was done with the purpose of identifying specific observable attributes that an examiner would want to see in the learner's response (Mertler, 2001). The process went through several stages. Initially as a way of piloting the NSSCO content standards of algebra was used as a basis to develop prototypical questions but extended with a few more examples from other topics. The process was conducted by the researcher with the assistance of two experienced NSCCO teachers and examiners. The research supervisor acted as a validator of this process. The group took the exemplar examination questions extracted from the past question papers as the initial question prototypes. These were edited where necessary to ensure close alignment to the expanded standards. They then used their knowledge and experience on question setting to extend the prototypes in order to ensure that:

- All solutions included all the steps necessary to obtain them.

- The new competencies (those reflected in the identified gaps in Table 5.3) were also covered in the set of question prototypes.

The process therefore started with the development of the prototypical or model solution for each question. The solutions developed were analysed to identify the competencies they were making visible. This analysis led to the allocation of levels and marks to the questions. The analysis also helped to identify overlaps and mismatches to the expanded process standards. Table 5.4 shows an extract from the set of examination question prototypes developed. The link to the expanded process standards highlights the overlap and the mismatch to the existing NSSCO process standards.

Table 5.4 Examples of Mathematical Proficiency Standards and Examination Question Prototypes developed

Mathematics Proficiency Standards	**Examination Question Prototypes with model response**	**Link to Process or Examination Standards**	
		overlaps	**Mismatch**
1. Work out and interpret the solution of linear inequalities in context	1. **Q**: Find the largest integer that satisfies $9 - 2y \geq 2$ **A**: $-2y \geq -7$ $y \leq 3.5$*(MC, KB =2, S=1)* the largest integer is 3 2. **Q**: The number of passengers in a kombi at one particular time is $x + 6 \leq 16$ determine and explain the value of x (LC *KB =2, S=1, M = 0*) **A:** $x \leq 10$ The number of passengers in the kombi before the 6 passengers got on should not exceed 10.	Work out the solution of a linear inequality Represented by the solution: $-2y \geq -7$ $y \leq 3.5$	Interpret the solution of the linear inequality in context **Links to the question prototype:** 1 A: largest integer is 3 2. A: The number of passengers in the kombi before the 6 passengers got on should not exceed 10.
2. Determine and interpret factors of algebraic expression	1. **Q**: factorise $x^2 - 4x + 4$ (LC *KB =2, S=1*) **A** : $(x-2)(x-2)$ 2. **Q**: Show and explain why $(x-2)$ is a factor of $x^2 + x - 6$ (MC) **A**: $x^2 + x - 6 = (x-2)(x+3)$ When $(x-2)$ is multiplied by another factor it gives $x^2 + x - 6$	Factorise a quadratic expression Prove that an expression is a factor of a quadratic expression Simplify an algebraic expression involving fraction	Interpret or Explain why a linear expression is a factor of a quadratic expression **Links to the question prototype:** 2. When $(x-2)$ is multiplied by another factor it gives $x^2 + x - 6$ 4. Whole prototype – explaining the answer

In Table 5.4, two examination question prototypes were developed for the mathematical proficiency standard 1. Question 1 is a question on solving linear inequalities, which should

reflect the following competencies; the knowledge of solving linear inequalities involving directed numbers linking it with the concept of integers. A learner's complete response would be scored in the following way; *two* marks for the knowledge base at a medium complexity level, *one* mark for a strategy at an atomic complexity level and *zero* score for metacognition since there is no assessable metacognitive skills displayed or expected. Therefore, the correct answer should be awarded a maximum of *three* marks under the newly developed framework.

(c) Using the developed prototypes to develop the complexity rubric.

According to Mertler (2001) a scoring rubric is a set of instructions that describes the qualities of the question response that are worth crediting. He went on to say a scoring rubric is a rating scale as opposed to a checklist that is used to rate the performance of an assessment task (Mertler, 2001). The examination prototypical questions with their response development process informed the development of the general proficiency-complexity rubric. This was done following a combination of Brookhart & Nitko's (2007) rubric development approaches but more specifically it was guided by the rubric development procedures as stipulated by Mertler (2001). The examination question prototypes developed provided specific observable attributes or traits that are envisaged in the learners' response to these questions. Using Mertler's (2001) developmental approach described in Table 2.3 in chapter 2 above, the following steps were followed;

- The prototypical answers were used to highlight specific observable attributes that was expected to be seen demonstrate in the learners' responses.
- A description was written to describe attributes that represent different complexity levels of proficiency as basic or low level of proficiency, medium level and attributes that show high level for each observable features.
- A thorough narrative was written to give a descriptions for the two extreme end levels of complexity as guided by the prototypical responses given. For example, writing that a certain response only requires knowledge of a concept at a low level and it involve interaction between two or three relationships forming a coherent structure with the higher level of complexity.
- The next step required to write descriptions for the intermediate level(s) of performance for each attribute separately, which in this case involved the interaction between two or more concepts.

- The described levels of complexity were applied to several developed prototypical question and responses that exemplify each level.
- The rubric was revised as necessary.

5.2.3.1 Cognitive complexity scale for the study

In developing the proficiency complexity rubric discussed in section 2.3.2.2, the literature on cognitive complexity discussed in chapter 3 was considered. This literature was used to define the complexity scale suitable for the proficiency categories demanded to be externalised and made visible to the markers in this study. The resulting proficiency complexity rubric was designed with three different complexity levels. The complexity was defined from the basic level of atomic or unitary knowledge to the second level which recognises the links between knowledge units. The third level acknowledged the creation of links between links requiring learners to externalise knowledge that would form a response with a coherent structure.

The complexity scale was therefore designed with the influence of both sources of cognitive complexity in mind. Thus, the researcher is of the view that each knowledge category, in this case the mathematical proficiency category, has an increasing level of complexity in the way that each category's knowledge is processed as they are both influenced by these sources of complexity. The complexity moves from an atomic or unitary level to a more interactive level of complexity. The complexity levels of the selected proficiency categories (knowledge base, strategy and metacognition), used in this study, are developed based on the response mode discussed in section 3.2.2.2 of this book and are explained below. Variations of demands per level per proficiency category are shown in Appendix G attached.

(a) Knowledge Base

An atomic level of the knowledge base category is represented by recalling facts of concept which is the level 1 of complexity. Examples of a level 1 question would require a learner to define an algebraic factor. Level 2 of the knowledge base category requires a learner to show competence where atomic knowledge is linked to one or more atomic knowledge. Example questions that demand a level 2 response is; *define a quadratic equation factor and give an example*. While level 3 of the knowledge base category requires a learner to show competence where there are links between links forming an organised structure. Example; learners are required to explain a solution to a linear inequality in context.

(b) Strategy

Level 1 of the strategy category or the atomic strategy requires learners to implement a single memorised relationship or process. Example; determine the difference between two numbers. Level 2 of the strategy category requires learners to implement a composite strategy showing how it can be used and provide a solution to a problem. Example; factorise a given quadratic equation. Level 3 of the strategy category requires learners to develop and implement a strategy showing how it can be used and provide a solution to the problem. Example; write the expression $\frac{3x^2 - 6x}{x^2 + 3x - 10}$ in its simplest form.

(c) Metacognition

Level 1 of the metacognition category refers to the process of making a choice, either by selecting the best solution or choosing the best method to follow to obtain a solution. Level 2 also involves the process of making a choice but giving a reason for your choice. Lastly, the strategy choice made for level 3 has to be supported by a process-based reasoning.

In both cases the complexity increases from level 1 through to level 3 because, the writing of a response requires activation of additional grapheme codes and takes relatively longer produce (Putnam & Roediger III, 2012). The complexity should therefore be viewed within an individual's proficiency category, not necessarily in an array of the proficiency categories themselves. For example, the conceptual knowledge as a knowledge category is expected to be internalised, stored and externalised at different levels of complexity.

In line with assessment principles, the main purpose of developing and adapting curriculum standards and the proficiency complexity rubric was to ensure that the examinations based on the selected proficiency elements would be valid and their outcomes dependable. The next section discusses validity in educational assessment in relation to this study.

The first draft of the rubric formed a two dimensional scale, with the scores considering separate proficiency categories and the complexity of the response required. The two dimensions were firstly the proficiency categories while the second dimension refers to the cognitive aspects of the scores. Table 5.5 shows examples of the examination question prototypes developed including the responses. The last column of Table 5.5 shows and explain

how the developed rubric was applied to these prototypes. Thus, the question responses developed as shown in Table 5.5 included the rating score to show how the solution can be scored in line with the interdependent proficiency categories.

Table 5.5 Examination question prototypes developed, with the question response and the rubric as applied

	Examination question Prototype	**Question response**	**Response Complexity Scoring Rubric (Score: 0 or 1 or 2 or 3)**
1	In an algebraic expression, define the term coefficient.	A Coefficient is a number multiplied by a variable	M = 0 (no monitoring and self-regulation) S = 0 (no strategy used or given) KB = 1 (remember or recall a definition) **Total score 1**
2	In an algebraic expression define the term variable or constant and giving an example.	A Variable is a symbol for a number we don't know yet, in 4*x*, *x* is a variable.	M = 0 (no monitoring and self-regulation) S = 0 (no strategy used or given) KB = 2 (remember or recall a definition and link to the expression 4*x*). **Total score 2**
3*	Write down the first three terms where n^{th} is $n^2 + 4$	$1^2 + 4 = 5$ $2^2 + 4 = 8$ $3^2 + 4 = 13$ $\therefore$ 5, 8, 13	M = 0 (no monitoring and self-regulation) S = 2 (implement a given strategy showing how it can be used and provide a solution to a problem) KB = 1 (remember or recall known mathematical facts, concepts and links). **Total score 3**
4*	Explain how to simplify $\frac{x^2 + x - 6}{x^2 + 2x - 3}$	factorise the numerator and the denominator to $\frac{(x+3)(x-2)}{(x+3)(x-1)}$, cross out the common factors. The remaining fraction $\frac{(x-2)}{(x-1)}$ is the simple form.	M = 1 (choosing a strategy) S = 3 (develop and implement a strategy showing how it can be used and provide a solution to the problem) KB = 3 (create an integrated knowledge structure incorporating a number of known facts and links) **Total score 7**

The general proficiency complexity rubric which was designed was refined and edited by the focus group of expert senior markers. The markers applied the general proficiency-complexity rubric to the mark schemes of the 2013 and 2014 NSSCO Mathematics examinations. As a result, they developed the specific question paper rubric for various question papers during the first round of the evaluation phase. As a result of the developments of the specific question paper rubric, the markers suggested a third dimension of the rubric as the length of the candidate's work. The markers argued that, without the length dimension, there might be two questions that are scored within the same complexity level of a particular category but one question's response might require a learner to go through more steps than the other. The length dimension was suggested to bring in the distinction between these two questions. The length dimension is demonstrated in the following examples;

(a) Simplify the two expressions

(i) $\dfrac{x-2}{x^2+3x-10}$

(ii) $\dfrac{3x^2-6x}{x^2+3x-10}$

The two expressions require the same strategy complexity level, however the second expression requires more work since the expression in the denominator has to be factorised first.

Therefore, it was found necessary and appropriate to extend the two dimensional scale to a three dimensional four score scale, proficiency-complexity scoring rubric. Three dimension in this case meant the scoring scale considers the separate proficiency categories as one dimension, the complexity of the required response and the length of the required response as the third dimension. The rubrics developed ensured that, although the three proficiency categories could be judged separately, they remained intertwined. This meant that the selected categories were scored concurrently per question response. Additionally, when the response required more than two steps in the working, it was awarded an extra score as a length score. The developed general complexity rubric is attached as Appendix G.

(d) Developing the specific question paper rubrics.

The researcher and the group of experienced markers used the developed proficiency-complexity rubric as a guide, to develop the prototypical responses to the questions in the specific question papers. The prototypical responses were analysed and linked to the

competencies in accordance with the selected categories as described by the developed proficiency complexity rubric, as explained in section 5.2.2 (b). The team allocated marks to each selected category separately per question as they judged it suitable for each prototypical response. The results of the process above were tabulated to show the evaluation marks' distribution per question and per category as shown in Table 5.6 below. Table 5.6 shows an example of the specific question paper scoring rubric, for question 1(a) of Paper 2 for 2013. Note that, the evaluation marks distribution per question is recorded in separate proficiency categories: KB (knowledge base), S (strategy) and M (metacognition) and then totalled to yield the evaluation total (ET) for the question. The original marks (OM) is also recorded in the last column. The column 'question response' gave some detail about the prototypical question response, together with the description of the expected competency per category that such a response makes visible. The marks allocated to each category were assigned according to the complexity of the response as explained in section 5.2.2 (b). The number of marks per category also reflected the level of complexity of the expected responses. Example, when only 1 mark is awarded, the complexity level 1 is expected in the respective category.

Table 5.6 Specific scoring rubric of paper 2 for 2013 for question 1 (a).

Qno.	**Question**	**Question response**	**Response Complexity Score (Score: 0 or 1 or 2 or 3)**				
			KB	S	M	**ET**	**OM**
1 **(a)**	Calculate the value of $\frac{1}{3}\sqrt{\frac{1}{3}+\frac{1}{3}\sqrt{\frac{1}{3}}}$ Write down your full calculator display,	0.15713484(03) or $\frac{\sqrt{2}}{9}$ *KB atom: roots* *S: Provide a solution without requiring to show a strategy used*	1	1	0	**2**	**2**
(b)	Write your answer correct to 4 significant figures.	0.1571 *KB atoms: significant figures, rounding*	2	0	0	**2**	**2**

The data generated during the process above was used in the initial investigation of the distinguishability of the chosen mathematical proficiency categories, as made visible in the existing NSSCO examination mark schemes. The process reflected how practicable it was for the markers to separate the three categories, using the prototypical responses in the marking scheme.

5.3 Distinguishability of different elements of proficiency

The data collected here was in the form of mark allocation per question per proficiency category as an initial stage of the evaluation process. As mentioned earlier, to generate the

specific rubric, the group of markers worked through the examination papers in order to decide how marks could be allocated to each question in accordance with the revised assessment framework. This section was aimed to present and analyse the data that was used to investigate the distinguishability of the selected proficiency categories tested in the NSSCO system, in relation to the chosen question papers. Therefore, this analysis addresses the objective: To investigate the distinguishability of different elements of proficiency, with the generated assessment tools when applied to the sample of recent examination question papers. The data used in this section was in the form of mark allocation as per the development of the specific question paper rubric discussed in section 5.2.2 (d) above.

The data allowed the comparison of the distribution of marks, according to proficiency categories and complexity, at the level of individual questions within each question paper. The mark distribution per category per complexity level, for each of the four papers and two levels (core and extended) was then analysed, to address the research objective of investigating the distinction of different elements of proficiency, by the generated assessment tools when applied to the sample of recent examination question papers. This was done using a chi-square analysis of the distribution of marks within each question paper.

5.3.1 Chi-square test comparison of the distinguishability of knowledge base and strategy categories

The raw mark frequency distributions per proficiency category and complexity level were used to quantify the possibility that the assessment of knowledge base and the assessment of strategy were statistically distinguishable measurements in each question paper. The investigation used the inferential statistical approach of the chi-square test (Hogg et al., 2013). The metacognition category was not included because the data in the specific question paper rubric shows that this contributed very little (and in some papers nothing at all) to the total marks. A chi-squared analysis was conducted on this data set, in order to test the significance of association between the two proficiency categories in each of the separate question papers across the two years (Kotari, 2004).

The aim of performing the chi-square test was to find out whether the distribution of marks between the two proficiency categories in each paper was statistically similar, or not (Hogg et al., 2013). For each paper, the test was carried out under the null hypothesis (H_0) that the distribution of marks of the two categories were the same in that paper. To conduct these chi-

square tests, the frequency counts of the marks allocated to each category and level, for each question in the paper, were calculated as explained in chapter 4. It should be noted that the degrees of freedom for the chi-square tests varied from paper to paper, because the mark allocations for the questions in the papers did not include all of the possible cases that could occur (Hogg et al., 2013). The degrees of freedom ranged between 4 and 2 and as a result, the level of significance for the chi-squared statistic also varied from one question paper to another (see Table 5.7).

Table 5.7 Chi-squared test comparison for the two categories for the year 2013

Hypothesis H_0: These two multinomial distributions are the same (KB = S).					
	P1	P2	P3	P4	consolidated
K-1 (df)	2	3	3	4	5
Test statistic	18.04444444	14.67567568	32.4	12.37397959	66.83438872
Chidist	0.000120698	0.002115868	4.30997E-07	0.014776497	4.66402E-13
Conclusion (95% confidence)	reject H_0	reject H_0	reject H_0	reject H_0	reject H_0

Table 5.7 shows that at a confidence level of 95 percent, the null hypothesis was rejected for all papers.. Thus, the mark distribution between knowledge base and strategy categories was statistically distinct in all papers, at a confidence level of 95 percent (Urdan, 201) Moreover, in three out of four papers the null hypothesis was rejected at a 99 percent confidence level. The chidist (p-value) of 0.00012 for Paper 1 indicated that there was only 0.012 percent chance that the null hypothesis was true and that of Paper 2 is 0.002, indicating that there was a 0.2 percent chance that the null hypothesis is true. The p-value for Paper 3 was even lower than the two papers. However, the probability that the two categories are the same in Paper 4 is 1.4 percent, showing that the null hypothesis cannot be rejected at 99 percent confidence level but it can be rejected at 95 percent. This shows that the degree of sameness in the two categories is higher in paper 4 than the other papers.

The comparison between the mark allocation profiles for knowledge base and strategy in the 2014 papers yielded a similar conclusion to that of 2013. The results of this analysis are presented in Table 5.8, where again, the null hypothesis was rejected in each case.

Table 5.8 Chi-squared comparison for the two proficiency categories for the year 2014

Hypothesis H_0: These two multinomial distributions are the same (KB = S).					
	P1	P2	P3	P4	consolidated
K-1 (df)	2	3	2	4	5
Test statistic	15.71428571	13.7826087	19.63483643	32.54061625	71.6105412
Chidist	0.000386978	0.003216502	5.44941E-05	8.58806E-08	4.73496E-14
Conclusion (99% confidence)	reject H_0	reject H_0	reject H_0	reject H_0	reject H_0

In all four papers in 2014, the null hypothesis was rejected at a 99 percent confidence level. The result therefore showed that there was a statistically significant difference between the mark distributions for the knowledge base and strategy categories, in all papers, with 99 percent confidence in all but paper 4 of 2013. This supported the conclusion that the knowledge base and strategy proficiency categories were conceptually separable when applied to the assessment of these examination papers.

The chi-square test analysis was also run for the consolidated papers. In this regard, the consolidated paper was represented by the marks of all four papers added together and treated as one paper representing the skills and competencies tested in the whole NSSCO Mathematics syllabus. Just like in the individual papers, the same observation was made with the mark distribution when the papers are consolidated. Questions for both papers were combined and analysed using a combined frequency table. The probability that the two categories are the same in the consolidated paper was much less than that of the individual papers, showing that the null hypothesis can be rejected at 99 percent. Thus, the degree of sameness in the two categories was much lower than the individual papers.

The observed significant distinguishability of the two proficiency categories necessitated the need to analyse and compare how the current question papers were structured in terms of the separate proficiency category marks. The raw marks allocated to different categories yielded different mark totals. In other words, the attainable maximum raw marks for every category in the different papers was different. In order to compare marks meaningfully in different categories, it was thus necessary to normalise the marks for each category considered. The raw marks were thus normalised by converting them to percentages, either by category or by

complexity level for each paper. This provided a measure of the relative weighting of each category and complexity level in the marks allocated for each paper, which facilitated the question papers structural comparison.

5.4 Conclusion

In this chapter, NSSCO Mathematics teaching and assessment process standards were analysed and reformulated. The aim of the reformulation was to improve the visibility of separate mathematical proficiency elements assessed in the NSSCO Mathematics examination. The reformulation followed two major steps. The first step was to establish or create a link between the current NSSCO process standards and the mathematical proficiency standards. The last step was to identify gaps between the current NSSCO process standards and the mathematical proficiency standards, in an effort to expand the process standards.

Subsequent to the development of the process standards, general proficiency-complexity scoring rubrics based on the developed process standards, were developed to enable the assessment of mathematical proficiency in the NSSCO Mathematics examination. To do this, a number of NSSCO Mathematics framework documents were analysed to develop a range of question prototypes using the process standards developed as a basis. The proficiency-complexity rubric developed was evaluated by applying it to the past NSSCO examinations for the years 2013 and 2014. Subsequently, specific question paper rubrics were generated, which allowed the investigation of the distinction of the mathematical proficiency elements assessed in these papers.

CHAPTER 6: Research Findings and Analysis: Investigate the principles and relationship of proficiency and Mathematics competence

6.1 Introduction

The data collected, presented and analysed in this section was used to respond to the second research objective: Using the designed tools to investigate the distinguishability of different proficiency elements in the learners' responses to the current question papers. The data was collected through the activities of evaluating the developed tools.

6.2 The possible distinction of proficiency in the learners' responses

The data analysed in this section was obtained through the evaluation of the specific question paper marking rubric developed and analysed in chapter 4. A sample of learners' NSSCO Mathematics answer scripts, 568 for 2013 and 618 for 2014, were remarked using the specific marking rubric for this purpose. This section presents and analyses these results in order to illuminate the issues that relate to the research objective of investigating the distinction of different proficiency elements in the learners' responses to the current examination question papers.

To generate the data analyses here, the focus group consisting of experienced senior markers remarked the sample of learners' NSSCO Mathematics examinations answer scripts for the years 2013 and 2014. The sample of the learners' scripts were remarked again by a smaller group of the focus group and the researcher as the principle marker. The remarking followed the criteria set by the specific question paper rubric developed by the focus group and reworked by the smaller group. For each learner, the markers recorded the original mark and the original grade obtained by the learner as well as the evaluation set of marks. The evaluation set of marks were recorded in different proficiency categories: knowledge base, strategy and metacognition and further split into the complexity levels. The sum of these three categories, which was termed the evaluation marks was also recorded. An example of the recording sheet used by the markers to record marks for a sample of learners for Paper 3, 2014 is shown in Table 6.1. The centre numbers and learner numbers shown in column 1 and 2 of the table are fictitious.

Table 6.1. The recording sheet for Paper 3, 2014

Centre No.	Candidate NO.	Paper	Previous Mark	Previous Grade	Knowledge Base				Strategy					New Total (KB + S + M)
					level 1 (47)	level 2 (29)	level 3 (0)	KB Total (76)	level 1 (30)	level 2 (4)	level 3 (0)	S Total (34)	M (3)	113
NA000	0000	3	61	C	36	13	0	49	16	2	0	18	2	69
NA000	0001	3	64	C	35	15	0	50	17	2	0	19	1	70
NA000	0002	3	60	C	34	13	0	47	16	4	0	20	1	68
NA000	0003	3	63	C	33	11	0	44	18	4	0	22	3	69
NA000	0004	3	63	C	35	11	0	46	17	5	0	22	0	68
NA000	0005	3	63	C	33	15	0	48	17	6	0	23	1	72
NA000	0006	3	59	C	34	12	0	46	16	4	0	20	0	66
NA000	0007	3	60	C	35	11	0	46	19	0	0	19	2	67
NA000	0008	3	60	C	34	14	0	48	18	4	0	22	1	71
NA000	0009	3	62	C	34	12	0	46	16	2	0	18	0	64
NA000	0010	3	62	C	36	15	0	51	14	0	0	14	0	65
NA000	0011	3	60	C	35	13	0	48	16	2	0	18	0	66
NA000	0012	3	63	C	35	14	0	49	18	2	0	20	1	70
NA000	0013	3	60	C	36	13	0	49	17	0	0	17	4	70

It was mentioned earlier that the third proficiency category of metacognition contributed very little or nothing at all to the total marks of the individual question papers. This could have been because the current assessment criteria did not include it and hence the question paper setting process does not accommodate it as a requirement in the current system. Therefore, the analysis in this section only included the learners' performance in the knowledge base and the strategy categories.

An initial analysis, in this regard, compared the mean marks of the learners' results in each one of the two categories in two groupings. The first grouping considered the results of the full sample of learners to investigate the first order effect. The second grouping analysed was according to the learners in each key grade, A, C, E and G to investigate the second order effect. Thereafter the mean mark analysis, a Pearson correlation analysis, followed by a regression analysis were performed with the above mentioned groupings in order to investigate similarities or differences in the performance of learners in these categories. The analysis in

this section was performed using the normalised marks. That is, the marks recorded were first expressed as a percentage of the maximum marks for that category. The marks were normalised in order to avoid the effect of different scales on the outcome of the analysis and thus ensuring a fairer comparison.

6.2.1 Analysis of the full sample: First order effects

This section presents the analysis of the full sample of the learners' results. The analysis was on the variation between the learners' original marks, evaluation marks and their two proficiency category marks. The results show an extensive variation in all four quantities, across all the learners. The variation for each learner across these dimensions was relatively small when compared with the variation across learners. As a result, there was a significant correlation between all these measures across the full sample of learners. This could be viewed as a first order effect, indicating coherence between all these measures for any learner when compared to the results of all the other learners.

The next two sub-sections present and discuss the coherence and variation between the two totals and the two proficiency categories within the full sample of learners, as indicated by the sample means, the sample correlation coefficients and the sample regressions.

6.2.1.1. The whole sample means: comparison per paper.

The performance of the learners in the original NSSCO examination system and the developed proficiency system showed some differences. As discussed earlier in section 5.3, the first step of this analysis was to normalise the learners' marks by changing them to percentages of the maximum for each category. This eliminated the effects of scale from the analysis. The normalised mean marks of the learners' totals from both systems are shown in Table 6.2.

Table 6.2. Mean marks for the two set of marks compared for both 2013 and 2014

	2013		2014	
	mean of the original mark	mean of the evaluation total	mean of the original mark	mean of the evaluation total
P1	54	58	60	60
P2	44	47	56	56
P3	50	44	46	44
P4	48	46	54	53

On the whole the mean marks of the evaluation marks were very similar to the mean marks of the original marks. The only noticeable difference detected was in Paper 3 of 2013. This is an indication that although the system was developed to distinguish between the two categories, their overall response outcomes were similar to each other and to the original marking system. This will ensure that the incorporation of this examination approach would show continuity with the functioning of the original examination system and the preservation of the reliability of the original examination outcomes

The mean percentages of learners' marks in the two selected categories were also calculated and compared. The whole group of learners seemed to clearly perform better in knowledge base than in the strategy category in all papers as shown in Table 6.3.

Table 6.3. Mean percentage marks for the two categories per paper

	Paper 1		Paper 2		Paper 3		Paper 4	
Whole sample year	mean KB	mean S	mean KB	mean S	mean KB	mean S	mean KB	mean S
2013	59	58	53	42	47	41	47	45
2014	63	56	61	52	47	43	59	47

The difference in mean marks for knowledge base and strategy categories for the whole group of learners had an average of about 6 percent. The differences between knowledge base and strategy categories are relatively small except those of Paper 2 of 2013 and Paper 4 of 2014. This could mean that there were papers that allowed learners to show a distinction in their performance in these two categories more than the others.

An additional observation was made when comparing the learners' performance in the two categories was that Paper 1 and Paper 3 showed less differences between the means of these categories than Papers 2 and 4. This highlights the distinction between core and extended levels of examinations. This could imply that either or both of the following: The extended papers bring out a better distinction between knowledge base and strategy categories as a result of the way the extended papers (Papers 2 and 4) were set and designed. Or, that the extended learners showed greater variance between categories. The difference in the learners' performance in the two proficiency categories were further investigated using correlation and regression analysis in the next section 6.2.1.2.

6.2.1.2. The full sample correlation and regression analysis

Pearson product-moment correlation coefficients for the full sample of learners per paper were calculated and analysed in this section. The analysis was done to examine the relationship between the learners' marks in the different categories for the full sample. The sets of learners' marks compared in this analysis were as follows; the learners' knowledge base (KB) and the strategy (S) marks, the learners' original marks (OM) and the learners' evaluation marks (EM). Finally, the learners' original marks and each of the two proficiency categories, knowledge base and strategy were also compared through the correlation test.

In addition to the correlation analysis, a simple linear regression was used to examine the strength of the relationship between knowledge base and strategy in terms of percentage variance explained (Higgins, 2005; Urdan, 2010). The result of both the correlation and the regression analyses is presented in Table 6.4.

Table 6.4. Full sample correlation analysis

Paper	Correlation coefficient KB vs S		Correlation coefficient OM vs EM		Correlation coefficient OM vs KB		Correlation coefficient OM vs S		Percentage significance KB vs S	
	2013	2014	2013	2014	2013	2014	2013	2014	2013	2014
1	0.994	0.948	0.991	0.991	0.988	0.983	0.976	0.972	100	100
2	0.935	0.893	0.983	0.979	0.969	0.940	0.964	0.963	100	100
3	0.890	0.848	0.979	0.990	0.970	0.984	0.926	0.895	100	100
4	0.930	0.967	0.984	0.983	0.956	0.978	0.973	0.973	100	100

The results in Table 6.4 show a very strong positive correlation between the performances of learners in all the different categories analysed. The correlation coefficients between knowledge base and strategy categories ranged from 0.994 to 0.890 for 2013 and from 0.967 to 0.848 for 2014 examinations. Thus, the two proficiency categories are strongly correlated in the full sample analysed, indicating that the two categories are not significantly different from each other in the global view. This was confirmed by the 100 percent chance of a significant relationship between the two categories in the whole group of all learners observed, as shown in Table 6.4. This showed that the hypothesis that the strong correlation observed resulted from a sample of an uncorrelated population can be rejected.

6.2.2 Analysis of key grade groups: Second order effects

Results in the two proficiency categories (knowledge base and strategy), were then compared for learners in groups having a similar performance, in order to investigate the possibility that

the distinction of these proficiency categories might be a local, second order effect. Comparisons were made within groups of learners in the same key grade. Results of these comparisons are presented in this section. Working within each key grade group provided means of controlling the global variation due to general performance that was identified in the previous section as having the major, first order influence in the learners' results.

6.2.2.1. Comparing the means of key grade groups

Initially the mean marks of each category were calculated and compared for each key grade group. These means are shown in Table 6.5. The results show no particular pattern in relation to the mean marks of the two categories across the key grades.

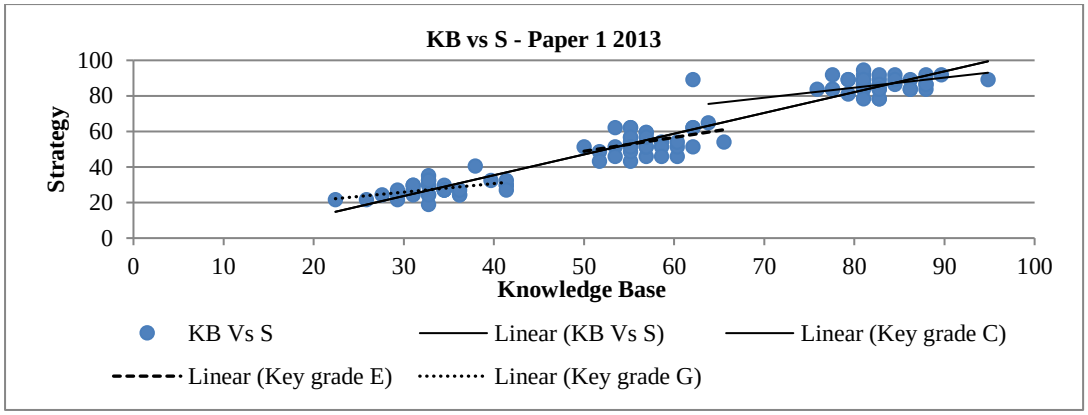
Table 6.5. Mean percentage marks for the year 2013.

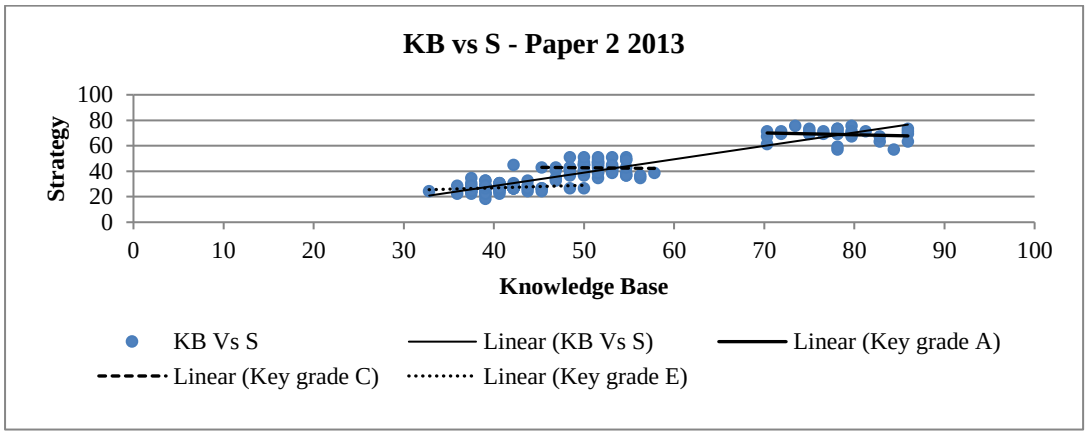
		Paper 1		Paper 2		Paper 3		Paper 4	
Year	Key Grade	mean KB	mean S	mean KB	mean S	mean KB	mean S	mean KB	mean S
2013	A			78	69			67	67
	C	82	86	51	43	63	58	50	45
	E	57	54	41	27	46	43	31	28
	G	34	28	------	------	32	24	------	-----
2014	A			72	68			81	68
	C	84	81	55	44	71	56	56	42
	E	63	56	44	28	44	47	37	26
	G	38	29	------	------	26	26	------	------

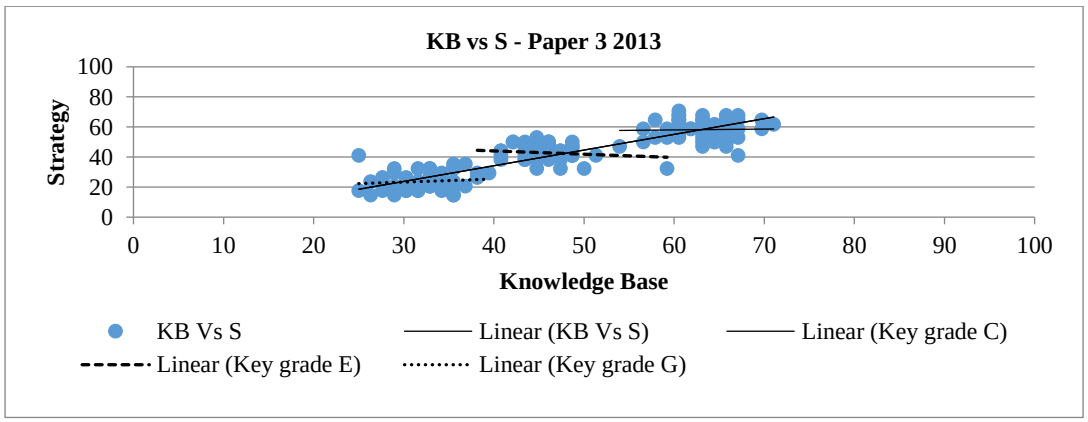
There was also a generally greater variation in the mean marks between the knowledge base and strategy category in the key grade groups as compared to the same variation in the whole sample shown in Table 6.5 above. This suggests an increased distinguishability of the two categories in the smaller key grade groups than in the whole sample of learners.

6.2.2.2. Correlation and regression analysis within key grade groups

In order to visualise the correlation of the learners' performance in knowledge base and strategy category, the learners' results in these categories were presented in scatterplots in Figures 6.1(a) and 6.1(b).







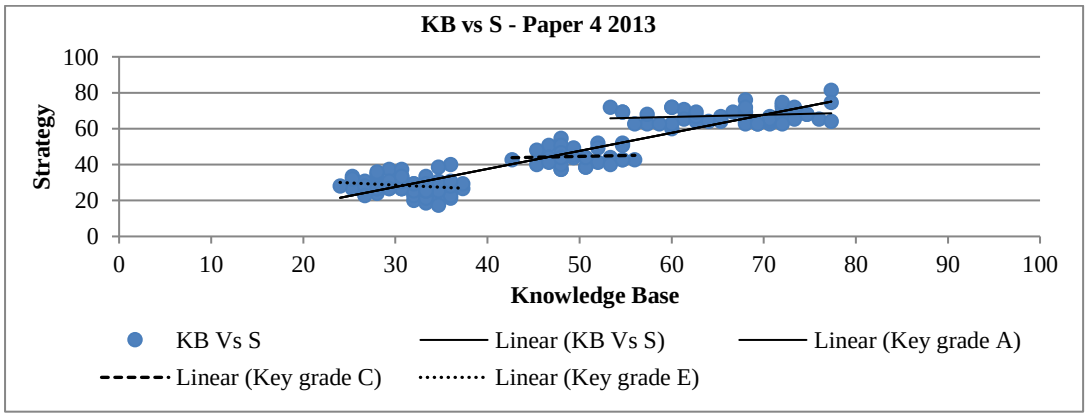
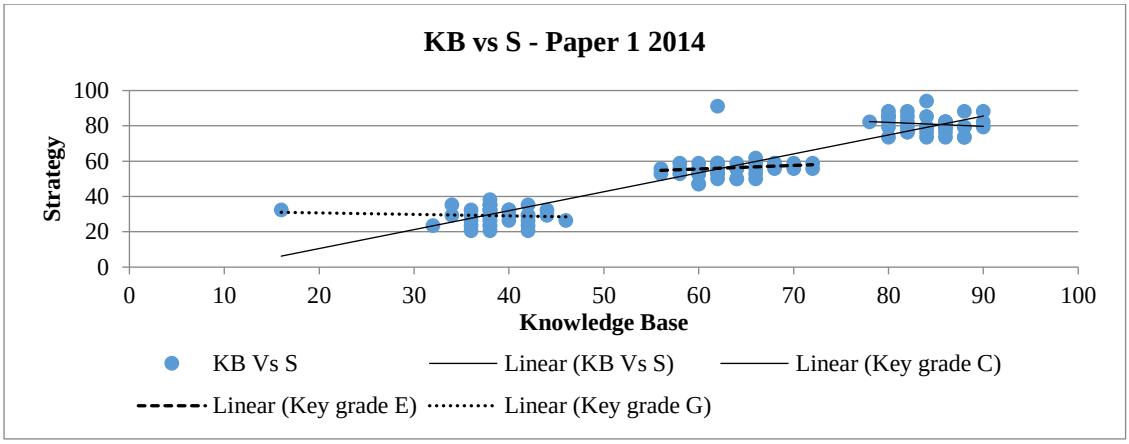
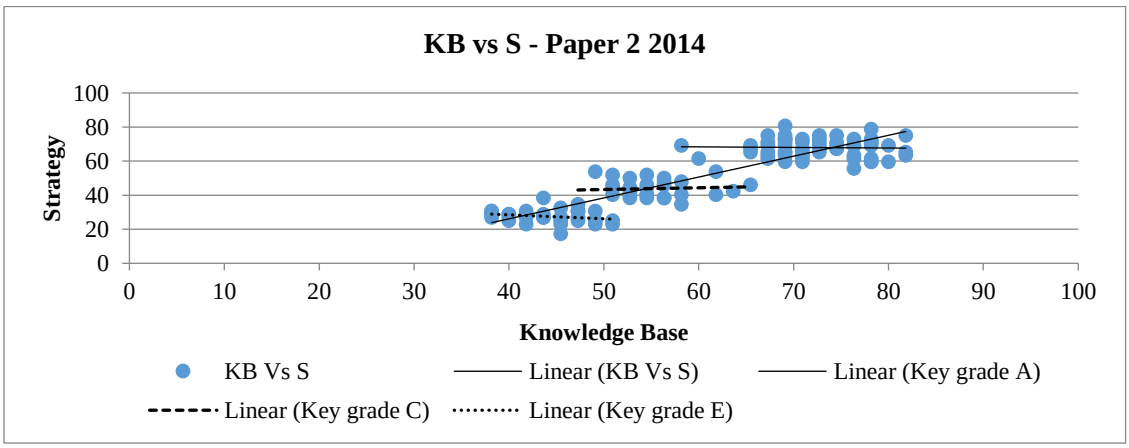
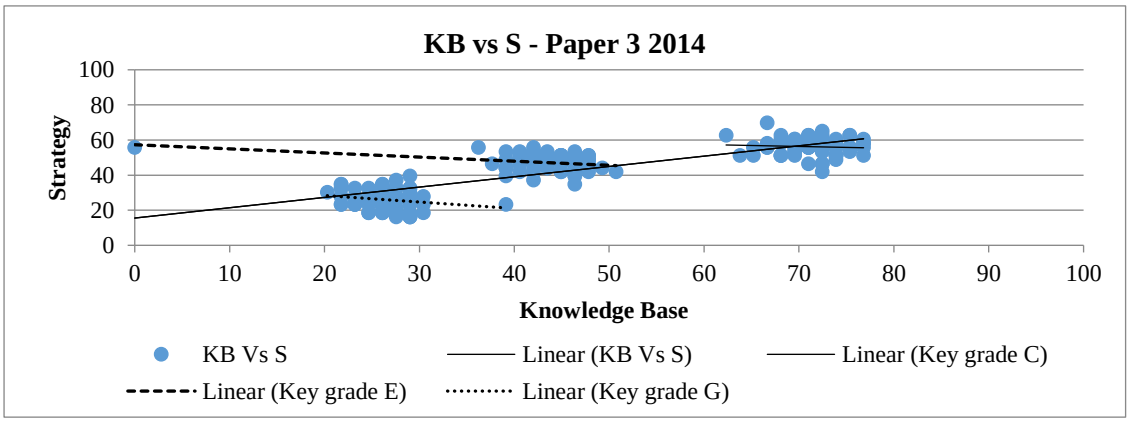


Figure 6.1(a). Scatter plots for all papers for the year 2013







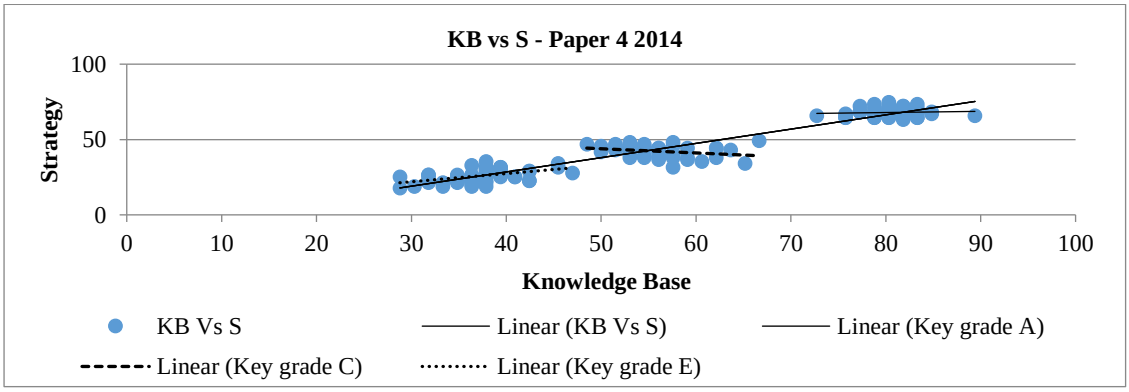


Figure 6.1(b). Scatter plots for all papers for the year 2014

Overall the scatter plots of all papers showed a very strong positive correlation at global level as shown by both Figures 6.1 (a) and (b). The data for the whole sample was correlated through the means of the key grade groups, since the clusters for each grade are quite compact and the line passes more or less through the centre of each key grade group. This means that on the whole, a good performance in knowledge base implies an equally good performance in the strategy category. However, when we take a closer look at the key grade groups, the scatter points are arranged in different ways, representing different correlations in the different groups. Each key grade group seems to have a different 'best' linear relationship and many of these clusters are quite 'round', indicating lower correlations within the key grade groups as compared to the global view. Examples of very round clusters are grade E of Paper 4 of 2013, grade C of Paper 1 of 2014 and grade G of Paper 3 of 2014.

The inter-relation between the learners' performance in knowledge base and strategy was also investigated through analysing correlation coefficients. Pearson-product correlation coefficients, r, and the regression percentage significance of the learners within the individual key grade groups were worked out in a similar way as was done for the full group of learners. Similar to the full sample analysis, the correlations were worked out between the following categories: the learners' knowledge base (KB) and the strategy (S) marks, the learners' original marks (OM) and the learners' evaluation marks (EM) and finally between the learners' original marks and each of the two categories, knowledge base and strategy. The correlation coefficient between knowledge base and strategy categories show the degree of scatter in each key grade group. The summaries of the correlation analysis and the simple linear regression analysis are presented in Table 6.6.

Table 6.6. Key grade groups correlation analysis.

P	Group	Correlation coefficient KB vs S		Correlation coefficient OM vs EM		Correlation coefficient OM vs KB		Correlation coefficient OM vs S		Percentage Significance KB vs S	
		2013	2014	2013	2014	2013	2014	2013	2014	2013	2014
1	Grade C	0.522	-0.150	0.865	0.559	0.799	0.353	0.702	0.379	99	65
	Grade E	0.324	0.122	0.598	0.595	0.455	0.546	0.532	0.390	92	57
	Grade G	0.493	-0.086	0.710	0.525	0.701	0.492	0.470	0.196	99	41
2	Grade A	-0.140	0.033	0.259	0.596	0.168	0.395	0.168	0.443	61	70
	Grade C	-0.380	0.656	-0.008	0.238	0.143	0.142	-0.008	0.203	20	54
	Grade E	0.159	-0.231	0.365	0.187	0.289	0.00	0.365	0.863	82	39
3	Grade C	0.030	-0.070	0.347	0.128	0.336	0.236	0.211	-0.066	18	44
	Grade E	-0.135	-0.236	0.233	0.294	0.241	0.496	0.047	-0.150	69	95
	Grade G	0.115	-0.196	0.603	0.377	0.626	0.294	0.324	0.195	62	90
4	Grade A	0.175	0.078	0.390	0.108	0.272	0.019	0.282	0.086	78	46
	Grade C	0.070	-0.287	0.382	0.228	0.279	0.051	0.290	0.180	30	95
	Grade E	-0.171	0.489	0.658	0.341	0.281	0.400	0.500	0.181	85	99

The data from the correlation analysis in Table 6.6 and the scatter plots in Figures 6.1 (a) and (b) was analysed and are presented in sub sections of separate question papers to highlight different observations at question paper level for the two years. Finally, the section discussed the implication the observation made has on the general design of question papers which provides further insight into the structures of the question papers.

(a) Paper 1

In Paper 1 of 2013, the learners in the key grade E group and the learners in the key grade C showed a very random distribution as compared to the learners in key grade G. This confirms the relatively weak correlation of the key grade E and C as compared to the key grade G group shown in Table 6.6. In 2014 however, the knowledge base and the strategy marks for learners in the key grade E seemed to be more strongly correlated to each other than the other two. In addition, the 2014 scatter diagrams in Figure 6.1 (b) showed a much more random distribution, the clusters are more round, in the key grade groups than that of 2013. This confirmed the two observations made earlier with the correlation and the regression, that the 2014 key grade group results have weaker a correlation than the 2013 results, as shown in Table 6.6. The above observation leads to the following conclusions: Firstly, the two proficiency categories could be more separable in the way learners performed in key grade groups than the combined group of all learners. Lastly, the degree of distinguishability was different from year to year and showed no particular trend.

(b) Paper 2

The results for Paper 2 of both 2013 and 2014 showed very weak correlation at all key grade groups except of the key grade group C which had a correlation coefficient of 0.656. There were some other notable observations that also suggested variation in the distinguishability of the two proficiency categories in the different groups. For example, in 2013 at both grades A and C, there was a negative and weak correlation, while at grade E there was a very weak positive correlation between the two proficiency categories' marks, see Table 6.6. However, Paper 2 of 2014 learners' results showed very weak correlations between the two categories in the two key grade groups A and E, but a stronger correlation in key grade C group, see Table 6.6. This once more gave an indication of greater variation between the two proficiency categories in the key grade groups. However, their variability showed no particular pattern.

Like Paper 1, there was also a general decrease in the level of significance as was shown by the outcomes of the regression analysis. The general decrease indicated a greater chance that the hypothesis that the population correlation is 0 could not be rejected. Thus, these results gave an indication that there was a greater chance that the two proficiency categories' marks were significantly different within the key grade groups. Additionally and in support of the above observation, the scatter plots in Figures 6.1(a) and 6.1(b) showed a relatively random distribution of the learners' marks within each key grade group.

(c) Paper 3

In both years, the correlation between the two proficiency categories was very weak within all key grades, with four out of six showing a weak negative correlation. Like in Papers 1 and 2 the very weak correlation indicated a greater chance of distinguishability of the two proficiency categories through the learners' performance within the key grade groups. However, Paper 3 seemed to show weaker correlation through all key grade groups than that Papers 1 and 2. Similarly, a drop in the percentage significance in the regression analysis also gave a greater indication that the outcomes were obtained from uncorrelated groups of data, see Table 6.6. Thus, indicating and confirming a greater chance distinguishability of the selected proficiency category in the individual key grade groups. The scatter plots of the individual key grade groups in Figures 6.1(a) and 6.1(b) also showed a very random distribution of points.

(d) Paper 4

Paper 4 showed a similar observation to the other three papers. The correlation coefficient between knowledge base and strategy were very weak in five out of the six key grade groups within the two years. Key grade E group of 2014, showed a weak correlation of 0.489 which was more collated than the other grade groups. Correspondingly the other the group had a 99 percent level of significance as was shown by the outcomes of the regression analysis. The scatter plots in Figures 6.1(a) and (b) also showed a distribution of points which is less random at grade E of 2014 than the other grades.

Finally, the results showed that the correlation coefficients of the learners' marks of the two different groups which were compared reduced considerably. That is, from the combined group of all learners to the group of learners according to the key grades, see Table 6.6. Firstly, in relation to the distinguishability of the marks relating to the different constructs, the weaker the correlation the stronger the chance to have the compared categories distinguishable. The comparisons between knowledge base and strategy showed correlation coefficients that were predominantly in the range of 0.656 to – 0.380. Thus ranging from a relatively strong positive correlation to a very weak negative correlation. The majority of the key grade groups ranged between a weak to a very weak positive correlation. This illustrated a very strong chance that the learners' performance in the two categories, knowledge base and strategy, could be separable within the key grade group as compared to the global observation made earlier.

Additionally, the data in Table 6.6 also showed that the comparison between the learners' original marks with their performance in the two selected categories had similar relationships. That is, the correlation between the original marks (OM) and the knowledge base marks (KB) is somewhat similar to the correlation between the original marks (OM) and the strategy (S) category marks. On average, these two comparisons both showed a weak to a very weak correlation between the compared constructs. Accordingly, this suggests that, in addition to the two categories having a strong chance of being separable, they also show a high chance of being distinctly different from the learners' original marks particularly in Papers 2, 3 and 4. However, the original marks for paper 1 seemed to be more strongly correlated to the learners' marks in the two categories than the other three papers, particularly in the year 2013. Further details on the variation between knowledge base and strategy are given in the next section which further investigates the distinguishability of knowledge base and strategy both at the global level and at the key grade group level.

6.2.3 Further analysis on the relationship between knowledge base and strategy in the learners' results

The analysis of the learners' results in key grade groups showed a relatively greater variation of their performance in the two separate proficiency categories, knowledge base and strategy categories compared with the global level observations. It was thus established that the learner to learner variations within each category (whether particular proficiency category, or total) were less correlated and this could be seen as related to a general measure of the learners' natural orientation in relation to the mathematical proficiency elements tested. However, the variability observed was not uniform and did not create a particular pattern. This section used the differences between the knowledge base and strategy marks to further investigate their variability.

6.2.3.1. The difference between knowledge base and strategy in the learners' performance

To investigate the variation between the two proficiency categories for learners showing similar general performance, the differences between the learners' performance in these categories were calculated as (the knowledge base mark – the strategy mark). Thus, a positive difference indicated that the learners' performance was better in knowledge base than in the strategy category, while a negative deference showed that the learners performed better in the strategy category. The results are presented as scatter plots shown in Figures 6.2(a) for the year 2013 and 6.2(b) for the year 2014. In the scatter plots the y-axis shows the differences and the x-axis represents the leaners according to their key grade groups.

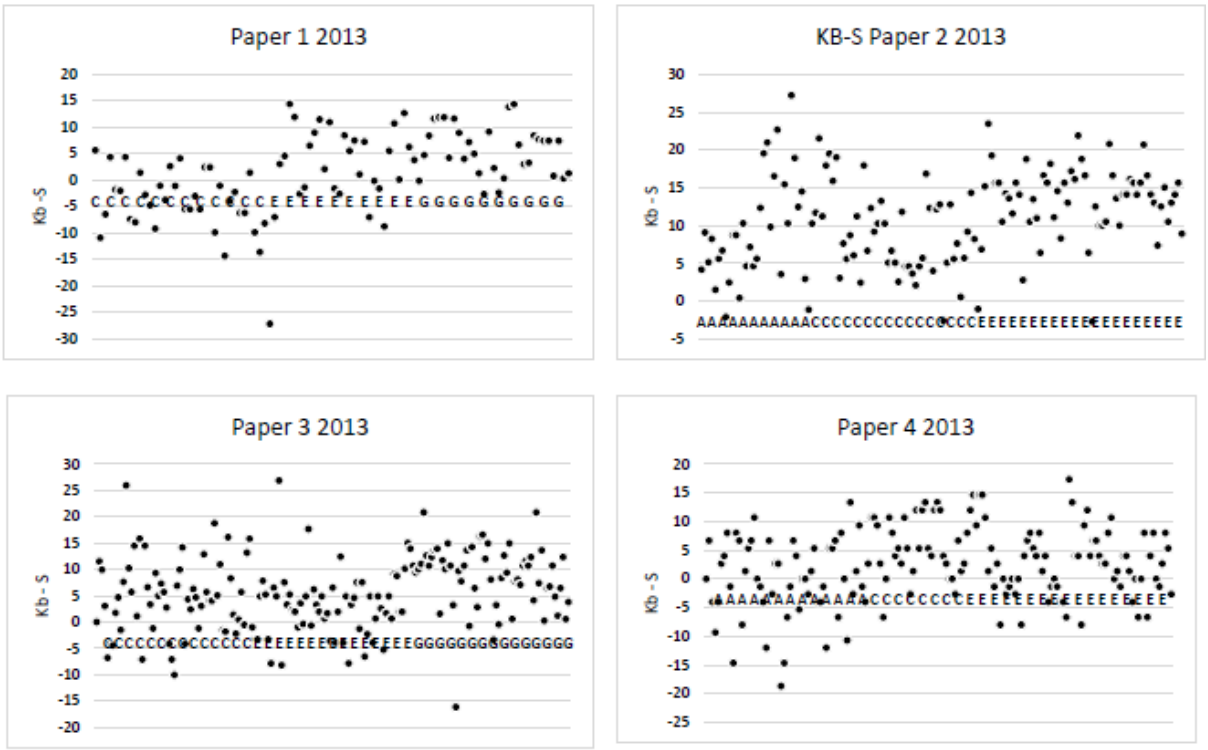


Figure 6.2(a). Scatter plots of the difference between knowledge base and strategy for 2013

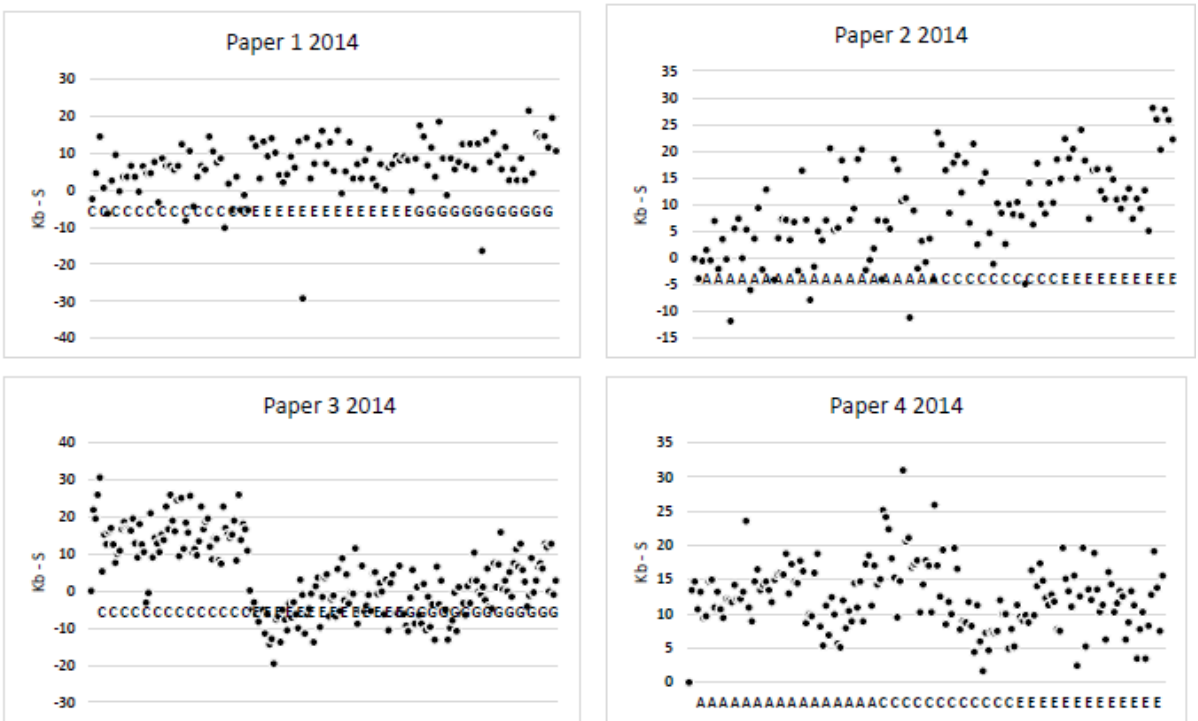


Figure 6.2(b). Scatter plots of the difference between knowledge base and strategy for 2014

The differences between the learners' performance in knowledge base and strategy category varied considerably within a particular paper, with a majority of learners performing better in

knowledge base than the strategy category across all the papers, see Figures 6.2(a) and (b). The scatter plots indicated that there was no clear jump between key grades group except for Paper 3 of 2014. This suggested a general pattern in terms of the distribution of differences. In addition, the difference seemed to be constrained within a certain range which is much less than the range of values across learners. The range within which 95 percent of the differences fall around the mean, is also given in Table 6.7. This implies that, assuming a normal distribution, the results in the two categories might vary independently of each category, but only within the observed range which is much smaller than the possible global variation.

One unexpected observation was that Papers 1 and 3 showed more negative differences than Papers 2 and 4. This means Papers 1 and 3 had more learners performing better in strategy that knowledge base compared with papers 2 and 4. For Paper 1 and 3 of 2013 the negative differences were mainly in the top performing learners. This observation was quite surprising for it showed that more learners at the core level (Papers 1 and 3) scored a greater proportion of strategy marks than knowledge base marks compared with the learners at extended level (Papers 2 and 4).

The data presented in Figures 6.2(a) and (b) can also be understood by working out the sample's mean of the differences and their standard deviation. This was done in order to understand how the differences were dispersed from their mean as their reference point (Urdan, 2010). The higher the standard deviation in relation to the mean, the greater the variation of the difference between knowledge base and strategy (Urdan, 2010). The results for 2013 and 2014 are shown in Table 6.7. Paper 3 shows the greatest variation, followed by Paper 1.

Table 6.7. Mean of the difference between knowledge base and strategy for 2013 and 2014

						Range within which 95% of observations fall			
		Mean of KB - S		Standard deviation of Kb-S		2013		2014	
Paper	groups	2013	2014	2013	2014	low limit	upper limit	low limit	upper limit
1	All	1	6	7.31	7	-13.62	15.62	-8	20
	Grade C	-4	3	5.37	5.61	-14.74	6.74	-8.22	14.22
	Grade E	3	7	5.46	6.61	-7.92	13.92	-6.22	20.22
	Grade G	6	10	4.33	5.49	-2.66	14.66	-0.98	20.98
2	All	11	9	6.02	9	-1.04	23.04	-9	27
	Grade A	9	11	6.33	7.38	-3.66	21.66	-3.76	25.76
	Grade C	9	16	6.47	7.46	-3.94	21.94	1.08	30.92
	Grade E	14	18	6.3	6.22	1.4	26.6	5.56	30.44
3	All	6	4	6.95	10	-7.9	19.9	-16	24
	Grade C	5	6	8.94	5.98	-12.88	22.88	-5.96	17.96
	Grade E	3	7	7.47	5.32	-11.94	17.94	-3.64	17.64
	Grade G	9	7	6.81	5.67	-4.62	22.62	-4.34	18.34
4	All	11	13	7	5	-3	25	3	23
	Grade A	0	13	4.68	5.61	-9.36	9.36	1.78	24.22
	Grade C	5	14	4.73	6.02	-4.46	14.46	1.96	26.04
	Grade E	3	11	4.4	4.7	-5.8	11.8	1.6	20.4

In general, the means and standard deviations of the differences shown in Table 6.7 just as the scatter graphs in Figures 6.2(a) and (b) showed less variation in the differences in Papers 2 and 4 and more variation in differences in Papers 1 and 3. This observation was made in both the full sample group of learners and in the key grade groups of learners.

On the whole, the mean differences between knowledge base and strategy was less than 10 in 5 out of the 8 papers and less than 15 in the other 3. The data also showed that on average a range of 7 standard deviations from the mean. This gives a good measure of the limits of the variation, which is much less than the variation in learners' general performance, suggesting that the difference between knowledge base and strategy made only a minor contribution to the difference between learners' variation in these papers. In addition, the two papers that show greater variation in the KB-S difference were papers with a higher proportion of level 2 strategy questions. That is, those questions that required learners to demonstrate more complex strategic proficiency. This suggested that the correlations and differences between knowledge base and strategy marks may be related to cognitive complexity. That is to say, there could be little

variation in low level strategy and knowledge base proficiency marks and a greater variation in higher level knowledge base and strategy marks. This observation brought out the need for further analysis. Therefore, the next level of analysis considered the observation of the distinguishability of knowledge base and strategy as observed through the learners' performance in these categories at different levels of complexity.

6.2.3.2. The learners' performance in these categories at different levels of complexity

This analysis considered a breakdown of the marks allocated to knowledge base and strategy into the two tested levels; level 1 and level 2. The analysis compared how the marks were distributed in the dominant levels 1 and 2 between the two proficiency categories. Moreover, the relationship between knowledge base and strategy category marks at different levels were also analysed. In order to understand the outcome of this analysis it was important to refer to the allocation of marks per level for each of the analysed papers as shown in Tables 6.4 (a) and (b) above. Particularly, it was important to note that the marks in these tables show an irregular distribution within the complexity level. Some levels were allocated more marks than others and some were allocated no marks at all. This posed a challenge in the comparison that followed as percentages of the total allocation were worked out for the marks achieved by the learners. However, the learners' results could still provide some insight into the distinguishability of the two categories according to complexity level performance.

The breaking down of the learners' marks into different complexity levels in each of the categories allowed for a better understanding of each of the proficiency category's different contribution to the learners' performance. For example, despite the fact that the scatter plots of KB Scaled per level and S Scaled per level for Paper 2, 2013 shown in Figure 6.3(a), showed a similar pattern, there were separate distinct features that each category showed. One of these features was that level 1 marks were more spread out in the strategy category than the knowledge base category. Secondly, the top performing learners seemed to be performing better in level 2 of knowledge base than in level 1.

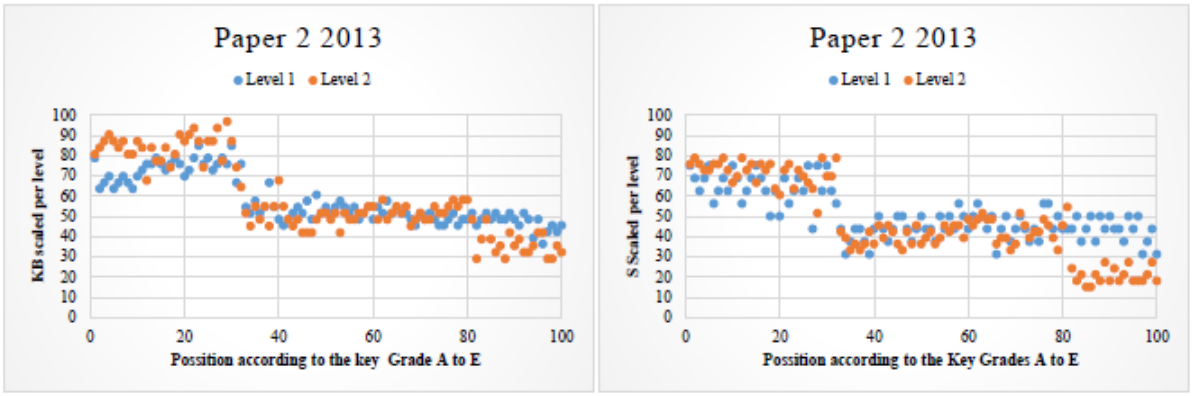


Figure 6.3(a). Scatterplots of knowledge base and strategy scaled per level for paper 2, 2013

Paper 2 of 2014 on the other hand showed a more noticeable and spread-out distribution of the level marks of both categories as shown in Figure 6.3(b). The difference in the learners' performance per level is more clearly shown. As in Paper 2 of 2013, learners performed better in level 2 of knowledge base than in the strategy category and they performed better in the complexity level 1. These two scenarios demonstrated the opportunities availed by assessing separate proficiency categories, which could thus allow for separate information to be extracted from the analysis of separate proficiency categories and further demonstrate their distinguishability.

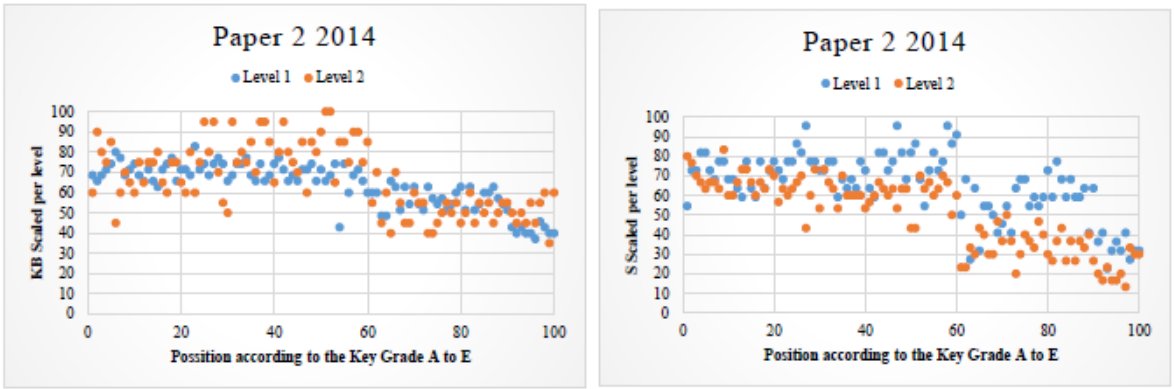


Figure 6.3(b). Scatterplots of knowledge base and strategy scaled per level for paper 2, 2014

Other features that reflected the distinguishability of the two categories are shown through analysing their relationship with regard to the learners' performance per complexity level. Figure 6.4(a) showed the learners' performance in knowledge base versus strategy category in levels 1 and 2 across the key grades. The two scatter plots showed distinct groupings that appeared to represent the learners' performance in particular key grades. Overall, there

appeared to be a linear relationship across the whole group of learners shown by how these groupings relate to each other through the means groups' (see the scatter plots of Kb versus S of Paper 2, 2014 in Figure 6.4(a)). The scatter plots of Kb versus S for Paper 4, 2013 and 2014 in Figure 6.4(b) showed similar features. However, the distinct formation in the grade by grade scatter plots showed some random distribution of points, details of these are given in the next section. This does not only demonstrate the distinguishability of the two proficiency categories, it also shows the distinguishability of assessing these categories at different complexity levels. Extending assessment to include separate components of mathematical proficiency at different complexity levels would enhance the quality of information obtained from an examination and would thus improve the examination feedback system to the classroom.

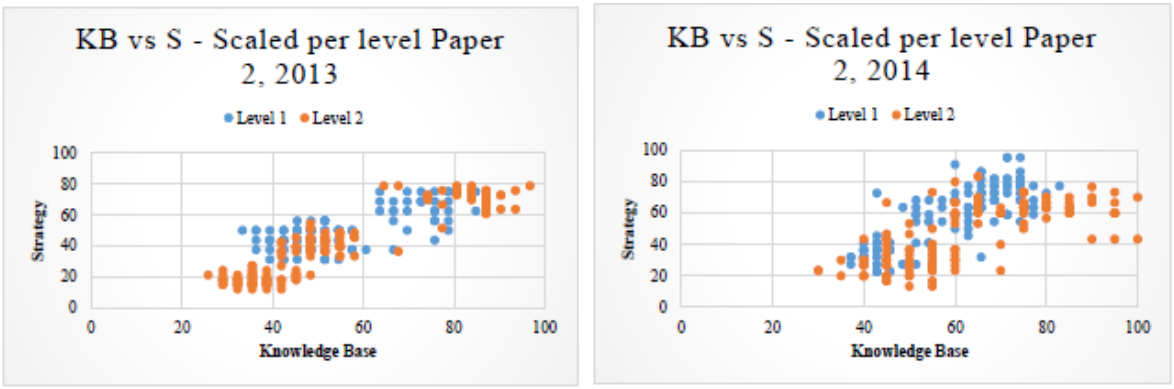


Figure 6.4(a). Composite scatterplots scaled per level for paper 2, 2013 and 2014

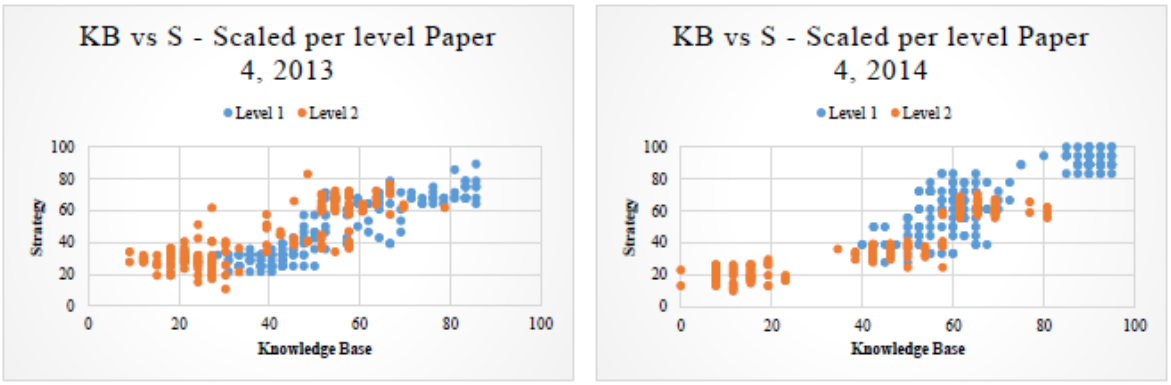


Figure 6.4(b). Composite scatterplots scaled per level for paper 4, 2013 and 2014.

The scatter graphs in Figures 6.4(a) and (b) showed two similar features that further highlight the variations in the learners' performance in relation to the two separate proficiency categories. In Figure 5.4(a), both graphs had the level 2 points more extending to the left of the graphs showing that there were a number of top performing learners who perfomed better in knowledge base than in the strategy category. While in Figure 6.4(b) both graphs showed that

the level 1 points seem to follow a regression line towards the top right corner, showing that in level 1 the learners seem to have performed more or less the same way in both categories.

When the same analysis was done at the key grade group level, more precise details were obtained. Figure 6.5 showed the scatterplots of knowledge base versus strategy for Paper 2 of 2013 at grades A, C and E. Grade A's scatterplot showed that learners seem to have performed more or less the same in both categories in level 1, but the majority of learners seem to have performed better in knowledge base at level 2 than in the strategy category. The middle grade C on the other hand showed not much difference in performance in both categories at both levels.

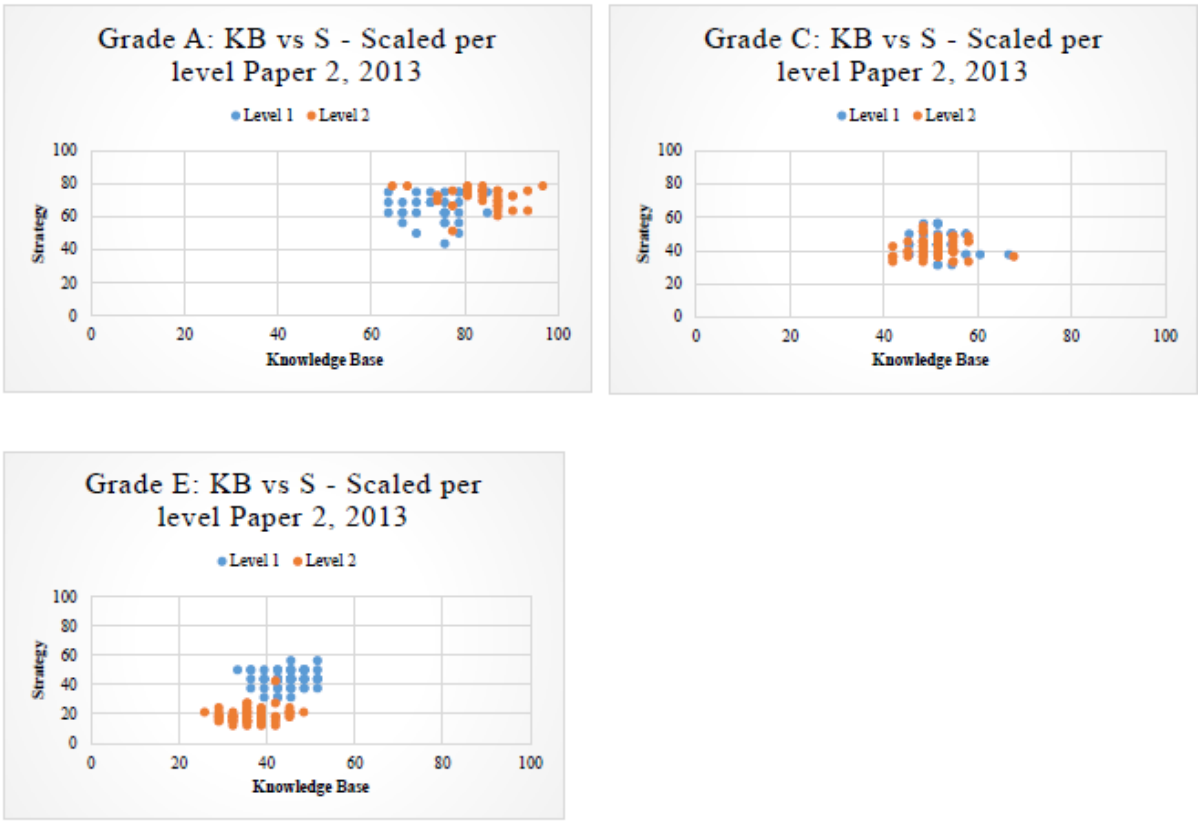


Figure 6.5. Key grade scatterplots scaled per level for paper 2, 2013.

These observations in these sections suggest three important considerations:

- The general mathematical engagement of the learners at the particular grade being assessed, might require some 'base level' alignment between the two categories of proficiency. Then:
    - the assessment of such a base level would require, and showed, very little variation between proficiency categories.

- a lack in any one category that reduces the learner's proficiency below such a base level, may override any strength in the other, so that the learners' overall performance will be effected.

- The questions could have been designed only to allow the visibility of a certain variability between the two categories. That is, although the difference in knowledge base and strategy requirements for different questions was sufficient to show the conceptual distinction in the marking structure, the questions do not allow as much distinction as could be measured in the general sense. In this regard, the issue could be a practical one and linked to the way these questions were set.
- The dominancy of complexity level 1 at the level of question papers design could have an influence on the observed distinguishability, in that the performance of learners in level 1 of knowledge base could have increased the variability between the two categories especially at key grade group level. Thus question papers needed to be adapted to enhance certain variability. The next section gives the general observations and the implication of the results on the question paper design.

6.2.4 General observations of the categories distinguishability in the learners' results

The data presented and discussed in sections 6.2.1, 6.2.2 and 6.2.3 brought out three general observations in relation to the question papers analysis and possible distinction of the two proficiency categories through the learners' results. These observation were;

(a) Broadly there were relatively small differences in the marks allocated to the two selected proficiency categories in most of the question papers when the two years were compared. This would suggest consistency in the mark distribution of these papers and could suggest a stable proficiency balance that could be used when designing the question papers. However, the suggested consistency was a bit weaker when a cross analysis of proficiency-complexity balance was done.

(b) There was a strong correlation between the learners' performance in the two categories within the full sample group of learners. In this regard, there was also a very strong correlation between the original and the evaluation marks within the complete sample. Hence, the distinguishability of knowledge base and strategy categories was less or not visible in the full sample group of all learners. Additionally, despite the fact that the developed system was designed to assess the selected components of mathematical proficiency separately, its outcome was not completely different from the outcomes of

the original system and preserved the original outcome of the NSSCO Mathematics examination.

(c) The correlation between the learners' knowledge base and strategy marks decreased considerably as the sample of learners was narrowed down to the learners in different key grade groups. The correlation in the key grade groups was generally weak in all four papers. Accordingly, this showed a greater visibility of the distinction of knowledge base and strategy categories within the key grade groups.

(d) The possible distinguishability of the two selected proficiency categories shown in the key grade groups of learners generally showed no particular inclination. However, any change in the degree of variability could give an indication of a difference in the test design. A relatively stronger correlation of 0.656 in key grade C of Paper 2 of 2014, as compared to other two grades, is a good example. In some papers the group average of learners showed a higher degree of distinguishability, showed through the weaker correlation than the other two key grade groups. Thus, for a particular paper a higher degree of distinguishability in any group could mean that the way in which the papers were designed discriminated learners better in that group than the learners in the other two groups.

The three observations had two different implications on the development of Mathematics education as follows:

(i) The interdependence of the proficiency categories appeared to preserve the overall learners' performance at the global level. However, the learners' abilities to do better in one category than the other observed in smaller grade groups could be useful in influencing the design of assessment tools that has the potential to enhance the discrimination in the learners' capabilities in the various proficiency categories. The latter has a potential to enhance the feedback strategies resulting from high stakes assessments which could be of greater benefit to the classroom teaching strategies.

(ii) The high chance of distinguishability observed in the key grade groups indicates that some learners might prefer knowledge base over strategy and vice versa as reflected by the greater variation in the average group of learners. This highlights the need for high stakes assessment strategies that will allow learners to take advantage of these fundamental differences. Mathematics assessment tasks should therefore be designed to contain an appropriate balance of the separate proficiency categories to allow the

takers to maximise their performance and reflect proper discrimination of their abilities. There is therefore a need to establish an appropriate mark distribution balance that allows all learners to have better access to questions with varying skills and abilities. In other words, the appropriate weighting for each category per complexity level needs to be determined which would allow for better discrimination amongst all groups of learners. In particular, there is a need to have for the question papers designed in such a way to provide better access to higher level knowledge base and higher level strategy in all papers. This is in order to better discriminate learners along the lines of their abilities in both categories.

6.3 Conclusion

In this chapter the data that was used to investigate the relationship of proficiency and maths competence was discussed. Particularly, the data that was used investigated the distinction of different proficiency elements in the specific marking schemes developed and in the learners' responses to the NSSCO examination. The constraints of the NSSCO Mathematics examination system in relation to the selected proficiency categories and the suggestions to mitigate these constraints are also presented and analysed in the next chapter 7.

CHAPTER 7: Research Findings and Analysis: Creating a better understanding of the current system

7.1. Introduction

This chapter presents and analyses the data collected during the last stage of the study. The data for the first two stages was presented and analysed in chapters 5 and 6 above. This data relates to the third research objective, which is: Using the designed tools to generate insight into the current examination system. This included the following objectives;

(a) To generate insight into the conceptualization of the current examinations, using the identified proficiency categories and developed tools.

(b) To investigate the constraints and affordances of the current examination system in relation to the identification and presentation of each of the identified proficiency categories, as well as the possibilities for overcoming these constraints by adapting question papers.

7.2. Insight into the conceptualization of the current examinations

In this section, the mark percentage distributions are presented and analysed in order to compare the structural balance of the question papers. Firstly, in terms of the proficiency categories, termed here the proficiency balance. Secondly, further analysis was done to examine how the proficiency category marks were distributed in complexity levels, this was referred to as complexity balance. These comparisons were done from year to year in order to generate insight into the conceptualization of the current examinations using the identified proficiency categories and developed tools.

7.2.1 Percentage distribution comparison: Proficiency Balance

This analysis examined how the percentages which were allocated per question paper, were distributed according to the proficiency categories. The allocation of marks to separate proficiency categories allowed the researcher to make inferences about the similarities and differences of the mark distributions of the separate categories across the different papers and the different years. This made it possible to gain insight into what proficiency categories appeared to have been originally tested by each paper of the NSSCO mathematical examination system. Table 7.1 shows the results of this comparison.

Table 7.1 Marks percentage per proficiency category for the four papers of 2013 and 2014.

Examination year	Paper	% KB	% S	% M
2013	1	60	38	2
	2	55	44	3
	3	67	31	3
	4	49	49	2
2014	1	59	41	0
	2	51	46	1
	3	59	40	0
	4	44	53	3

The knowledge base category contributed a higher percentage of marks for most of the question papers, particularly in Papers 1 and 3. For both years (2013 and 2014), these two papers had the higher percentage of knowledge base, at about 60 percent. Papers 2 and 4 on the other hand showed a slightly different proficiency balance. The knowledge base percentage for Paper 2 was slightly lower than that of the Papers 1 and 3. Moreover, Paper 4 for the year 2013 had an equal distribution of marks between knowledge base and strategy categories, while Paper 4 for 2014 had a lower percentage allocated to knowledge base (44 percent) than strategy (53 percent). By design, Papers 1 and 2 were expected to consist of questions which demand short answers: questions that mainly demand recall of concepts and less complex problem solving skills. Thus, it was not completely surprising that the two papers had the higher percentage of their marks assigned to knowledge base. Additionally, the questions of Papers 1 and 2 questions included some assessment that focused solely on knowledge base, with some questions only demanding knowledge of concepts.

Paper 3, a core level paper and Paper 4, an extended level paper, were designed to have more complex problem solving questions that expected more structured answers. In other words, Papers 3 and 4 were expected to have mostly questions that targeted more complex problem solving skills and no questions that addressed knowledge base alone. Thus, papers 3 and 4, were less likely to have the knowledge base category dominating in terms of the mark distribution. However, it was surprising to notice that Paper 3 for both years had the highest percentage of their marks assigned to the knowledge base category. Papers 3 for both years had the highest percentage of their marks (67 and 59 for 2013 and 2014 respectively) assigned to the knowledge base category. The big difference between the mark allocations for paper 3 for the two years also shows inconsistency in the setting of the paper across the years. The balance for Paper 4 seemed to reflect the expected observation. Paper 4 for the year 2014 shows a greater percentage of marks allocated to the strategy category than to the knowledge base category, while that of 2013 showed these categories in equal weighting (see Table 7.1).

Further details on the core and extended examination level and the expectation of the four papers are given in chapter 2 section 2.2.1.

Another valuable observation that could be made from Table 7.1 is that the proficiency balance for each paper was similar over the selected two year period. Very small differences were detected between the mark distributions in each paper from year to year. The differences between the knowledge base marks and the difference between strategy marks of the three out of the four papers ranged between 1 and 5 marks. The fourth paper (Paper 3) had a differences of 8 and 9 marks in knowledge base and strategy respectively. Such a similar mark distribution in a paper for both years could be a reflection of consistency in both the original design of the question papers and the application of the developed marking system. The slightly higher differences in Paper 3 could be a reflection of a lack of consistency in the setting of the two question papers across the two years.

Looking at this analysis in terms of the examination levels shows that the core level papers (1 and 3) together had more marks targeting the knowledge base category. While the extended papers (2 and 4) together, showed a more even balance, with knowledge base slightly higher than strategy in 2013 and slightly lower in 2014 –see Table 7.2. Overall, this suggests an insight into the expected proficiency capacity of learners at the different examination levels. It appears that core level learners were expected to show more knowledge base competencies than the strategy category.

Table 7.2 Percentages of scores testing different proficiency categories per examination level

		Knowledge base	Strategy	Metacognition
2013	Core (1+3)	64	34	2
	Extended (2+4)	51	46	3
2014	Core (1+3)	59	40	1
	Extended (2+4)	47	51	3

In conclusion, the relative consistency in the mark distribution between the two proficiency categories in each paper shows evidence that could suggest that the developed assessment system could be used as a useful tool to better understand the structural design of individual papers and the whole examination. Further insight was obtained from analysing the allocation of marks in these categories according to the complexity of the required response.

7.2.2 Percentage distribution comparison: Complexity Balance

The distribution of marks per complexity level was also compared amongst the four papers. To aid this second level comparison, two dimensional arrays were formed from the normalised marks for each paper. The first dimension considered in this analysis, was the allocation of marks according to the three proficiency categories; knowledge base, strategy and metacognition. The second dimension was the allocation of marks according to the three complexity levels.

As mentioned earlier the complexity scale was designed with three levels as follows (see Appendix G);

- Level 1, indicated proficient performance using elementary components of that category (atoms) – atomic proficiency. This was awarded 1 mark.
- Level 2, indicated the awareness and deliberate use of links (mathematical relationships or operations) between elementary components to enable the performance of these actions. This performance was awarded a maximum of 2 marks.
- Level 3, for which the questions required responses that involve the relating of mathematical links to form a coherent relational structure. This was awarded a maximum of 3 marks.

The analysis of the complexity balance was performed using two approaches. The first, was to compare the marks allocated to each proficiency category within a specific complexity level. That is, the analysis considered how the marks at each complexity level were distributed according to the proficiency categories. These distributions were then compared from paper to paper. Table 7.3 shows the percentage distribution of proficiency category marks per complexity level for each paper. Note that the metacognition category contributed very little to all the papers and all the marks assigned to metacognition were at level 1.

Table 7.3 Percentages distribution of marks per proficiency categories within a complexity level for each paper.

	Scaled (percentage) marks per proficiency category						
		knowledge base		strategy		metacognition	
Paper	Year	2013	2014	2013	2014	2013	2014
1	Level 1	59	88	78	88	100	100
	Level 2	41	12	22	12	0	0
	Total	100	100	100	100	100	100
2	Level 1	52	65	33	44	100	100
	Level 2	48	35	67	56	0	0
	Total	100	100	100	100	100	100
3	Level 1	62	80	88	79	100	100
	Level 2	38	20	12	21	0	0
	Total	100	100	100	100	100	100
4	Level 1	56	61	37	23	100	100
	Level 2	44	39	63	77	0	0
	Total	100	100	100	100	100	100

The distribution of these proficiency categories' marks into different complexity levels brought out a number of features that generated further insight into the design of the current examination question papers. Firstly, all the papers had a highest percent of their knowledge base category marks allocated to level 1 for both years. The distribution of the strategy marks showed a different pattern; in the core Papers 1 and 3 the highest percent of strategy marks were allocated to level 1. However, the extended Papers 2 and 4 had their highest percent of strategy marks allocated to level 2, see Table 7.3. The above shows that the distribution of marks amongst the different papers observed, reflected that the two examination levels, core and extended, were designed to have different expectations in relation to the two proficiency categories. The core Papers, 1 and 3 seem to demand less complex strategy skills and the extended papers 2 and 4 seem to demand more complex strategy skills.

In order to confirm the observation made in Table 7.3 a second approach of the mark distribution analysis was made by observing how the marks allocated to different complexity levels were allocated to the two predominant proficiency categories. For this to be done, another two dimensional array was developed as shown in Table 7.4.

Table 7.4 percentage distribution of marks per complexity level within each proficiency category for each paper.

Paper	Scaled (percentage) marks per complexity level				
		Level 1		Level 2	
	Year	2013	2014	2013	2014
1	Knowledge base	52	59	75	60
	Strategy	45	40	25	40
	Metacognition	3	1	0	0
	Total	100	100	100	100
2	Knowledge base	62	60	48	40
	Strategy	30	39	52	60
	Metacognition	8	2	0	0
	Total	100	100	100	100
3	Knowledge base	54	59	88	58
	Strategy	46	40	12	42
	Metacognition	0	1	0	0
	Total	100	100	100	100
4	Knowledge base	58	65	41	30
	Strategy	38	29	59	70
	Metacognition	4	6	0	0
	Total	100	91	100	100

Consistent with the observations made using the data presented in Table 7.3 above, the complexity level distribution of marks shown in Table 7.4 allows deeper comparison of the marks allocated to the knowledge base category and strategy category within the four papers. All four papers had the highest percent of their level 1 marks allocated to the knowledge base category for both years. In addition, the highest percent of level 2 marks for Papers 1 and 3 were also allocated to knowledge base for both years. The extended papers, 2 and 4 on the other hand, showed a different pattern. They all had the highest percent of their level 1 marks allocated to knowledge base and the highest percentage of their level 2 marks allocated to the strategy category. Despite the fact that these observations suggest a general consistency on how the level marks were distributed amongst the categories between the two years, there were differences in the actual percentages of marks distributed within levels. For example, 75 percent of level 2 marks for Paper 1 in 2013 were allocated to the knowledge base category but only 60 percent of level 2 marks for Paper 1 in 2014 were allocated to the knowledge base. This shows a difference of 15 percent.

Paper 3, unexpectedly and consistent with the earlier observations, had the highest percentage of both their level 1 and 2 marks allocated to knowledge base consistently for both years. Paper

3 and Paper 4, were expected to have the percentage of their marks at both levels allocated in more or less the same way for both knowledge base and the strategy category. This was because the two papers 3 and 4 were designed to have questions that demands structured answers and target more complex problem-solving questions. The two papers were not expected to have simple knowledge recall questions whose marks could be allocated to knowledge base only. In other words, Paper 3 was expected to have the marks targeted for the strategy category and those targeted for the knowledge base category at both levels to be more or less the same. Knowledge base was not expected to dominate at any complexity level. The unexpected observation confirms similar trends for the core level Papers 1 and 3 made earlier in that more marks were allocated to knowledge base in these papers.

The above observations provide insight that could be valuable to the design and development of the question papers. Moreover, the analysis of the proficiency and the complexity distribution of marks allows inferences to be made using the proficiency-complexity profiles of each paper within the two years of examinations analysed.

7.3. Constraints of the current examination system and the possibilities of overcoming them

This section presents an analysis of the markers' qualitative comments and judgments of the two assessment systems and suggests possible adaptations to the existing NSSCO Mathematics question papers and marking schemes in order to enhance the distinguishability of the two proficiency categories. This analysis was based on the data gathered through the discussion of the focus group before and after the evaluation phase. The summary of the markers' comments is attached as Appendices I and J. The data collected for this section was presented and analysed in order to address issues relating to the fifth research objective. This objective was to investigate the constraints of the current examination system in relation to the identification and presentation of each of the identified proficiency categories, as well as the possibilities for overcoming these constraints by adapting question papers.

The data was presented and analysed in relation to:

(a) Constraints identified through markers' experience.

(b) Possibilities for overcoming these constraints.

7.3.1 Constraints identified through markers' experience

This subsection provides the markers' understanding of the possibility of assessing the two selected proficiency categories separately. The observations the markers made in this regard were informed by their experience when setting up the marking schemes and when remarking the learners' responses. In this light, the markers provided a critical analysis of the design of the NSSCO examination in terms of paper construction and balance.

In general, the markers' observed that the current NSSCO Mathematics examinations indirectly demand that learners give correct answers only. The proficiency system however, gives credit for various separate skills and competencies displayed within the learners' working to get a response (Ma1, Ma2 & Ma6). Marker Ma1 gave an example of question 9(c) in Paper 4 for 2014 examination, where the learners were previously only awarded one mark for identifying the regions by shading. However, through the new system they were awarded two separate marks; one for the knowledge of showing that the region satisfies the inequality and the other mark for the strategy employed in shading regions were also awarded (Ma1). Table 7.5 shows an extract of the specific question paper rubric of paper 4. 2014 question 9.

Table 7.5 question 9, paper 4 2014.

Q.no	Question	Question response	Response Complexity Score (Score: 0 or 1 or 2 or 3)			
			KB	S	M	**Total**
9	**(a grid is provided)** A dog breeder has *x* Dobermans and *y* Bulldogs. The breeder has 7 or fewer Dobermans and no more than 8 Bulldogs. She has enough room in her kennels for not more than 10 dogs in total at a time.					
(a)	Express the three conditions above as inequalities.	$x \leq 7$; $y \leq 8$; $x + y \leq 10$ *KB atoms: Linear programming* *S: Implement an atomic strategy*	2	1	0	**3**
(b)	Using 1 cm = 2 dogs on the *x*-axis and 2 cm = 1 dog on the *y*-axis draw the three inequalities on the axes.	Scale of *x*-axis Scale of *y*-axis Correct ruled lines *KB atoms: Linear programming: drawing in equalities- scale, drawing a line* *S: Implement a composite strategy + length*	2	2+1	0	**5**
(c)	Shade the regions which do not satisfy the inequalities, leaving the region which satisfies the inequalities unshaded.	Correct shading *KB atoms: shading a region :Linear programming* *S: Implement an atomic strategy*	1	1	0	**2**

In another example, Marker Ma7 stated that the current NSSCO system of marking does not award marks for most of the knowledge required by a question, but the proficiency system does under the category of knowledge base. Example for the question: 31% of a class 45 learners passed a test. How many learners passed the test? The expected answer with working is $\frac{31}{100} \times 45$ =14. This was worth one mark in the original marking, but would now be worth 2

marks, one mark for an indication of the atomic knowledge and the other for applying a strategy and obtaining the answer.

Most of the questions in the NSSCO did not directly require learners to show their working or calculations. Thus, the calculation steps required for the learner to give their solution was not always required by the questions. Hence, in line with the strategy as proficiency category, the only mark awarded in cases where the calculation was not shown was level 1. In addition, the original mark allocation did not indicate that learners were required to do more, by showing separate skills when working out their solution (Ma2 and Ma6). This resulted in many learners losing their strategy category marks which they could have been awarded if the working had been shown. This was a particular concern for level 2 strategy category responses where the learners' final answers were wrong and they have not shown working. If a learner showed his or her correct working he or she could have been awarded a separate strategy and or knowledge base mark, depending on the work shown, even if his or her final answer was wrong (Ma6). The above observation affected mostly the learners in the lower grades E and G who appeared to have been the ones giving only the final answers without showing their working (Ma4 & Ma5).

Furthermore, the NSSCO question papers seemed not to have tested higher level knowledge base and strategic skills. According to Marker Ma4 the new system would help to discourage mechanical calculator work and to give credit to alternative thinking and innovation. The markers also observed that the question papers tested little or no metacognitive competencies (Ma4, Ma5 & Ma7). There was no description and explanation asked and the only cases which could be awarded a metacognition mark were where learners chose one of the available strategies that could be used to solve a given mathematical problem (Ma4, Ma7). Figure 7.1 gives Marker Ma7's comment regarding the above. In the comment the marker also gave the suggestion of a question that could better assess metacognition. He also emphasised that demanding an explanation from the learners might make their understanding of concepts clearer, which could allow the separate assessment of argumentation clearer.

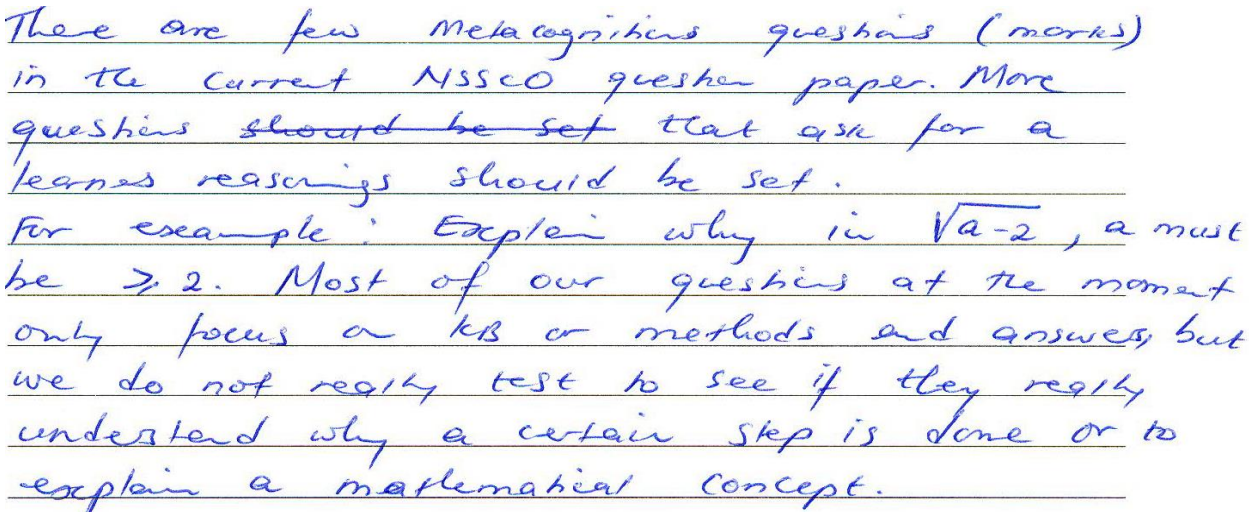


Figure 7.1. Marker Ma7's comment on testing high level competencies and skills

Despite the relative ease in the way the markers allocated marks to separate categories when developing the specific rubrics, Markers Ma3 and Ma4 felt that it was improper to award a mark for a choice of a strategy only. Marker Ma3 stated "Metacognition mark should not be awarded for just a choice of a method, every question might have two or three methods to solve it". He further suggested that awarding marks for the choice of strategy would in his view only be appropriate "if the learner uses a strategy that has an advantage over another method". However, further discussions in the group established that Ma3's suggestion above might challenge the reliability of the metacognition component, particularly through the marking.

Some markers further observed the assessed complexity of the knowledge base and strategy categories and said the awarding knowledge base and some strategy marks was limited to only 1 mark or 2 marks per question (Ma3, Ma4, Ma5). They then called for the assessment of these categories at varied complexity levels, in particular, Markers Ma4 and Ma6 suggested that learners should be asked to justify their answers and given explanations which lead the discussion into the next session that gave suggestions for adaptation of the tools.

7.3.2 Possibilities for overcoming these constraints

This section presents the suggestions made by the markers on the adaptation of the NSSCO Mathematics examination system to ensure enhance visibility of mathematical proficiency. The suggestions on adaptation given were targeted at ensuring clearer visibility of the variation of the selected categories through the learners' performance. The proposed changes are directed at the adaptation of;

(a) the general question papers' construct, by suggesting a different possible composition of the question papers that would allow for enhanced visibility of the separate categories.

(b) the particular question's construct, by suggesting a difference in question design in order to ensure that the learners make the separate proficiency categories more visible through their responses to these questions.

7.3.2.1. Question papers' construct

The question papers should be designed such that a specific number of marks are assigned to the separate competencies; knowledge base, strategy and metacognition in a predetermined proportion (Ma1, Ma2, Ma6 and Ma7). Marker Ma1 was concerned that there were no marks that could be allocated for explanations or descriptions of concepts and strategies in the current papers. He was also concerned that the marks allocated to the metacognition category were very low. He went on to say this needs to be increased to allow proper assessment through the proficiency system (Ma1). Thus, a specific number of marks should be made available during the setting for a specific competency, with clear instructions on how such features could be made visible by the learner when answering the question (Ma1 & Ma2). Papers 3 and 4 should be made up of questions or sub questions which demand the learners to show more of their strategy and should then have more than 1 or 2 marks (Ma1, Ma5 & Ma7). Additionally, the markers suggested these question papers to have more marks assigned to metacognition since these question papers are set for more complex questions that demand structured answers. Consequently, the question papers designed on the proficiency assessment system might need more marks and more time for the learners to respond compared with to the original NSSCO assessment system (Ma3 & Ma4).

7.3.2.2. Question construct

Markers suggested that questions should be constructed in such a way that the competency required is apparent and should clearly show how each competency is going to be credited (Ma6 and Ma7). The markers felt, since the question papers used for the evaluation process, were not set on the developed system this was not always the case in some questions. They stated that, although it was relatively clear what the markers were to look out for to award a mark (Ma6, Ma7), it might not have been clear to the learners which competencies were required in line with the new proficiency system (Ma1, Ma2, Ma3 & Ma4). Ma4 stated that when a learner gives a correct answer despite them having not shown their strategy they are

awarded 1 mark for knowledge base and another mark for strategy. He gave the example of question 8(b) Paper 4. Question 8 (b) required learners to solve for x in an equation $\log_3^{x-1} = \log 11$, which $x = 3.18$ a solution as the solution. On the other hand, a second learner who got a wrong answer, while not having showed his/her working, would not be awarded any of the two available marks. If such question directly demanded the learners to show their working it would have given the second learner a better chance to be awarded 1 mark out of the two available marks (Ma1 & Ma5).

The markers also stated that questions which demand strategies should clearly state that the learners are required to show their working with the marks corresponding to the separate competencies required (Ma1). The example in Figure 7.2 illustrates this point.

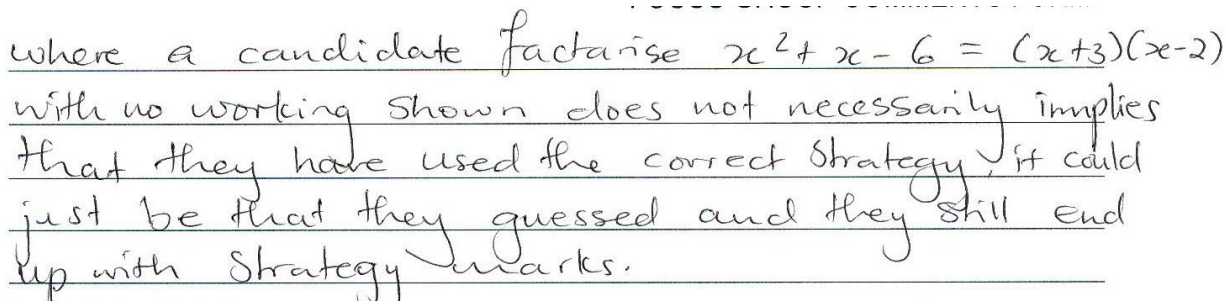


Figure 7.2. Marker Ma1's comment on the demands of the questions

In another example, Marker Ma7 was concerned about how the old NSSCO Mathematics marking penalises accuracy errors. In her example to calculate N$30 − N$6.40, the learners were required to give the answers to two decimal places for 1 mark. A learner who presents his or her answer as N$23.6 was not awarded a mark since he/she used only one decimal place in his/her answer. She said that, this was not proper since N$23.60 and N23.6 have the same value. Marker Ma7 was of the opinion that there should have been two separate marks available for this question, 1 mark for strategy and 1 mark based on the knowledge of accuracy. Therefore, each question should be designed to clearly be able to objectively award marks for all the competencies demanded i.e both knowledge base, strategy and metacognition should and could be made clearly visible to be awarded a separate mark (Ma1, Ma3 & Ma7).

The markers gave several examples of how the current questions could be adapted to fairly assess the competencies in the new developed system. These examples are;

- Question 11(a), Paper 4 of 2013, would have been better if learners had been asked to convert ¼ litre to millilitres so that knowledge base can be shown by the learners instead

of asking them to choose which cup contains more as some learners might guess their answer (Ma4).

- Question 15 on comparison, would have been better if the learners had been asked to work out the volumes first before comparing the amount of liquids in the cups so that the strategy mark could easily be separable and fairly awarded (Ma1).

On Question 17, Paper 1 of 2013 it was difficult to identify the knowledge base since the bearing asked involves working with angles. The expected answer was already a three digit number and some learners were writing any three digit numbers. It would be better if the answers could have only two digits and learners who added a zero in front of the two digits, could be the ones who clearly show their knowledge of bearing and can then be judged easily and fairly (Ma6).

7.4. Conclusion

In closing, this chapter presented and analysed the information that was aimed at generating insight into the conceptualisation of the current examinations using the identified proficiency categories and the associated tools. The chapter finally presented the constraints of the current examination system in relation to the identification and presentation of each of the identified proficiency categories, as well as the possibilities for overcoming these constraints by adapting question papers as highlighted by the focus group. The next chapter, chapter 8 will discuss the research findings in order to shed light on the issues raised by the five research objectives.

CHAPTER 8: Discussion of the Research Findings

8.1 Introduction

This chapter discusses the research findings generated from the analysis in chapters 5, 6 and 7. These findings were discussed in relation to the research objectives and the relevant literature. The discussion was broken down into several sections in order to address the following matters.

(a) The feasibility of deepening Mathematics assessment by adapting the examination tools in order to assess mathematical proficiency through high stakes assessment more effectively. The issues discussed in this section are raised by the research objectives 1(a) and (b) as shown in sections 1.7 and 4.2.

(b) The investigation of the relationship between general Mathematics capability and mathematical proficiency as evident in the developed tools and the learners' answers' scripts. The discussion here strived to develop deeper insight into the distinctiveness and relatedness of aspects of mathematical proficiency. The issues discussed in this section were raised by the research objectives 1(c) and (2) as shown in sections 1.7 and 4.2.

(c) Finally, the chapter discusses the insights into the current system, emerging from the use of the proficiency tools as related to the theories guiding this study. This discussion focused on two fundamental issues of assessment which were important to this project. These fundamental issues were:

- The possibility of using proficiency as the basis for a high stakes assessment framework.
- The concern of assessing what is easy, rather than what is important.

The discussion here used the identified constraints of the current examination and the possibilities for overcoming these constraints to discuss the issues above. Thus the discussion was guided by the research objective 3(b) as shown in sections 1.7 and 4.2.

8.2 The feasibility of deepening Mathematics high stakes assessment

The possibility of assessing distinct mathematical proficiency strategies explored and revealed through the findings of this study has deepened the Mathematics assessment currently assessed by the NSSCO Mathematics examination. The current NSSCO Mathematics examination scores represented learners' competencies in particular processes that were not formally categorised to allow a better understanding of what proficiencies are assessed and how best to assess all important aspects of Mathematics (Burkhardt, 2007). Therefore, the result discussed here shows that the adaptations that were done to the NSSCO examination tools achieved two goals;

- They described competencies that could be categorised into distinguishable mathematical proficiency categories.
- They allowed the characterisation of what the current NSSCO Mathematics examination is assessing in relation to the selected proficiency aspects. This extends the assessment outcomes in that they not only award uncharacterised scores per question, but allow assessment results to describe learners' competencies with regard to the distinguishable proficiency aspects.

The above was achieved through the adaptation of NSSCO Mathematics examination artefacts in order to deepen high stakes Mathematics assessment. The NSSCO Mathematics examination tools; the broad teaching, learning and assessment standards or the process standards and the marking schemes were adapted to incorporate the distinct proficiency elements and make proficiency more visible. In addition, the changes made to the artefacts were seen to better enable learners to make their mathematical proficiency visible through examinations. The adaptations made were based on Allan Schoenfeld's (2007a&b) proficiency formulation and incorporated three of his proficiency categories: knowledge base, strategy and metacognition. In addition, more general literature on what defines a mathematically proficient learner was explored and informed the extraction of traits that characterised the three categories. Additionally, competency in argumentation from other formulations of mathematical proficiency, was incorporated into the selected categories by extending them to demand that learners reflect their understating through explanation and reasoning. The research findings are discussed here in two distinct ways;

- Understanding the distinct aspects of Mathematics assessed by the NSSCO examination.
- Extending what Mathematics is assessed by the NSSCO Mathematics examination.

8.2.1 Understanding the distinct aspects of Mathematics assessed by the NSSCO examination

The data obtained through the literature search was the first indicators of how distinct the selected proficiency categories were from each other. The authors of the literature reviewed could distinguishably provide several descriptive competencies that were aligned with the selected categories. These descriptions could be used in the development of prototypical responses to the specific question paper that distinguished features that could be associated to separate proficiency categories. This led to the development of the general proficiency rubric

and hence the adaptation of the marking scheme to develop the specific question paper rubrics with distinct question response features awarded marks separately. The specific question paper rubrics were designed in such a way that the expected mathematical proficiency aspects, e.g. strategies used to get to the answers, are all written down. This further facilitated the characterization of the structure of different papers used in the evaluation process, a practice needed for consistency in assessment from one year to another as a way of maintaining standards. Knowing how the marks were distributed between the categories within the current examination would enable a process of working out an appropriate structure for future examinations.

The distinguishability of the distribution of marks for the two selected categories, knowledge base and strategy proved to be statistically significant within the four papers as tested using chi-square tests. These tests which were run under a null hypothesis (H0) that the probabilities of the distribution of marks of the two categories were the same in each paper. The significance of this distinction was indicated by the fact that the values of test statistics were all larger than the expected test statistics at the 95% significant level. These values lead to the rejection of the null hypothesis for all papers as recommended by (Hogg et al., 2013). Therefore, there was a statistical indication that there was a significant distinguishable difference between the distribution of knowledge base and strategy marks in all four different papers for both 2013 and 2014.

The resulting distinguishable mark distributions also revealed that the NSSCO Mathematics examination question papers were limited to testing mainly knowledge base and strategy categories of mathematical proficiency at their lower level of complexity. The more complex levels of these proficiency categories were barely tested. The metacognition category, which demands the learner's reflection, self-evaluation and self-monitoring process during problem solving seemed not to be tested at all. The above observation corresponds with Marker Ma2's comment that, the manner in which the NSSCO question papers and their corresponding marking schemes were constructed were found to give more emphasis to correct answers. The above was particularly important for this study in order to identify the strength and weakness of the NSSCO Mathematics examination and find ways of deepening it.

Assessing distinguishable elements of mathematical proficiency would strengthen the NSSCO examination system in that, it would allow identification of possible shortcomings of question

used. For example, demanding a correct answer only does not allow the learners to show their fluency in mathematical strategies (Ma2). The observed emphasis of correct answer only was particularly the case with the three papers, 1, 2 and 3. All steps taken to work out the answer, needed to be clearly shown and assessed through the learners' responses in order to make their proficiency visible as suggested by (Bazerman, 2012). In addition, the learners' thinking through solving a question needs to be made more visible, in order to ensure that Mathematics examinations would not continue promoting rote learning (Burkhardt, 2007). The above considerations highlighted some of the shortcomings of the current NSSCO Mathematics examination system in relation to assessing proficiency and confirmed the identified gaps between the two systems that were filled during the process standard adaptation procedure. Accordingly, the weakness identified indicated what aspects of proficiency future assessment should incorporate in order to encourage teaching and learning for the proficiency traits not tested but important to teach and assess as promoted by (Burkhardt, 2007). Thus, understanding distinct aspects of the Mathematics assessed through NSSCO leads to understanding how it can be extended.

8.2.2 Extending what Mathematics is assessed by the NSSCO Mathematics examination

The deepened assessment process addressed by this study strived to ensure valid assessment of all aspects of mathematical proficiency in order to encourage teaching for proficiency (Burkhardt, 2007; Schoenfeld, 2007 a). The specific question paper rubrics reflected improved visibility of distinguishable mathematical proficiency aspects tested. In the context of this study, the distinguishability of these aspects was particularly important because, in order to enable valid assessment of these aspects through high stakes examinations, they needed to be distinguishable in relation to the outputs produced. As mentioned earlier, due the continuous operational nature of the NSSCO examinations, the adaptation was limited to the process standards and the marking schemes, but not the question papers. Therefore, despite the fact that the incorporated proficiency elements were distinguishable and brought a different view to the learners' results, it did not allow the extension of Mathematics assessment system to assess all competencies desired within the selected elements of mathematical proficiency. Incorporating all identified competencies of proficiency would require adaptation of the question papers as well.

All the mathematical competencies and skills expected to be seen in the learners' responses to an assessment task need to be clearly demanded by the questions (MoE, 2010 c). For example,

the literature studied expects that learners should be able to describe the strategy to be used rather than only providing a mechanical arithmetic solution as was currently observed for most questions in the NSSCO Mathematics examination. However, for this to be validly assessed it should be clearly stated in the process standards linked to the content standards and also clearly specified within questions in a particular question paper. Incorporating such a change which would require the adaptation of question papers that leads to a major systematic change would require further study.

8.3 The relationship between Mathematics capability and mathematical proficiency

Adapting the marking schemes alone through the development of the specific question paper rubric generated another notable observation in relation to the resulting learners' marks. The proficiency aspects were only distinguishable at the key grade group levels or as a second order effect, but not at global level. The selected mathematical proficiency categories used in the developmental stage seem to add value to the assessment results at the local level of learners displaying similar global (aggregate) performance. At the global group level, all the proficiency categories which were made visible through the change of marking scheme only, were strongly correlated to the generalised, aggregated, Mathematics performance for the whole group of learners. This was referred to as a first order effect. The two distinct observations at two separate levels of learners' groupings that emerged from these analysis are discussed in the next sections, to highlight their usefulness in deepening assessment and informing teaching for mathematical proficiency.

8.3.1 The first order effect

The lack of distinguishability of the proficiency elements at global level was shown by the fact that the correlation test and the regression analysis showed a significant alignment between the performance of learners in the knowledge base and strategy categories when the whole sample of the remarked learners' scripts was analysed as one group. This observation was made with all four examination papers for the two years. The strong alignment at the global level suggested that assessing the separate proficiency categories, based on the current papers set, would not have an influence on the learners' overall performance in Mathematics. This has positive implications for the continuity of aggregate results, should these different proficiency categories be incorporated into the current system.

On the other hand, this first order effect does not seem to have any positive impact on the teaching of Mathematics. The lack of distinguishability at this level seems to suggest and encourage teaching to maintain the status quo. However, linking it to the distinguishability observed in learners in key grade groups, described as the second order effect, might significantly contribute to both development in assessment and Mathematics teaching as discussed in the next subsection.

8.3.2 The second order effect

As a second order effect observation, the two selected proficiency categories were found to be distinguishable as a result of the correlation tests and the simple linear regression analysis within separate key grade groupings. The results showed a different observation from that of the first order effect identified in the previous section. At this level, there were clear indications of possible differences between the ways in which learners' performed in the two selected categories, in all papers. Although the distinguishability of the learners' performance in these categories was significant, the differences were not the same at each key grade within and across a paper and also they did not show any particular pattern.

The observations above allow two inferences to be made that are important for this study. The first of these inferences is that the individual learners within one key grade group responded differently to the distinct proficiency categories. Some learners performed better in one category than the other. This could be as a result of differences in learners' learning and examination situations, as referred to by Bazerman (2012), who states that an individual externalisation process is uniquely influenced by situations. Secondly, each key grade group was made up of learners whose performance in the selected categories was different from each other, indicating a dynamic of abilities within the key grade groups.

The above observations were made despite the earlier observed dominancy of one category over the other in some of the examination question papers. Of particular concern was the dominancy of knowledge base over strategy in both papers 1 and 3 which constitute the Core level of the NSSCO examination. This could have disadvantaged learners whose thinking was more strategically oriented. Broadly, these differences did not show a particular group distinction as shown by the distribution of learners' category marks. This could be an indication that the questions did not discriminate well between the learners for all categories within each

key grade. Achieving a different, improved balance between the categories during the setting of examination question papers could lead to a different distribution of learners' category marks

The discussion in subsections 8.3.1 and 8.3.2 above indicates that just changing examination marking in line with the selected proficiency categories did not change much at the global level. Thus, at the level of all the marks or all leaners' performances, all the proficiency categories were strongly correlated to a generalized Mathematics performance. However, it resulted in distinct performances of learners in the two proficiency categories at the key grade level. These outcomes lead to some important contributions for this study in relation to developments in assessment and the Mathematics teaching process.

8.3.3 Important contributions to assessment and Mathematics teaching

The lack of distinguishability in the first order effect suggested that the relevancy of the overall NSSCO examination results were preserved and also assured the reliability of these results, in that the overall outcome of the examination can thus be expected to remain the same. This was despite the distinguishability of the proficiency elements through the distribution of marks in the adapted marking schemes. However, this would overcome the uncertainty that would be produced by any systematics change in high stakes assessment, as pointed out by Grewel (1993) and discussed in section 2.4.2.2. This could however be because adaptations were only made to the marking scheme. Extending these to the setting of question papers might result in a different observation.

The second order effect showing distinguishability of the proficiency elements, might be useful for assessment in that it could improve assessment feedback to include learners' achievements in individual proficiency elements. An additional important implication of the first and second order effects to assessment, in this regard, is that if the general performance only improves when all aspects of proficiency have improved then a lack in one aspect of proficiency might result in retarding the other aspects and so hold back improvements in general performance. In this case, the awareness of, and teaching to improve the retarding proficiency elements would be important for the teaching system. Therefore, a deliberate effort to balance the teaching of all aspects of proficiency could be valuable in improving the general performance.

Another interesting assertion was that the second order effect could also have been reduced by the dominance of one proficiency category over the other in some papers. Reducing the

dominance of one category over the other in some papers and testing the higher complexity levels of these competencies might allow for better discrimination of the learners in the key grade groups. Further research is thus needed to investigate appropriate Mathematics examination question paper design that would allow better discrimination within the key grade groupings in all papers in line with all aspects of the selected proficiency categories. The question papers need to include questions that test these categories at varied complexity levels. This was supported by the suggestions given by the markers that questions requiring argumentation, explanation and descriptions of the concepts, facts and operations needed to be included in the question papers to discriminate better between the learners' performance in both knowledge base and strategy categories (Ma2, Ma4).

The lack of distinguishability of proficiency elements, in the general performance of learners could also have been influenced by the fact that the examination questions were not designed to achieve a performance in different elements mathematics. The papers were designed to achieve a single measure of mathematical ability presented as the NSSCO Mathematics examination results. In which case, the examination questions might be a constraining factor for deepening insight into Mathematics assessment and its impact on Mathematics teaching. This also suggests a need for more research to explore how adapting examination questions would influence the distinguishability of the proficiency elements both in the general and in the group performance.

8.4 Insight into the NSSCO examination system

The mark distributions in the developed question paper rubrics together with the markers comments brought out insights into the current NSSCO examination system. This study looked at the possibility of using mathematical proficiency as an assessment framework base. The study was anchored on the general debates of development in Mathematics with a particular alignment to mathematical proficiency as formulated by Alan Schoenfeld (2007a&b) and other related literature. The second concept informed by the data analysed was the concern of assessing what is easy, rather than assessing all important aspects of mathematical proficiency. The underlying question guiding this part of discussion was: Can we make it easier to assess what is important? This section discusses these issues in relation to the data presented and analysed in chapters 5, 6 and 7.

8.4.1 Mathematics Proficiency as an assessment framework base

The NSSCO Mathematics examination has the general object of identifying learners' Mathematical competence, at the end of Grade 12. The purpose of the NSSCO examination system is to certify and allow for the selection of learners for further education at tertiary institutions or for employment on the job market. Extending this objective to include the detailed assessment of separate mathematical proficiency elements requires an understanding of the conditions for assessing mathematical proficiency. Assessment of a learner's mathematical proficiency would need to consider several aspects of mathematical proficiency (Schoenfeld, 2007a). The study considered and selected what were regarded as important elements of the mathematical proficiency according to Schoenfeld (2007a&b). It purposely excluded the aspect of disposition, not because this was viewed as less important but because of the challenges it was likely to pose to high stakes assessment at this stage. Schoenfeld (2007b) further stated that "knowledge plays a central role" (pp.71). This was also indicated by the study's results through the fact that the knowledge base category dominated in most cases, both in mark allocation and in the learners' performance. However, the learners' ability to use problem solving strategies and the ability to make use of what they know is equally important and should also be assessed.

An assessment framework based on mathematical proficiency would need to have the expectations for all aspects to be tested clearly spelled out (MoE, 2010c). The development of the expanded process standards and hence the generation of the general proficiency-complexity rubric in chapter 5 was aimed at clearly spelling out the possible expectations of the aspects selected to be assessed through an assessment framework based on mathematical proficiency. The proficiency traits expected to be visible at each level of complexity for each proficiency category were clearly described, see Appendix G. For example, 'implementing a strategy based on the application of a single memorised principle/process' is a general description of an atomic or basic strategy, which is at a complexity level 1. The expected specific features from the learners' work is shown through the following example: *Changing 256 kilometres per hour (km/h) to metres per second (m/s) by multiplying by 1000 and dividing by 3600.* Being an atomic strategy, the visibility of the strategy can also be shown by giving the correct answer only. However, to solve a problem like: *'A heptagon has four angles which are each 120°. The other three angles are equal. Calculate the size of one of these angles',* a more complex, composite strategy was needed to be made visible in the response. The expected response was

$5 \times 180 = 900$, $\frac{900-4(120)}{3}=140$, for paper 2, 2013. The response was described by the specific rubric as composite because it is a two-step strategy and involves a combination of complementary basic operations, see Appendix G. The general description given in the general rubric did not only enhance the validity of what was tested (Moss et al., 2006) but also its translation to specific question paper rubrics enhanced the reliability of the marking outcome. The general descriptions of all categories at all three complexity levels were given in the general rubric in Appendix G.

Despite the fact that some aspects of proficiency were seen as subtle and difficult to assess, assessing them is required and necessary as emphasized by (Schoenfeld, 2007a). However, to successfully assess them, it might require longer and wordier question stems and might also demand longer and more descriptive learners' responses to make what is required more visible. The research process of this study worked towards developing the assessment of the more subtle aspects of mathematical proficiency through two stages. Firstly, at the standard development or reformulation stage, where descriptive words were used to describe the expected measurable traits. The second stage was the development of the proficiency-complexity rubric which ensured that each of the expected levels of achievement were clearly described to visibly distinguish every creditworthy feature in the learners' response. A marker can only apply the marking scheme reliably if he or she was able to understand the requirements of the marking scheme and apply it consistently (MoE, 2010c). The markers indicated that after the training and the subsequent discussions during the initial round of the study, they could understand the requirements of the marking rubric – this process was aimed at standardising the rubrics (MoE, 2010c). The application of the adapted process standards to the NSSCO Mathematics examination through the remarking a sample of learners' scripts for past examinations showed the possibility of extending its application to any Mathematics examination system.

The process followed in this research revealed that, in order to successfully enhance the assessment of mathematical proficiency according to the description given by the applied categories, the tools and artefacts used by the NSSCO Mathematics examination or a targeted examination system need to be adapted. The adaption suggested by the markers, targeted the design of individual questions and that of the question papers. An examination question should clearly state what the examiner expects to see in the learners' response to that question (MoE,

2010c). In other words, if the examiners expect to see more proficiency to be visible in the learners' responses to the questions, the questions should be set in a way to demand high proficiency from the learners. For that reason, the adaptation proposed by the markers for the questions and thus the question papers will contribute to ensuring that the competencies described by the expanded process standards are sufficiently assessed.

In addition to the observed applicability, the markers stressed that it brings a necessary difference on how examination was viewed before (Ma1, Ma2, Ma4 and Ma6). The markers highlighted that one of the advantages of the proficiency system was that learners were required to show the details of what they are able to do (Ma2, Ma6 and Ma7). As a result, teachers who are to be used for the marking process, should be able to identify the learners' challenges and weakness easily and be able to advise future teaching through the examiners' report (Ma6) which is currently used as an assessment feedback system.

8.4.2 Assessing valued aspects of mathematical proficiency not just what is easy to assess

The three categories chosen as the aspects of mathematical proficiency investigated in this are said to be intertwined (Schoenfeld, 2007a) and hence are all important. The intertwinement could suggest that the assessment of one category cannot be done separately from the others. Although the two dimensional grid of the proficiency-complexity rubric which was developed, shown in Appendix G, shows a complexity continuum of individual categories. Its application was done simultaneously to reflect the relationship between these categories. The strategy category, for example, cannot be tested without knowledge base. An example from Paper 1 of 2013, question 9 shows this; "Factorise completely $4x^2 + 12xy$" the answer is $4x(x + 3y)$. The following categories were assessed: Level 1 knowledge base as the knowledge of factors of a binomial expression and level 1 strategy, crediting the actual process of getting both correct factors. Likewise, the metacognition category cannot be assessed without assessing knowledge base alone or both knowledge base and strategy. The above, in relation to this study, reflects the intertwinement of the proficiency categories as described in the literature studied.

Measuring just the separate categories does not tell how proficient a learner is (Burkhartdt, 2007; Schoenfeld, 2007a). Therefore, considering all three aspects of Schoenfeld's (2007a) proficiency and incorporating argumentation was to ensure that important aspects of mathematical proficiency were assessed and not just what is easy to judge. The findings reflect

the possibility of distinguishing individual elements of proficiency through high stakes examinations. Thus, despite the important interrelationship between the categories, their observed distinctiveness shows their potential to be validly assessed and be used to describe the learners' performance in Mathematics assessed through high stakes assessment.

Congruent to Burkhardt's (2007) observation and concern, yet contrary to the principle of mathematical proficiency, NSSCO Mathematics examinations seem to assess more what is easy to assess than what is important for the learners to know. It was observed that the knowledge of concepts or relationships used was never explicitly made visible through either definitions or any verbal representation. The learner's knowledge of concepts and relationships was only shown through either graphical representation or as part of an arithmetic calculation which that as stated earlier was mainly presented as a final answer (Ma6, Ma3). One of the concerns of this study is posed in the questions: How can we easily assess what seems difficult to assess? How well can we describe the traits expected by proficiency in an effort to make them more visible? In the context of this study as explained in chapter 2, visibility of proficiency is defined as the extent to which the tools enable proficiency to be shown by the learners and noticed by the markers in the learners' responses. Therefore, it has been established through the study that it is possible to assess separate and important elements of mathematical proficiency through the NSSCO Mathematics high stakes examination, but the tools that drive the examination need to be adapted in order to make the desired aspects of proficiency clearly visible and validly assessed.

Burkhardt (2007) stressed that assessment should assess what is important to assess and not just what is easy to measure. It was not always clear whether the leaners were truly knowledgeable about the concepts or whether they merely memorised the representations. Hence the major change to the NSSCO Mathematics process standards, through this study, was to ensure that the learners also respond in a verbal form which was seen to enhance the visibility of their thinking and encourage teaching with understanding. Additionally, it came out through the markers discussion that knowledge base features were not always clearly visible to the markers in the learners' response. Some expressed their concern by saying that, sometimes the knowledge base marks were awarded on assumption that the learners had the knowledge of the concept (Ma7). In other words, if a learner has successfully solved a mathematical problem, he or she was assumed to have acquired the knowledge of the concept on which that problem was based.

One of the principles of reliable assessment and marking is that of ensuring that the knowledge and skills which are to be examined and the awarding of marks during the marking process, should be made explicit through the questioning and also through the marking criteria (MoE, 2010c). It was clear from the literature studied that the knowledge of concept/fact or relationships seems to be made more visible through a definition, description or an explanation not just through procedural, arithmetic calculation (Burkhardt, 2007; Milgram, 2007). As urged by Milgram (2007) mathematical procedures and arithmetic calculations can easily be memorised and regurgitated on demand during an examination. In line with this study, the NSSCO Mathematics examination should allow learners to show how proficient they are in Mathematics. Learners should show how they are able to devise alternative approaches when they get stuck or they hit road blocks (Schoenfeld, 2007a). They should be able to show that they are efficient with and make use of what they know in problem solving.

8.5 Conclusion

The data collected for the study and discussed in this chapter shows the viability of assessing distinct elements of mathematical proficiency through high stakes assessment. The possible distinction of the selected mathematical proficiency categories observed in this study, highlights the possibility of focusing the assessment system on assessing what is found to be important to assess. Thus in relation to this study, if it is important to assess metacognition then the examination papers should be designed to include questions that better assess metacognition. There is a greater need to shift the NSSCO Mathematics examinations from placing more emphasis on the final arithmetic answer to also assessing the learners' thinking which lead them to specific answers. The latter would enhance classroom teaching to direct the teaching of required and important aspects of mathematical proficiency.

CHAPTER 9: Research Conclusion

9.1 Introduction

This chapter first summarises the context and the process of the study. It continues by presenting the research findings and conclusions on the issues researched. The way forward and recommendations for future studies are also given. The chapter ends by underlining the limitations of the study.

9.2 The summary of the study's context and process

This study was aimed at gaining an understanding of how mathematical proficiency can be assessed better in the NSSCO Mathematics examination. The study was conducted in the Senior Secondary school Mathematics high stakes examination context. The NSSCO Mathematics Examinations are high stakes examinations written through a paper and pencil method, where learners are required to respond to a series of questions. High stake examinations are seen to drive teaching and learning. Therefore, in the view of study the examination might be the best tool to motivate the teaching and learning for mathematical proficiency.

The study followed a developmental research method in an exploratory paradigm. Examination tools were adapted to enable the distinct assessment of different aspects of mathematical proficiency. The developed tools were then evaluated using the learners' responses to the NSSCO Mathematics question papers to investigate the distinguishability of the selected aspects of mathematical proficiency.

9.3 The research findings

The study was designed with the general aim of gaining an understanding of how the assessment of mathematical proficiency may be improved in high stakes assessment, particularly in the case of the NSSCO Mathematics examination. This section presents the final conclusions arising from the research. The findings showed the possibilities for improvement of the NSSCO Mathematics examination system that were offered by adapting the process standards, the marking criteria and the marking rubric only. Suggestions as to how the adaptation could be extended to the examination question papers were also given.

9.3.1 Identifying assessable proficiency categories

During the first part of the research process, the current examination tools were adapted in order to extend the external validity of mathematical proficiency assessment using the NSSCO Mathematics examination papers and learners' answer scripts. To do this, different concepts and ideas from the literature and research on mathematical proficiency were incorporated into the NSSCO Mathematics examination tools. Various interpretations of the selected mathematical proficiency categories were used to modify the NSSCO examination tools in line with these categories. The research sought for a distinction of different elements of proficiency and an alignment of the examination tools with these proficiency elements.

The literature consulted, clearly and separately described the competencies that characterise individual mathematical proficiency categories. The current competencies highlighted by the NSSCO tools, particularly the process standards, were matched and aligned with this literature, allowing a conceptual alignment of the two sets of competencies. The process resulted in the generation of an adapted NSSCO examination system to enable the alignment of the cognitive process and the subject content in the examination tools with the process described in the selected proficiency categories (Brookhart and Nikto, 2007).

The proficiency categories selected for this study were three of those proposed by Schoenfeld (2001b): knowledge base, strategy and metacognition. His fourth category, disposition, was not included in this study because this related more to the orientation of the learner to Mathematics, than to explicitly presented mathematical performance and was therefore seen to be challenging to assess through high stakes examinations. To ensure conceptual distinction the literature consulted was used to generate the following specific characteristics of each:

(a) Knowledge base; refers to comprehension and understanding of mathematical concepts, facts and operations and includes using precise definitions and description of all terms, operations and properties of the operations.

(b) Strategy; refers to methods used to arrive at an answer and includes a learner's capacity for computation, logical thought, reflection, explanation, and reasoning to highlight the need for argumentation.

(c) Metacognition; refers to using what one knows effectively and monitoring one's own thinking. Metacognition may also be defined as 'thinking about thinking' (Schoenfeld, 2007a).

9.3.2 Generating assessment tools

Tools generated during this phase of the research were the *adapted process standards*, the *general proficiency-complexity rubric* and the *specific question paper rubrics*. Based on the aligned process standards, prototypical responses were developed for questions from each of the past NSSCO Mathematics question papers. These responses were then analysed to identify mathematical competencies that could be categorised within the separate proficiency categories of the aligned process standards. This enabled the development of *a general proficiency-complexity rubric* that described the features expected to be noticed by markers when evaluating the learners' response. Additionally, the complexity of the prototypical responses for each proficiency category were rated. This process was followed by the researcher and the focus group of markers, for each of the four Papers, 1, 2, 3 and 4 used in the research for the two years 2013 and 2014. In this way, a *specific question paper rubric* was generated for each of these papers. In the specific question paper rubric, the expected marks were allocated to each proficiency category, at different levels of complexity, depending on the questions' expected response.

9.3.3 Distinction of proficiency aspects

The mark allocations for the different aspects of proficiency in the specific question paper rubric revealed information that suggested that the selected elements were distinguishable. A chi-squared analysis was carried out to investigate the statistical distinction of the mark distributions for each proficiency category in each paper. The results indicated that the different response categories were statistically separable in each paper. This supported the judgement based on the literature that the two selected response categories (knowledge base and strategy) were indeed conceptually distinguishable.

The distinguishability of the selected proficiency categories was essential to ensure the content and construct validity of the rubric designed on the developed standards and also the reliability of assessment outcomes assessing the individual categories separately (Moss et Al., 2006). The adapted system yielded distinct mark distributions for every NSSCO examination question paper within a system that was not set with this distinction in mind. This suggests an externally valid alignment of the developed assessment system.

9.3.4 Understanding distinction through learners' results

The markers used the specific question paper rubrics to assess and remark a sample of the learners' answer scripts to the corresponding NSSCO examination question papers. They observed and matched the features shown in the prototypical responses to the features in the learners' responses and awarded them marks corresponding to the separate categories at different levels of complexity.

The learners' results, obtained by remarking responses to these papers, set under the previous system, were then analysed. During the analysis it was observed that only the two categories, knowledge base and the strategy category contributed significantly to the total marks for each paper as discussed in chapter 5 and 6. Therefore, the third category, metacognition, which contributed very little or nothing at all in some papers was excluded in the analysis of learners' marks. This analysis generated two major conclusions:

(a) When the complete sample of learners was analysed as a single entity (at the global level) the two proficiency categories were not statistically distinguishable from each other, or from the previous marks obtained. That is, the overall outcome of the examination for each learner did not appear to be affected by the separate contribution of the two categories. This indicated that the existing NSSCO single mark results were consistent with those that would be obtained using the adapted marking system for examination papers set under the previous system. This alignment of results at the level of the full group of learners was regarded as a first order effect of the adaptations, on learners' marks.

(b) A second order effect of the adaptation of the system was observed within the sub-groups of learners selected according to the key grades. At this level of analysis, the two proficiency categories were not highly correlated, suggesting a possible distinction of the two proficiency categories within groups of learners showing similar general performance. Within the same key grade group, the two selected categories each had a distinct contribution to an individual learner's performance.

The above observations have direct positive implications for teaching for mathematical proficiency in the senior secondary Mathematics classrooms in Namibia. The second order effect highlights the importance of distinguishability of the proficiency categories in relation

to the learners' natural orientation to Mathematics and suggests that particular learners might have natural preferences. Distinguishing proficiency categories and ensuring each category is taught properly and sufficiently will benefit learners differently on individual levels. In addition, the first order effect gives an assurance that the learners' single examination mark will remain comparable to the traditional set of standards in relation to the other learners. However, possible limitations on the adaptation of tools, that were identified by the focus group of markers, may have had a negative influence on the distinguishability at both the first second order level, indicating that the adapted tools were still too blunt and needed to be further developed to allow for an effective distinction.

The NSSCO curriculum process standards can be adapted by using and incorporating the description of the mathematical proficiency categories given by different authors and matching them to the current standards. The resulting set of standards can be used to more effectively assess mathematical proficiency through high stakes assessment, particularly through the NSSCO. However, other examination tools might need to be adapted as well in order to ensure that the selected proficiency categories may be effectively distinguished and assessed within the examination structure. The adaptations done and used in the remarking process were restricted to some tools only, due to the functioning nature of the NSSCO examination used in the study. However, markers performed a qualitative evaluation to highlight constraints and suggested adaptations that could be made to the examination question papers.

9.3.5 Understanding the conceptualization of current examinations

The structure of the question papers as related to the identified proficiency categories and complexity levels was explored using the mark allocations in the specific question paper rubrics. This yielded the following insight into the possible proficiency structure of the existing exam papers.

(a) The knowledge base category made up the highest percentage of the skills tested in Papers 1 and 3. The two papers were originally designed for the Core level in the NSSCO Mathematics examination. The skills and competencies tested in these two papers were also dominantly pitched at complexity level 1 of this category, the complexity level 2 of the knowledge base category was hardly tested.

(b) The majority of the skills and competencies tested in the extended level papers of the NSSCO Mathematics examination (Papers 2 and 4 together), were shown through structured responses, resulting in the majority of the marks being allocated to the strategy

category. However, Paper 2 also had a higher percentage of knowledge base tested. Only Paper 4 assessed more strategy than knowledge base. The strategy skills tested were also predominantly pitched at the lower levels of complexity of the category that is complexity levels 1 and 2, level 3 was not tested.

Paper 1 and 2 were originally designed to have questions with short to semi-structured answers. Questions that mainly demanded a final answer, involving either recalling of concepts or one step to two steps arithmetic calculation questions. The previous marking system also encouraged the learners to give a final answer only since the final answer was always given full credit with or without working in these two papers (Ma2 and Ma6). In contrast to this, Paper 3 and 4 were originally designed to have questions that required structured answers. Questions set for these papers demanded learners to show the procedures followed in order to get the solution. It was therefore not expected to have the majority of Paper 3 percentage marks allocated to knowledge base.

The original examinations which, as mentioned in chapter 3, were designed to have each examination level consisting of two papers; a question paper that demanded short answers and a paper with questions that demanded structured answers. The Core examination level consisted of Papers 1 and 3 and Extended examination level consisted of Papers 2 and 4. Therefore, the two levels were designed to have a similar composition in terms of the demand in relation with the type of answers demanded by the questions. Hence, it was expected to have a similar distribution of marks in terms of the proficiency categories. The dominance of knowledge base category in the core examination level as compared to the Extended examination level, reflects some of the valuable additional insight that the new system provides.

The above inferences do not only give indications of the current expectations of the two NSSCO Mathematics examination levels but could also serve as a starting point for possible adaptations that would bring an improved mathematical assessment process in relation to the setting of question papers. The emerging proficiency structure would allow for a clearer specification of the criteria for the expectations of the Mathematics high stakes examination papers in line with the selected categories.

9.3.6 NSSCO constraints and adaptations

After the development of the specific question paper rubrics and after the remarking process, each marker wrote a qualitative report. The reports included their general observations on different aspects of the research process in order to compare the two systems of marking and the usability of the proficiency marking system in marking separate categories of proficiency. The markers also included suggestions on the possible adaptation of the NSSCO Mathematics examination system, for both the design of the questions and the overall construction of the question papers. Possible adaptations were also suggested during the markers' oral discussions during the marking. These were recorded by the researcher who acted as a moderator during these discussions. These reports were analysed and the following conclusions emerged;

(a) Most questions in the NSSCO did not request learners to show their working and full credit was given to a final answer only, while the proficiency system gave credit for various separate skills and competence displayed within the learners' full response, including working to get the final answer.

(b) The NSSCO Mathematics question papers did not appear to assess the complete spectrum of complexity of the separate proficiency categories including their higher complexity levels.

(c) The question papers could be designed such that specific number of marks are assigned to the separate competencies; knowledge base, strategy and metacognition in a predetermined proportion.

(d) The questions should be constructed in such a way that each separate competency demanded is apparent and should clearly show how each competency is going to be credited. One marker, Ma4 gave a suggestion using Question 11(a), paper 4, 2013, as an example. Question 11(a) required the learners to use the diagram given to choose a cup that contained more liquid. He suggested that the question could have been set to ask learners to first write the two volumes in a common unit, for example millilitres. This could have allowed better assessment of knowledge of capacity and the strategy used in conversion leading them to their conclusion. Otherwise, if the learners are just asked to choose which cup contains more of the liquid, some learners might guess their answer (Ma4) and the competencies demanded are not immediately apparent.

The above deductions indicated the possibility of improving the assessment of the separate elements of proficiency, by adapting the design of some examination questions and the construction of the examination question papers. The questions can be designed to ensure that

they clearly demanded to have specific proficiency-complexity elements visible. The changes to the questions would also ensure the inclusion of questions that ask for reasons, explanations or discussions as indicated by markers Ma4 and Ma7, as some of the competencies that were missing within the old examination question design. Having such adapted questions would also give a chance to extend the testing of a broader range of metacognition skills since it was identified that the skills tested in the current system under metacognition was only the choice of the strategies that could be used to solve a given mathematical problem (Ma4, Ma7).

The adaptation of the question papers, could also ensure the appropriate representation of the separate proficiency categories. Having a predetermined mark allocation per category and complexity level was identified as the main way of improving the assessment of proficiency which would encourage teaching for proficiency. Both elements of proficiency need to be proportionally represented at different complexity levels. Therefore, the contribution of each category needs to be pre-determined prior to the design of the question papers.

9.4 The research recommendations and the way forward

This section presents the way forward and recommendations arising from the study in order to address the constraints and limitations identified in the findings. Addressing constraints and limitations will possibly require piloting the complete set of adapted curriculum process standards through teaching and the adaptation of questions in papers.

It was shown in this study that the adapted tools can be used to guide the marking process only. It provides a total mark allocation that was consistent with the original marks but at the same time provides extra information on the performance of the learners in relation to the separate proficiency categories. However high stakes assessment outcomes need to be valid and reliable. Therefore, in order to ensure validity of the outcomes of this assessment system, the learners need to be aware of the skills demanded through the style of questioning. The questions need to be set in such a way that learners are able to work out how much is demanded from them. The general recommendations emerging from the study are listed as follow:

(a) There is a greater need to adapt the NSSCO Mathematics process standards to include all distinguishable proficiency categories and ensuring each category is taught properly and sufficiently will benefit learners differently on individual levels in preparation for the NSSCO high stakes examinations.

(b) There is a greater need to shift the NSSCO Mathematics examinations from placing more emphasis on the final arithmetic answer to also assessing the learners' mathematical thinking.

(c) The NSSCO Mathematics examinations question papers should be designed to include a specified proficiency balance with a specified complexity balance which examines a complete spectrum of proficiency categories at all complexity levels.

(d) The NSSCO Mathematics questions should be designed to ensure that the expected proficiency competencies and complexity levels are explicitly and clearly demanded to ensure that the learners have an equal chance of making them visible through their responses.

The above recommendations were further expounded to indicate whether minor or major adaptations would be required before they are implemented. In the cases where major adaptations would be required they are classified as areas for further research. The recommendations arising out of this study are in three sets;

- Minor adaptations that could be done on the NSSCO examination questions.
- Using the developed proficiency assessment tools for marking.
- Changes that require further research studies.

The discussions on the adaptations to the NSSCO Mathematics examination framework are given in the next two sections, 9.4.1 and 9.4.2.

9.4.1 Minor adaptations to the NSSCO Mathematics examination questions

The style of questioning used to implement the current NSSCO curriculum standards for Mathematics concentrates more on the application of several mathematical facts and concepts to arithmetic calculations. Therefore, some adaptation to the system can be done through changing the style of questioning without having to change the process standards. For example:

(a) NSSCO Mathematics questions should be designed to required learners to define facts or concepts or describe strategies or methods of obtaining a certain answer.

- Requiring the learners to define or give a description of some facts and concepts might not require any change to the existing standards but only formulating questions which require definitions and descriptions.
- Asking more complex levels for both knowledge base and strategy category might only need question adaptation and not the adaptation of standards.

(b) The NSSCO Mathematics question papers should be designed with a reasonable balance of the proficiency competencies already described by the current process standards. Ensuring a reasonable balance between knowledge base and strategy category might only need adaptation to the question paper structure and not the adaptation of standards.

These were regarded as minor adaptations that could be implemented at the level of question paper construction. Therefore, it is recommended that these minor changes be incrementally introduced to the NSSCO Mathematics system.

Together with the above changes, the marking should follow the new marking system to ensure that the different categories are scored differently. The adapted marking system would retain compatibility with previous marks, provided that the adaptations for the setting of the question papers is controlled to be not too drastic, as indicated by the analysis of the learners' results at the global level.

9.4.2 Areas for further research

The focus of this study was to provide a better understanding on the possibility of improving the assessment of mathematical proficiency based on the current NSSCO examination papers. The distinction of the selected proficiency categories observed through the second order effect reflected the potential for improving the assessment of these categories at different complexity levels, which could lead to improved discrimination of the learners' performance within the key grade groups. This does not only demonstrate that the selected proficiency categories provide deeper insight or function as useful tools for the current exam structure, but also shows its potential to enhance the feedback provided to schools as a result of the high stakes examinations. The feedback can include additional information in terms of two or more different conceptual aspects, in addition to a single aspect represented by one single mark given by the global observation. It is clear that a number of possibilities exist but some improvements did not appear to be attainable using the current NSSCO examination papers structure.

A follow up study would be necessary to investigate possibilities for improvement that involve adapting the questions and the question papers and possibly reworking the current process standards. Research into this issue is important in order to ensure the validity of such changes in the wide NSSCO examination context. The process should be designed with the following objectives;

(a) To determine the style of questioning to be done to ensure that the gaps identified are tested and,

(b) To determine the appropriate proficiency-complexity structure suitable for the assessment at each level and each paper in order to better discriminate the learners' capabilities in the selected proficiency categories.

(c) To determine possible ways to work out a desired relationship between the categories in order to improve the setting of papers specification and evaluation.

In the light of further studies the following points are important and need to be built into the recommended further research goals and objectives.

(a) Some learners' responses to more complex knowledge base and strategy category would have subjective elements and would need to be assessed accordingly.

(b) Some questions based on more complex knowledge base and strategy category would be wordy.

9.5 Limitations of the study

(a) While the findings of the study provided valuable insight that suggest the external validity of this assessment as related to conceptual knowledge and strategic thinking, it should be noted that the possibility of objective verification of external validity is a contentious issue.

(b) The study was an exploratory investigation using a small sample from the NSSCO Mathematics examination system. The conclusions of the study should thus be interpreted as indicative, rather than as objectively verified. Follow up studies using different designs would be required to validate and generalise these outcomes.

REFERENCES

Agaton, P. P. Jr. (2016). The implication of Learning Theories to Assessment and Instructional Scaffolding Techniques. *American Journal of Educational Research*. 4(9): 695 – 700.

Airasian, P.W (1988). Symbolic Validation The case of State-Mandated, High stakes Testing, *"Education Evaluation and Policy Analysis"* 10(4): 301 – 313.

Aliaga, M. & Ginderson, B. (2000). Interactive statistics. *Saddle River*, p3-15.

Amiel, T. & Reeves, T. C. (2008) Design-Base Research and Educational Technology: Rethinking Technology and the Research Agenda. *Educational Technology & Society*, 11(4), 29-40.

Avenstrup, R. (1993). The challenges of examination localization. *In proceedings for National Conference on IGCSE and HIGCSE*. Swakopmund. MoEC.

Baker, E.L & Stites, R. (1990). "Trends in Testing in the USA" in Fuhrman S.H & Malen, B eds., *The politics of curriculum and testing (Yearbook, 1990)* 139 -157, London, England: Farmer Press.

Ball, D. L. (2003). *Mathematical Proficiency for All Students: Towards a Strstegic Research and Development Program in Mathematics Education*. RAND. Sant Monica, CA.

Bazerman, C (2012). "Writing with Concepts: Communal, Internalized, and Externalized*".* *Mind, Culture, and Activity,* 19: 259–272.

Beaco, J., Bryam, M., Cavalli, M., Coste, D., Cuenant, M. E., Guollier, F. & Panthier, J. (2010). Guide for the development and implementation of curricula for plurilingual and intercultural education. www.coe.int/lang. (Retrieved on 29 January 2017).

Black, P. ; Burkhardt, H.; Daro, P.; Jones, I.; Lappan, G.; Pead, D. & Stephens, M. (2012). *High stakes Examinations to Support Policy. Educational Designer, 2(5).* http://www.educationaldesigner.org/ed/volume2/issue5/article16/. (Retrieved on 7 August 2013).

Black, B., Suto, I. & Bramley, T. (2011). The interpretation of features of questions, marking schemes and examinee response and their impact upon marker agreement. *Assessment in Education: Principles, Policy & Practice*. 18(3): 295 – 318.

Biggs, J. (2003). Teaching for quality learning at university: *What the student does*. Second Edition, Society for Research into Higher Education: Open University Press, Buckingham.

Brookhart, S.M & Nitko, A.J (2007). *Assessment and Grading in Classrooms,* New Jersey: Pearson Education, Inc.

Bude, U. & Lewin, K. M. (1997). *Assessment of Science and Agriculture in primary school in Africa;* 12 countries cases reviewed. 1(2) Improving test design.

Burkhardt, H. (2007). Mathematical Proficiency: What is important? How can it be assessed? In Schoenfeld A.H. Eds. *Assessing Mathematical Proficiency,* 77 - 97 United Kingdom: Cambridge University Press.

Carlson, M.P. & Bloom, I. (2005). The cyclic nature of problem solving: An emergent Multidimensional Problem-solving Framework. *Educational Studies in Mathematics*, 58(1) 45 – 75.

Clarke, D., Goos, M. & Morony, W. (2007). Problem solving and working mathematically: An Australian perspective. *ZDM Mathematics Education,* 39, 475 – 490.

Cohen, L., Manion, L. & Morrison, K. R. B. (2007) *A Guide to Teaching Practice* (sixth edition).London: Routledge.

Crotty, M. (2003): The Foundations of Social Research: Meaning and Perspectives in the Research Process, London: Sage Publications, 3rd edition, 10.

Daniels, H. (2001). *Vygotsky and Pedagogy*. London: Ruotledge Farmer.

Dawson, C. (2002).Practical Research Methods, New Delhi, UBS Publishers' Distributors.

Dey, I. (2005).Qualitative Data Analysis: A User – Friendly Guide for Social Scientists, New York: Taylor & Francis e-library.

Earl, L. (2003). *Assessment as Learning: Using Classroom Assessment to Maximise Student Learning.* Thousand Oaks, CA, Corwin Press.

Ellis, T. J. & Levy, Y. (2009). Towards a guide for novice researchers on research methodology: Review and proposed methods. *Issues in Information Science and Information Technology*, 6, 323-337.

Ericsson, K. A & Simon, H. A. (1998). How to study thinking in everyday life: *Contrasting Think Aloud Protocols with Descriptions and Explanations of Thinking* Mind Culture and Activity.

Ferrara, S., Svetina, D., Skucha, S., & Murphy, A. (2011). Test design with performance standards and achievement growth in mind. *Educational Measurement: Issues and Practice,* 30 (4), 3-15.

Green, S. (2002). *Criterion Referenced Assessment As A Guide To Learning -The Importance of Progression and Reliability*; A paper presented at the Association for the Study of Evaluation in Education in Southern Africa International Conference, Johannesburg.

Grix, J. (2004). The foundations of research. London: Palgrave Macmillan.

Groves, S. (2012). Developing mathematical proficiency, Journal of science and mathematics education in Southeast Asia, 35(2) pp. 119-145.

Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report,* 8 (4), 597-607.

Gourgey, A. F. (1998). Metacognition in basic skills instruction. *Instructional Science: An International Journal of the Learning Sciences*, 26(1), 81-96.

Govil, P. (2013). Ethical Considerations in Educational Research. *International Journal of Advancement in Education and Social Sciences.* 2(1) 17 – 22.

Haapasalo, L. (2003). The Conflict between Conceptual and Procedural Knowledge: Should We Need to Understand in Order to be Able to Do, or vice versa? In L. Haapasalo & K. Sormunen (eds.) *Towards Meaningful Mathematics and Science Education*. Proceedings on the 19th Symposium of the Finnish Mathematics and Science Education Research Association. University of Joensuu. Bulletins of the Faculty of Education 86, pp. 1-20.

Haapasalo, L. & Kadijevich, D. (2000). Two types of mathematical knowledge and their relation. JMD, Journal for Mathematic – Didaktik, 21, 139 -157.

Hardman, J. (2008). Researching pedagogy: an Activity Theory Approach. *Journal of Education*, 48, 65 – 95.

Harel, G. (2008). DNR perspective on Mathematics curriculum and instruction, Part I: focus on proving, *ZDM Mathematics Education*, 40, 487–500.

Harlen, W. (2005). Teacher's summative practices and assessment for learning – tensions and synegy. *The Curriculum Journal.*16, 207 – 223.

Hasan, H. (2003). Information systems development as a research method. *Australasian Journal of Information Systems,* 11(1), 4 – 12.

Hiebert, E. H. (2011). *Readability and the common core's staircase of text complexity*. Text Project & University of California, Santa Cruz.

Hogg, R. V., Mckean, J. W. & Craig, A. T. (2013). *Introduction to mathematical statistics*, (7^{th} ed). Pearson Education, Inc

Hoover, M. (2014). *Assessing Math to know Math.* Mathematical Sciences and Research Institute. Berkeley, California.

Huba, M & Freed, J. (2000). *Learner-Centered Assessment on College Campuses*. Boston: Allyn & Bacon.

Ibrahim, A.M. (2012) Thematic Analysis: A Critical Review of its Process and Evaluation, *Academic Conference Proceedings, WEI International European*. October 14-17, 2012, Zagreb, Croatia.

Iipinge, R. (2012). *An investigation into the perceptions of the first year Mathematics students towards the alternative mode intervention: UNAM case study:* A thesis submitted in partial fulfilment of the requirement for the degree of Master of Education. Rhodes University.

Iipinge, S. M & Likando, G. N. (2012). The Education assessment reforms in post-independence Namibia: A critical analysis. *SA-eDUC journal* 9(2).

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement (4th ed., pp. 17-64).* Washington, DC: The National Council on Measurement in Education & the American Council on Education.

Kellagaghan, T and Greaney, V (2003); *Monitoring Performance: Assessment and Examinations in Africa*. Unpublished report for the association for the Development of Education in Africa (ADEA) presented in December 2003, Grand Baie, Mauritius.

Kellogg, R. T. (2007). Are written and spoken recall of text equivalent? *The American Journal of Psychology*, 120, 415–428.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory into Practice,* (41)4.

Kilpatrick, J., Swafford, J., Findell, B. (Eds.). Mathematics Learning Study Committee (2001) *Adding it up: Helping children learn Mathematics.* Washington DC: National Academy Press.

Kothari, C. R. (2004). *Research Methodology: Methods and Techniques* (second revised edition). New Delhi. New Age International (P) Limited, Publishers.

Kumar, R. (2011). *Research Methodology-A Step-by-Step Guide for Beginners*, (3nd.ed). London EC1Y 1SP, Sage Publication Ltd.

Lawrence, J. A. & Valsiner, J. (2003). Making personal sense: An account of Basic internalization and externalization process. *Theory and Psychology*, 13(6), 723 – 752.

Le Cornu, A. (2009). Meaning, Internalization and Externalization. Toward a Fuller Understanding of the Process of Reflection and Its Role in the Construction of the Self. *Adult Education Quarterly* 59(4), 279-297.

Leyendecker, R.W. (2008).*Curriculum reform in Sub-Saharan Africa: Opportunities and obstacles*. German: Print Partners Ipskamp-Enschede.

Lester, F. K. (1994). Musing about Mathematics problem solving research: 1970-1994. *Journal for Research in Mathematics Education,* 25, 660 - 675.

Lewin, K. & Dunne, M. (2000). Policy and Practice in Assessment in Anglophone Africa: does globalization explain convergence? *Assessment in Education: principles, policy and practice.* 7(3), 363 – 377.

Lin, G. (2009). Higher Education Research Methodology – Literature Method. *International Education Studies*, 2(4), 179 – 181.

Linn, R. L. (1993). Educational Assessment: Expanded Expectations and Challenges. *Educational Evaluation and Policy Analysis*, 15(1) 1-16.

Lutz, S. & Huitt, W. (2003). Information processing and memory: Theory and application. *Educational Psychology Interaction.*Valdosta, GA: Valdosta State University. Retrieved [19/01/2018), from http://www.edpsycinteractive.org/papers/infoproc.pdf

MacDonald, M. (2001).Finding a critical perspective in GT. In R. Schreiber & P. N. Stern (Eds), *Using grounded theory in nursing* (pp. 113 – 158). New York, NY: Springer.

Meadows, M. & Billington, L. (2005). *A Review of the Literature on Marking Reliability*. Report for the National Assessment Agency by AQA Centre for Education Research and Policy.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7 (3).

Messick, S. (1995). Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning. *American Psychological Association*, Inc. 0003-066X/95/S2.00 50(9) 741-749

Milgram, J.R. (2007). What is Mathematical Proficiency? In Schoenfeld A.H. Eds. *Assessing Mathematical Proficiency*, 31– 58, United Kingdom: Cambridge University Press.

Moss, P. A., Girard, B. J. & Haniford, L. C. (2006). Validity in Educational Assessment. *Review of Research in Education, Rethinking Learning: What Counts as Learning and What Learning Counts*. 30, 109-16

Namibia. Ministry of Education (2012); Procedures for Grading and Grade Review: Namibia Senior Secondary Certificate, Ordinary and Higher Level, DNEA, Namibia.

Namibia. Ministry of Education (2010 c). *Guidelines for setting and marking assessment tasks DNEA*, Windhoek, Namibia.

Namibia. Ministry of Education (2010 b). *Namibia Senior Secondary Certificate (NSSC) Mathematics Syllabus Ordinary Level Syllabus*, Grades 11 - 12. NIED, Okahandja, Namibia.

Namibia. Ministry of Education (2010 a). *The National Curriculum for Basic Education*. NIED, Okahandja, Namibia.

Namibia. Ministry of Education (2006). *Education and training sector improvement programme (ETSIP): Planning for a learning nation*. Windhoek: Ministry of Education.

Namibia. Ministry of Education and Culture (1993). *Towards education for all: A development brief for education, culture and training.* Windhoek: Gamsberg Macmillan.

National Research Council. (1993). *Measuring what counts: A conceptual guide for Mathematics assessment.* Washington, DC: National Academy Press.

Niebling, B. C. (2013). The role of Cognitive Complexity in Competency – Based Education. A paper presented at the Iowa ASCD conference, June 26th 2013.

Nitko, A. J. (2001). *Educational assessment of students* (3rd ed.). Upper Saddle River, NJ: Merrill.

Ottevanger, W., van den Akker, J. & de Feiter, L. (2007). Developing Science, Mathematics, and ICT Education in Sub-Saharan Africa: Patterns and Promising Practices. *World Bank Working Paper* No. 101.

Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods*, (2nd ed.) Newbury Park, California. Sage Publication, Inc.

Pape, J.S. & Tchoshanov, M. A. (2001). The role of Representation(s) in Developing Mathematics Understanding. Theory into Practice, 40(2) 118 – 127.

Plomp, T. (2013). Educational Design Research: An Introduction. In T. Plomp & N. Nieveen, *Educational Design Research-Part A: An Introduction*, 10 - 51. Netherlands, SLO

Polya, G. (1957). *How to solve it: A new aspects of mathematical methods*. United States of America, Princeton University Press.

Pollitt, A., Elliot, G, & Ahmed, A. (2004). *Let's stop marking exams*. Paper presented at the IAEA Conference, Philadelphia. June 2004.

Pugalee, D. K (2001) "Writing, Mathematics and Metacognition" *School Science and Mathematics* 101(5), 236-254.

Putnam, A. L. & Roediger III, H. L (2012). Does response mode affect amount recalled or the magnitude of the testing effect? *Memory Cognition,* 41:36–48

Resnick, D & Resnick, L (1992). "Assessing the Thinking Curriculum: New Tools for Education Reform", in Gifford, B.R and O'Connor, M eds., *Changing Assessment: Alternative views of Aptitude, Achievement, and Instruction,* 37 – 75; Boston: Kluwer Academic Publishers.

Richey, R. C., Klein, J. D. & Nelson, W. (2004). Development research: Studies of instructional design and development. In D. Jonassen (Ed.) *Handbook of Research for Education Communications and Technology* (2nd ed.) (1099 -1130). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Richey, R. C., & Klein, J. D. (2005). Developmental Research Methods: Creating Knowledge from Instructional Design and Development Practice. *Journal of Computing in Higher Education,* 16(2), 23 – 38.

Richey, R. C., & Klein, J. D. (2007). Design and development research. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers. Singapore, Pearson Education.

Ritchie, J., Lewis, J., Nicholls, C. M.& Ormston, R. (2007). Qualitative Research Practice: A Guide for Social Science Students and Researcher, London. Sage Publications.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346-362.

Rittle-Johnson, B. & Alibali, M. W. (1999). Conceptual and procedural knowledge of Mathematics: Does one lead to the other? *Journal of Educational Psychology*, 91(1). 175 - 189.

Ross, J.A.; McDougall, D. & Hogaboam-Gray, A. (2002). Research of reform in Mathematics education, 1993 – 2000.*Alberta Journal of Educational Research*, 48(2), 122 – 138.

Sadler, DR (2005). Interpretations of criteria-based assessment and grading in higher education, *Assessment and Evaluation in Higher Education*, 30(2), 175–194.

Schneider, M. & Stern, E. (2010a). The cognitive perspective on learning. Ten cornerstone findings. In Organisation for Economic Co-Operation and Development (OECD) (Ed). *The nature of learning: Using research to inspire practice* (pp. 69-90). Paris: OECD.

Schneider, M., & Stern, E. (2010b). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology*, 46, 178-192.

Schoenfeld, A. H. & Kilpatric, J. (2008) Towards a theory of proficienct in teaching mathematics in Wood, T. & Tirosh, D. Ed. *International handbook of mathematics teachers (2). Tools and Processes in Mathematics Teacher Education*.

Schoenfeld, A. H. (2007 a). "Issues and Tensions in the Assessment of Mathematical Proficiency" in Schoenfeld A.H. Eds. *Assessing Mathematical Proficiency,* 3 – 15 United Kingdom: Cambridge University Press.

Schoenfeld, A. H. (2007 b). "What is Mathematical Proficiency and How Can It Be Assessed?" in Schoenfeld A.H. Eds. *Assessing Mathematical Proficiency,* 59 – 73, United Kingdom: Cambridge University Press.

Schwartz, J.L. (1994.). The role of research in reforming Mathematics education: A different approach. *Mathematics thinking and problem solving*, New York: Routledge Falmer.

Scotland, J. (2012). *Exploring the Philosophy Underpinnings of Research: Relating Ontology and Epistemology to the Methodology and Methods of Scientific, Interpretative and Critical Research Paradigms*. Published on: http://dx.doi.org/10.5539/elt.v5n9p9 (Retrieved on 21 August 2014).

Seguin, R. (1993). *Curriculum Development and Implementation of teaching programmes: Methodology guide*. UNESCO publications.

Shanahan, T. (2013). Letting the text take center stage: How the Common Core State Standards will transform English Language Arts instruction. *American Educator*, 43, 4-11.

Shannon, A (1999). *Keeping Score*, National Research Council: Mathematical Sciences Education Board, Washington, D.C.: National Academy Press.

Skemp, R. R. (1979). Goals of Learning and Quality of Understanding. *Mathematics teaching*, 88, 44 – 49.

Smith, G.; Wood, L.; Coupland, M.; Stephenson, B.; Crawford, K. & Ball, G. (2010). *Constructing mathematical examinations to assess a range of knowledge and skills*. Published on http://dx.doi.org/10.1080/0020739960270109 (Retrieved on 20 August 2013).

Smith, E. E.; Shoben, E. J. & Rips, L. J. (1974). Structure and Process in the Semantic Decisions. *Psychological Review*, 81(3) 214 -241.

Star, J. R. & Stylianider, G. J. (2013). Procedural and conceptual knowledge: Exploring the gap between knowledge type and knowledge quality. *Canadian Journal of Science, Mathematics and Technology Education.* 13(2), 169-181.

Stiggins, R. (2007). Assessment through the eyes of a child. *Education Leadership: Educating the whole child*, 64(8), 22-26.

Stiggins, R. (2001). *Student Involved Classroom Assessment* (3rd Edition). Columbus, OH: Merrill Publishing.

The Charles A. Dana Center, (2012) at the University of Texas, *common core standards for Mathematics*.http://ccsstoolbox.agleming.com/pdf/CCSSI_Math%20Standards.pdf. (Retrieved on 15 August 2013).

United Nations Educational, Scientific and Cultural Organization (UNESCO) (1984). *Mathematics for all: Problems of cultural sensitivity and unequal distribution of Mathematics education and future perspective on mathematic teaching for the majority*: Report and papers presented at the 5th international congress on Mathematics education, Adelaide August 24-29, 1984.

Urdan, C. T (2010). *Statistics in Plain English*. Third Edition. New York, Routledge Taylor& Francis Group.

Van der Merwe, I. F. J. (1993). The Senior Secondary Curriculum in Namibia for IGCSE and HIGCSE. *In proceedings for National Conference on IGCSE and HIGCSE*. Swakopmund. MoEC.

Vaughn, S., Schumm, J. S. & Sinagub, J. (1996). *Focus group interviews in education and psychology*. Thousand Oaks, Sage.

Verenikina, I. (2010). Vygotsky in Twenty-First-Century research. In J. Herrington & B. Hunter (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 16-25). Chesapeake, VA: AACE.

Voskoglou, M. G. (2008). Problem-Solving in Mathematics Education: Recent Trends and Development. *Quaderni di Ricerca in Didattica (Scienze Mathematiche),* University of Palermo, 18, 22-28.

Webb, N. L. (1992). Assessment of students' knowledge of Mathematics: Steps towards a theory. In D.A. Grouws (Ed.), *Handbook of research on Mathematics teaching and learning* (pp. 661-683). New York: Macmillan.

Webb, N. L. (1999). Research monograph No. 18: Alignment of science and Mathematics standards and assessment in the four states. Washington D.C.: CCSSO.

Wiliam, D.; Lee, C.; Harrison, C. & Black, P. (2004). Teacher developing assessment for learning: impact on student achievement. *Assessment in Education*, 2(1), 49 – 65.

Wilson, L. D. (2007). "High stakes testing in Mathematics" in Lester, F. K Eds. *Second Handbook of Research on Mathematics Teaching and learning*, 1099 – 1110, United State of America: Information Age publisher.

Winter, G, (2000). A Comparative Discussion of the Notion of 'Validity' in Qualitative and Quantitative Research. The Qualitative Report, 4(3&4). Published on http://www.nova.edu/ssss/QR/QR4-3/winter.html) (Retrieved on 27 July 2016).

Wragg, E. C. (2001). *Assessment and learning in the secondary school,* New York: Routledge Falmer.

Wu, H. (1997). The Mathematics Education Reform: Why you should be concerned and what you can do. *The American Mathematical Monthly*, 104 (10), 946-954.

Yunus, Z.; Razak, N. A.; Khairani, A. Z. (2012); Application of Mathematics proficiency model in test development. Retrieved from http://www.recsam.edu.my on 20 August 2013.

APPENDICES

APPENDIX A

	NSSCO PROCESS STANDARDS
1	Organise, interpret and present information accurately in written, tabular, graphical and diagrammatic forms
2	Perform calculations by suitable methods
3	Use an electronic calculator
4	Understand systems of measurement in everyday use and make use of them in the solution of problems
5	Estimate, approximate and work to degrees of accuracy appropriate to the context
6	Use mathematical and other instruments to measure and to draw to an acceptable degree of accuracy
7	Interpret, transform and make appropriate use of mathematical statements expressed in words or symbols
8	Recognise and use spatial relationships in two and three dimensions, particularly in solving problems
9	Recall, apply and interpret mathematical knowledge in the context of everyday situations
10	Make logical deductions from given mathematical data
11	Recognise patterns and structures in a variety of situations, and form generalisations
12	Respond to a problem relating to a relatively unstructured situation by translating it into an appropriately structured form
13	Analyse a problem, select a suitable strategy and apply an appropriate technique to obtain its solution
14	Apply combinations of mathematical skills and techniques in problem solving
15	Set out mathematical work, including the solution of problems, in a logical and clear form using appropriate symbols and terminology.

APPENDIX B

NSSCO MATHEMATICS CONTENT STANDRDS	
SPECIFIC CONTENT OBJECTIVES	
NUMBERS AND OPERATIONS	
Numbers	
	Learners will be able to:
1	identify and use natural numbers, integers (positive, negative and zero), prime numbers, common factors and common multiples, national and irrational numbers and real numbers (e.g. π, 2)
Squares, square roots, cubes and cube roots	
1	calculate squares, square roots, cubes and cube roots of real numbers
Directed numbers	
1	use directed numbers in practical situations (e.g. temperature change, flood levels)
2	apply the four basic operations on directed numbers
Vulgar and decimal fractions and percentages	
1	use the language and notation of simple vulgar and decimal fractions and percentages in appropriate context
2	recognise equivalence and convert between these forms
3	calculate a given percentage of a quantity
4	express one quantity as a percentage of another
5	calculate percentage increase and decrease
6	carry out calculations involving reverse percentages, e.g. finding the cost price, given the selling price and the percentage profit
Ordering	
1	order quantities by magnitude and demonstrate familiarity with the symbols $=, \neq, >, <, \leq, \geq$
Standard form	
1	use the standard form $A \times 10^n$ where n is a positive or negative integer, and $1 \leq A \leq 10$
Four basic operations	
	apply the four basic operations for calculations with integers, decimal fractions and vulgar (and mixed) fractions, including correct order of operations and use of brackets
Estimation	
1	make estimates of numbers and quantities
2	give approximations to specified numbers of significant figures and decimal places
3	round off to reasonable accuracy in the context of the given problem
Limits of accuracy	
1	give appropriate upper and lower bounds for data given to specified accuracy (e.g. measured lengths)
2	obtain appropriate upper and lower bounds to solutions of simple problems (e.g. the calculation of the perimeter or the area of a rectangle) given data to specified accuracy
Ratio, proportion and rate	
1	apply the ideas and notation of ratio, proportion and rate to practical situations
2	divide a quantity in a given ratio
3	use scales in practical situations
4	calculate average speed
5	increase and decrease a quantity by a given ratio
6	express direct and inverse variation in algebraic terms and use this form of expression to find unknown quantities
Use of an electronic calculator	
1	use an electronic calculator efficiently
2	apply appropriate checks of accuracy
Money and finance	
1	solve problems involving money and convert from one currency to another
2	use given data to solve problems on personal and household finance involving earnings, simple interest, compound interest (knowledge of the formula is not required, can be provided in examinations),discount, profit and loss, tax and budgeting
3	extract data from tables and charts
MEASURES	
SI units of measures	
1	use SI units of mass, length, area, volume and capacity in practical situations
2	express quantities in larger or smaller units

3 calculate time in terms of the 24-hour and 12-hour clock (notation for time e.g. 14:00 or 2.00 pm) 4 read clocks, dials and timetables

MENSURATION

Perimeters, areas and volumes

1 carry out calculations involving
- the perimeter and area of a rectangle and triangle
- the circumference and area of a circle
- the perimeter and area of a parallelogram and a trapezium
- the volume of a cuboid, prism and cylinder
- the surface area of a cuboid and a cylinder

2 solve problems involving the arc length and sector area as fractions of the circumference and area of a circle, the surface area and volume of a sphere, pyramid and cone (given formulae for the sphere, pyramid and cone)

GEOMETRY

Geometrical terms and relationships

1 use and interpret the geometrical terms: point, line, parallel, intersecting, bearing, right angle, acute, obtuse and reflex angles, perpendicular, similarity, congruence

2 use and interpret vocabulary of triangles, quadrilaterals, circles, polygons and simple solid figures, including nets

3 use the relationships between areas of similar triangles, with corresponding results for similar figures and extension to volumes and surface areas of similar solids

Geometrical constructions

1 measure lines and angles
2 construct a triangle given the three sides using a straight edge and compasses only
3 construct other simple geometrical figures from given data using protractors and set squares as necessary
4 construct angle bisectors and perpendicular bisectors using straight edges and compasses only
5 read and make scale drawings

Symmetry

1 recognise line and rotational symmetry (including order of rotational symmetry) in two dimensions
2 recognise properties of triangles, quadrilaterals and circles directly related to their symmetries
3 use the following symmetry properties of circles:
- equal chords are equidistant from the centre
- the perpendicular bisector of a chord passes through the centre of a circle
- tangents from an external point are equal in length

Angle properties

calculate unknown angles using the following geometrical properties (reasons may be required but no formal proofs):
1 angles at a point
2 angles on a straight line and intersecting straight lines
3 angles formed within parallel lines
4 angle properties of triangles and quadrilaterals
5 angle properties of regular polygons
6 angle in a semi-circle
7 angle between tangent and radius
8 angle properties of irregular polygons
9 angle at the centre of a circle is twice the angle at the circumference
10 angles in the same segment are equal
11 angles in opposite segments are supplementary

Locus

use the following loci and the method of intersecting loci for sets of points in two dimensions:
1 which are at a given distance from a given point
2 which are at a given distance from a given straight line
3 which are equidistant from two given points
4 which are equidistant from two given intersecting straight lines

ALGEBRA

Algebraic representation and formulae

	Learners will be able to:
2	use letters to express generalised numbers
3	express basic processes algebraically
4	substitute numbers for words and letters in formulae
5	transform simple formulae
6	construct linear equations from given situations
7	*transform formulae involving roots, powers, fractions and factorization*
	Algebraic manipulation
	Learners will be able to:
1	multiply a monomial by a polynomial
2	use brackets and extract common factors
3	*expand products of algebraic expressions*
4	*factorise where possible expressions of the form* $ax+ay;\ \ ax+bx+kay;+kby;\ \ a^2x^2-b^2y^2;$ $a^2+2ab+b^2;\ \ ax^2+bx+c$
5	*manipulate algebraic fractions, e.g.* $\frac{x}{3}+\frac{x-4}{2},\ \frac{2x}{3}-\frac{2}{x-3},\ \frac{3a}{4}\times\frac{5ab}{3},\ \frac{1}{x-2}-\frac{2}{x-3}$
6	*factorise and simplify expressions such as* $\frac{x^2-2x}{x^2-5x+6}$
7	*write quadratic expressions in the form* $a(x+p)^2+q$
	Polynomials
	Learners will be able to:
1	*carry out operations of addition, subtraction and multiplication of polynomials*
2	*carry out division of a polynomial by a binomial expression, and identify the quotient and the remainder*
	Equations and inequalities
	Learners will be able to:
1	solve simple linear equations in one unknown
2	solve simultaneous linear equations in two unknowns
3	solve quadratic equations given in factorized form, e.g. $(x-2)(x+3)=0$
4	*solve quadratic equations by factorisation, by completing the square and by use of the formula* $\frac{-b\pm\sqrt{b^2-4ac}}{2a}$
5	*solve equations with fractions, e.g.* $\frac{x-2}{x-1}=1-\frac{x+8}{x+14}$
6	*solve simultaneous equations, one linear and one quadratic*
7	*solve simple linear inequalities*
	Sequences
	Learners will be able to:
1	recognise patterns in sequences
2	find the n^{th} term of a linear sequence
3	*distinguish between arithmetic and geometric sequences* *use the formulae for the n^{th} term and the sum of the first n terms to solve problems involving arithmetic and geometric progressions*
	Indices
	Learners will be able to:
1	use and interpret positive, negative and zero indices
2	*apply the laws of indices correctly*
3	*use and interpret fractional indices, e.g.* $2^{\frac{1}{2}}=\sqrt{2}$
4	*simplify expressions involving indices, e.g.* $\frac{3^{x-y}\times3^{x+y}}{3^{2x}}$

5	*solve simple exponential equations, e.g.* $2^x = \frac{1}{4}$
	Logarithms
	Learners will be able to:
1	*use the laws of logarithms to simplify expressions*
2	*solve equations of the form* $a^x = b$ *using logarithms*
	Graphs in practical situations
	Learners will be able to:
1	demonstrate familiarity with the Cartesian coordinates in two dimensions
2	interpret and use graphs in practical situations, including travel graphs and conversion graphs
3	draw graphs from given data
4	*apply the idea of rate of change to easy kinematics involving distance-time and speed-time graphs, acceleration and deceleration*
5	*calculate distance travelled as area under a linear speed-time graph*
	Linear programming
	Learners will be able to:
1	*represent inequalities graphically and use this representation in the solution of simple linear programming problems (the convention of using broken lines for strict inequalities and shading unwanted regions will be expected)*
	Functions
	Learners will be able to:
1	*use function notation, e.g.* $f(x) = 3x - 5$; $f : x \mapsto 3x - 5$ *to describe simple functions, and the notation* $f^{-1}(x)$ *to describe their inverses*
2	*form composite functions as defined by* $gf(x) = g(f(x))$
	Graphs of functions
	Learners will be able to:
1	construct tables of values for functions of the form $y = ax + b$, $y = \pm x^2 + ax + b$, $y = a/x$, $(x \neq 0)$ where *a* and *b* are integral constants
2	draw and interpret such graphs
3	find the gradient of a straight line graph and *determine the equation of a straight line in the form* $y = mx + c$
4	solve linear and quadratic equations approximately by graphical methods
5	*construct tables of values and draw graphs for functions of the form* $y = ax^n$ *where a is a rational constant and n = –2, –1, 0, 1, 2,3 and simple sums of not more than three of these and for functions of the form* $y = a^x$ *where a is a positive integer*
6	*estimate gradients of curves by drawing tangents*
7	*solve associated equations approximately by graphical methods*
	GRAPHS AND FUNCTIONS
	Graphs in practical situations
1	demonstrate familiarity with the Cartesian coordinates in two dimensions
2	interpret and use graphs in practical situations, including travel graphs and conversion graphs
3	draw graphs from given data
4	apply the idea of rate of change to easy kinematics involving distance-time and speed-time graphs, acceleration and deceleration
5	calculate distance travelled as area under a linear speed-time graph
	Linear programming
1	represent inequalities graphically and use this representation in the solution of simple linear programming problems (the convention of using broken lines for strict inequalities and shading unwanted regions will be expected)
	Functions
1	use function notation, e.g. f (x) = 3x − 5 ; f : x a3x − 5 to describe simple functions, and the notation f – 1(x) to describe theirinverses

2 form composite functions as defined by gf (x) = g(f (x))

Graphs of functions

1 construct tables of values for functions of the formy = ax + b, y = ±x2 + ax + b , y = a x , (x ≠ 0) where a b are integral constants

2 draw and interpret such graphs

3 find the gradient of a straight line graph and determine the equation of a straight line in the form y = mx + c

4 solve linear and quadratic equations approximately by graphical methods

5 construct tables of values and draw graphs for functions of the form y = axn where a is a rational constant and n = –2, –1, 0, 1, 2,3 and simple sums of not more than three of these and for functions of the form y = ax where a is a positive integer

6 estimate gradients of curves by drawing tangents

7 solve associated equations approximately by graphical methods

COORDINATE GEOMETRY

Graphs and Cartesian coordinates

1 *calculate the distance between two points given in coordinate form, the gradient of the line-segment joining them, and the coordinates of their midpoint*

2 *find the equation of a straight line given sufficient information (e.g. the coordinates of two points on it or one point on it and its gradient)*

3 *interpret and use equations of the form* $ax + by + c = 0$ *, including knowledge of the relationships involving gradients of parallel and perpendicular lines*

4 *apply coordinate geometry to quadrilaterals*

TRIGONOMETRY

Bearings and the trigonometrical ratios

1 interpret and use three-figure bearings measured clockwise from the north (i.e. 000 to 360)

2 apply Pythagoras' theorem and the sine, cosine and tangent ratios for acute angles to the calculation of a side or of an angle of a right-angled triangle (angles will be quoted in, and answers required in, degrees)

3 solve trigonometrical problems in two dimensions involving angles of elevation and depression

4 *extend trigonometrical ratios to angles between 90o and 360o*

5 *solve problems using sine and cosine rules for any triangle and the formula: area of triangle* $ab\sin C = \frac{1}{2}$

6 *solve simple trigonometrical problems in three dimensions including angle between a line and a plane*

VECTORS IN TWO DIMENSIONS AND TRANSFORMATIONS

Vectors

1 describe a translation using a vector represented by **a**, add and subtract vectors and multiply a vector by a scalar

2 calculate the magnitude of a vector x as $x^2 + y^2$ (Vectors will be printed as AB or **a** and their magnitudes as AB or |a|. In their answers to questions candidates are expected to indicate a in some definite way, e.g. by an arrow or by underlining, thus AB or a)

3 represent vectors by directed line segments

4 use the sum and difference of two vectors to express given vectors in terms of two coplanar vectors

5 use position vectors

Transformations

1 reflect simple plane figures in horizontal and vertical lines

2 rotate simple plane figures about the origin, vertices and midpoints of edges of the figures, through multiples of 90°

3 construct given translations and enlargements of simple plane figures

4 recognise and describe reflections, rotations, translations and enlargements

5 use the following transformations of the plane: reflection, rotation, translation and enlargement and their combinations

6 identify and give precise descriptions of transformations connecting given figures

STATISTICS AND PROBABILITY

Statistics

1 collect, classify and tabulate statistical data

2 read, interpret and draw simple inferences from tables and statistical diagrams

3 construct and use bar charts, pie charts, pictograms, simple frequency distributions and histograms with equal intervals

4 calculate the range, mean, median and mode for individual and discrete data and distinguish between the purposes for which they are used

5 *construct and read histograms with equal and unequal intervals (areas proportional to frequencies and vertical axis labelled 'frequency density')*

6 *construct and use cumulative frequency diagrams*

7 *estimate and interpret the median, percentiles, quartiles and interquartile range*

8 *calculate an estimate of the mean for grouped and continuous data*

9 *identify the modal class from a grouped frequency distribution*

Probability

1 calculate the probability of a single event as either a fraction or a decimal (not ratio)

2 understand and use the probability scale from 0 to 1

3 understand that (the probability of an event occurring) = 1 - (the probability of the event not occurring)

4 *understand probability in practice, e.g. relative frequency*

5 *calculate the probability of simple combined events, using possibility diagrams and tree diagrams where appropriate (in possibility diagrams outcomes will be represented by points on a grid and in tree diagrams outcomes will be written at the end of branches and probabilities by the side of the branches)*

APPINDIX C

KEY GRADE DESCRIPTORS
At Grade A the learner is expected to:
- Express any number to 1, 2 or 3 significant figures. Relate a percentage change to a multiplying factor and vice versa, e.g. multiplication by 1.03 results in a 3% increase.
- Relate scale factors to situations in both two and three dimensions. Calculate actual lengths, areas and volumes from scale models. Carry out calculations involving the use of right-angled triangles as part of work in three dimensions.
- Add, subtract, multiply and divide algebraic fractions. Manipulate algebraic equations – linear, simultaneous and quadratic. Use positive, negative and fractional indices in both numerical and algebraic work. Write down algebraic formulae and equations from a description of a situation.
- Process data, discriminating between necessary and redundant information. Make quantitative and qualitative deductions from distance-time and speed-time graphs.
- Make clear, concise and accurate mathematical statements, demonstrating ease and confidence in the use of symbolic forms and accuracy in algebraic or arithmetic manipulation.
At Grade C the learner is expected to:
- Perform calculations involving several operations. Calculate percentage change. Use a calculator fluently. Give a reasonable approximation to a calculation involving the four rules. Do compound interest calculations. Use and understand the standard form of a number. Apply ratio and proportion concepts.
- Use area and volume units. Find the volume and surface area of a prism and a cylinder. Use a scale diagram to solve a two-dimensional problem. Calculate the length of the third side of a right-angled triangle. Apply trigonometry to solve triangles. Calculate angles in geometrical figures. Use constructions to solve a locus problem.
- Recognise, and in simple cases formulate, rules for generating a pattern or sequence. Solve simple simultaneous linear equations in two unknowns. Solve quadratic equations. Transform simple formulae. Substitute numbers in complicated formulae and evaluate the remaining term.
- Construct a pie-chart from simple data. Plot and interpret graphs, including graphs of linear and simple quadratic functions.
At Grade E a learner is expected to:
- Apply the four rules of number to positive and negative integers and decimal and vulgar fractions. Identify cube and prime numbers. Calculate a quantity as a percentage of another quantity. Convert between currencies. Use natural number indices. Understand and apply simple proportion. Find average speed.
- Find the perimeter and area of rectilinear shapes. Find the volume and surface area of a cuboid. Calculate angles in triangles and quadrilaterals. Use a map and scales to find distances and bearings. Construct triangles and nets of cuboids and triangular prisms.

Recognise reflections, rotations and enlargements.
- Continue a sequence progressing into negative numbers. Write a simple expression using two variables. Simplify a linear expression using two variables. Substitute numbers into formulae and evaluate the remaining term. Use brackets and extract common factors from algebraic expressions. Solve simple linear equations.
- Use and understand the concepts mean, median and mode. Interpret a simple pie chart. Draw and interpret straight line graphs, including conversion and travel graphs.
At Grade G a learner is expected to:
- Perform the four rules on positive integers and decimal fractions using a calculator where necessary. Convert simple fractions to decimals and percentages. Identify square numbers. Give factors of numbers. Order a list of decimal numbers. Calculate simple percentages. Perform simple money calculations. Work out the time of a simple journey. Do simple estimations. Identify the smallest integer in a list containing negative numbers.
- Recognise and name simple plane figures and common solid shapes. Measure a line or an angle accurately. Find the perimeter of rectilinear shapes. Find the area of a rectangle and the volume of a cuboid. Draw the line of symmetry in a shape. Draw the reflection of a shape about a given vertical or horizontal mirror line.
- Continue a sequence in numbers or diagrams. Substitute numbers into simple formulae.
- Interpret tables. Draw and interpret a bar chart. Calculate the mean of a small set of discrete data. Identify the mode from a set of discrete data or a bar chart. Plot or give coordinates of points in the first quadrant.

***Purple print is algebra

APPENDIX D

KEY GRADE DESCRIPTORS - ALGEBRA
At Grade A the learner is expected to:
• Add, subtract, multiply and divide algebraic fractions. • Manipulate algebraic equations – linear, simultaneous and quadratic. • Use positive, negative and fractional indices in both numerical and algebraic work. • Write down algebraic formulae and equations from a description of a situation. • Generate strategies and use it to solve algebraic problems • Justify a strategy used in solving an algebraic problem and reflect on a solution to an algebraic problem • Explain whether or not a mathematical problem is truly a problem • Explain how a certain solution was reached and find other ways of getting the same answer.
At Grade C the learner is expected to:
• Transform word/novice problems into algebraic equations and solve them • Describe mathematical application or strategy. • Recognise, and in simple cases formulate, rules for generating a pattern or sequence. • Solve simple simultaneous linear equations in two unknowns. Solve quadratic equations. • Transform simple formulae. Substitute numbers in complicated formulae and evaluate the remaining term.
At Grade E a learner is expected to:
• Define mathematics concepts • Continue a sequence progressing into negative numbers. • Write a simple expression using two variables. • Simplify a linear expression using two variables. • Substitute numbers into formulae and evaluate the remaining term. • Use brackets and extract common factors from algebraic expressions. • Solve simple linear equations.
At Grade G a learner is expected to:
• Continue a sequence in numbers or diagrams. • Substitute numbers into simple formulae.

APENDIX E

Table 1 NSSCO STANDARDS TRANSATION TABLE

	Current Standard (Process)	**Interpretation of the current Process standards: Discussion with possible examples**	**Link to <u>Proficiency Standards</u> **Knowledge based, strategies and metacognition****	**Resulting process standard (Expanded Process Standard)**
1(a)	<u>*Organise*</u>, information accurately in written, tabular, graphical and diagrammatic forms; Production – for thinking? (1) is about representation	<u>*Given statistical information or a list of numbers - rearrange or position them in a systematics order*</u> either in a written, tabular / graphical or diagrammatic form. Example given a list of number, order them by magnitude or statistical information	*Presenting or representing solutions in a tabular form, graphs and other diagrams*** ***Strategy *****	• Present or represent solutions in a tabular form, graphs, diagrams and any other organised form and *write down conclusions drawn from the information in such presentations*
1 (b)	<u>*interpret*</u> information accurately in written, tabular, graphical and diagrammatic forms; Interpretation?	<u>*Assign a meaning to or make sense of*</u> information in written tabular, graphical and diagrammatic forms. Example annual rainfall distribution by month.	Use knowledge of mathematics concepts to come up with mathematical representation (tables, graphs and other diagrams like Venn diagrams etc. Translating to *Explain the meaning concept, fact and procedures* *And use it to Present or represent solutions in a tabular form, graphs and other diagrams* ***** knowledge based and Strategy*****	
1 (c)	<u>*present*</u> information accurately in written, tabular, graphical and diagrammatic forms; Production – for communicating?	<u>*show or display*</u> of information accurately in written, tabular, graphical and diagrammatic forms	Use more specified strategies and procedures to produced solutions to a problem __ As a way to *Present or represent solutions in a tabular form, graphs and other diagrams *****strategy*****	
2	<u>*perform*</u> calculations by suitable methods	use a suitable method <u>*to do or to carry out*</u> calculation	Use a known procedure to a known solution to solve a similar but problem - exploiting solutions to similar problems	• Identify and use basic computational facts • Use a given strategy to solve a mathematical problem.

			Decide/generate a strategy (calculation procedure) *and use it* ***strategy****	• Decide/Generate a strategy and use it in solving a problem. • Describe a strategy used and justify its use in solving a mathematical problem • Reflecting on the solution or method used to solve a problem and suggest other ways of solving it.
3	use an electronic calculator	***to be incorporated in all other strands perhaps?)***		
4	*understand* systems of measurement in everyday use and make use of them in the solution **or solving a** of problems;	*Know and comprehend the nature or meaning of* systems of measurement in everyday use and make use of them in the solution of problems **(how is it measurable? Through description of how the instruments are used)**	*Use knowledge of mathematics concepts to comprehend* (through description or definition) *and to apply semantic knowledge in solving problems in context* ** ***Knowledge based****	• Describe systems of measurement in everyday use and make use of them in the solution or solving a of problems
5	*estimate, approximate* and *work to degrees of accuracy* appropriate to the context; ***it needs not to be estimate, approximate rather estimate or approximate***	*estimating or approximating or taking an educated/close guess usually with some though or calculation with regard to worth, size, quality or quantity*	*Use knowledge of mathematics concepts to comprehend* (through description or definition) *and to apply semantic knowledge in solving problems in context* - example estimating half of a quantity, they should know what half is ***knowledge based****. *But at the same time this involve 'Generating a strategy and use it procedurally'* ***strategies **** either mentally or written out that involves procedures of rounding off and dividing.	• Estimate and approximate and work to degrees of accuracy appropriate to the context.
6	use mathematical and other instruments to *measure* and to *draw* to an acceptable degree of accuracy;	use mathematical and other instruments to *find out the size, amount or degree of something by comparing it with a standard and to produce a picture or a diagram on paper* to an acceptable degree of accuracy	*Use a given strategy* (measuring) *to solve a mathematical problem Use knowledge of mathematics concepts and procedures* - to read measurements using a measuring instrument- *in mathematics problem solving; be able to handle instruments to represent different mathematical*	• Use mathematical instruments and other instruments to measure and to draw to an acceptable degree of accuracy

			concept diagrammatically *****strategy*****	
7(a)	*interpret* mathematical statements expressed in words or symbols;	*assign a meaning to or make sense of* mathematical statements expressed in words or symbols *Use a mathematical formula to express to rewrite a mathematical statement and vice versa or retell or write in your own words*	*Use knowledge of mathematics concepts to comprehend and/ come up with definition*** ***knowledge based*****	• Demonstrate the understanding of mathematical statements expressed in words or symbols. (through transformation, representation or explanation) • *Visualise through explanation the mathematics behind and mathematical representation*. • Interpret what the mathematical problem mean.
7(b)	*Transform and make appropriate use of* mathematical statements expressed in words or symbols;	*Change by interpreting* a mathematical statements expressed in words to symbols or vice versa.	*Use knowledge of mathematics concepts to comprehend and/ come up meaning/definition of facts *** ***knowledge based*****	
8 (a)	*recognise* spatial relationships in two and three dimensions,	*know something (from having come across it before); be aware of* spatial relationships in two and three dimensions	*Use knowledge of mathematics concepts to come up with facts or definition *** ***knowledge based***** *and using what you know effectively* ******metacognition*****	• Identify and use spatial relationships in two and three dimensions, particularly in solving problems. • Make use of similarities or symmetry and analogies to solve spatial and other mathematical problems.
8 (b)	use spatial relationships in two and three dimensions, particularly in *solving* problems;	use spatial relationships in two and three dimensions, particularly in *finding ways/**strategies** of getting the answer to the problems*	*Using the knowledge of spatial relationship – by comparing similarities (making analogies) to device strategies to solve problems.*	
9 (a)	*recall* mathematical knowledge in the context of everyday situations;	*retrieve from memory or remembering* mathematical knowledge in the context of everyday situations	*Use knowledge of mathematics concepts to come up with facts or definition *** ***knowledge based*****	• Define or describe mathematical concepts, facts and relationships • Recall through definition or representation, apply and interpret mathematical knowledge in the context of everyday situations
9 (b)	*apply* mathematical knowledge in the context of everyday situations;	*put* the mathematical knowledge *to practical use* in context	*Use knowledge of mathematics concepts and procedures and apply them in context* ******knowledge based*****	
9 (c)	*interpret* mathematical knowledge in the context of everyday situations;	*assign a meaning to or make sense of* mathematical knowledge in context	*Use knowledge of mathematics concepts to come up with facts or definition *** ***knowledge based*****	
10	*make logical deductions* from given mathematical data;	*use the mathematical knowledge to make sense of* mathematical data	Interpret and explain relationship in the mathematical data.	• Interpret and explain relationship in the mathematical data.

11	*recognise* patterns and structures in a variety of situations, and form generalisations;	*know something from having come across it before*; be aware of a fact or problem in relation to patterns and structures in a variety of situations, and form generalisations	*Use knowledge of mathematics concepts to come up with facts or definition or identification* ***knowledge based*** *and using what you know effectively* in different mathematical situation ***metacognition***	• Recognise and identify patterns and structures in a variety of situations, and form generalisations.
12	*respond* to a problem relating to a relatively unstructured situation by *translating* it into an appropriately structured form;	*act in response to* or *answer* problems related to a relatively unstructured situation by translating (*assign a meaning to or make sense of)* in into an appropriate structured form	*Use knowledge of mathematics concepts to come up with facts or definition****knowledge based*** *and using what you know effectively* in different mathematical situation ***metacognition***	• Transform novice/context arithmetic problems by interpreting it into an appropriately structured form.
13 (a)	*analyse* a problem	*examine in details in order to discover meaning or essential features* of a given problem	*Use knowledge of mathematics concepts to come up with facts or definition* ***knowledge based***	• Analyse a problem, select a suitable strategy and apply an appropriate technique to obtain its solution. • Explain how a certain solution was reached and find other ways that can be used to get the same solution • Explain why a certain answer is not a solution to a problem.
13 (b)	*select* a suitable strategy to obtain a solution to a problem	*choose one strategy in preference to another* to obtain a solution to a problem	*Decide/generate a strategy and use it* ***OR*** *Be able to formulate, represent and solve mathematical problems* ***strategy***	
13 (c)	*apply* an appropriate technique or strategy to obtain a solution to a problem	*use* the appropriate strategy/methods chosen in order to obtain a solution to a problem	*Decide/generate a strategy and use it* ***OR*** *be able to formulate, represent a chosen strategies in solve mathematical problems* ***strategy***	
14	*apply* combinations of mathematical skills and techniques in problem solving;	*put* a combinations of mathematical skills and techniques *in practice* in problem solving;	*Decide/generate a strategy and use it* ***OR*** *be able to formulate, represent skills and techniques used in solve mathematical problems* ***strategy***	• Use a combinations of mathematical skills and techniques in problem solving;
15 Not clear	***set out*** **??** mathematical work, including the solution of problems, in a logical and clear form using appropriate symbols and terminology			

15 rewritten	***present*** mathematical working; solution to problems, in a logical and clear form using appropriate symbols and terminology	*Demonstrate or show* mathematical working or solution to problems in a logical and clear form using appropriate symbols and terminology.	*Decide/generate a strategy and use it **OR** be able to formulate, represent and solve mathematical problems* ****strategy****	• Present mathematical working; solution to problems, in a logical and clear form using appropriate symbols and terminology

Table 2. Gaps not related to the current NSSCO process standards

	Other gaps (Proficiency standards not related to the current NSSCO standards)	**Description of the gap and related proficiency category**	**Resulting process standard**
1	Pose questions	The ability of the candidate to formulate or reformulate a problem ****strategy****.	Restating or reformulating problems Generate a relevant and appropriate easier related problem, and use its solution to solve the related problem. Or Explaining whether or not a mathematical problem to be solved exists or explain whether or not a mathematical problem is truly a problem.
2	Identify irrelevancies	The ability to identify irrelevant issues either in the problem or in a given solution to a problem ****strategy****.	
3	Model the problem	Present a simplified description of a problem in order to assist the calculation or in solving the problem ****strategy and metacognition****.	Restate the problem in simplified words or context in order to solve it.
4	Identify obstacles	The learners' ability to identify an obstacles to getting the correct solution through a given procedure.	Given the working to a solution of a specific problem that might have reached a dead end – reflect on what has been done wrong and suggest ways on how to fix it
5	Explain or Justify solutions or strategies used and produce alternative strategies, including describing multiple strategies	Make a fact or a concept or strategy used comprehensible OR communicate plainly the strategy used OR provide adequate grounds for the solution obtained.	Explain how he or she reached a certain solution **knowledge base** and find other ways that can get you the same solution **strategy** and the whole process can be seen as (Reflect on what is being done and change the strategies in the course of solving the problem) **metacognition** Explain how a certain solution was reached and find other ways that can get you the same solution **metacognition**
6	Solve backwards	(could be there but not standing out) – Given the final steps of the solution and work back to the initial steps,	Working backwards as a strategy to reformulate, represent and solve mathematical problems **strategy**
7	Evaluate quality of solution	Assess or apprise the valuable features of the solution given – to be given in a form of description or explanation	Explain or describe the valuable features of certain solution and **metacognition**

KEY

Proficiency Category

- ***Knowledge Based*** **-** the understanding of concepts and procedural skills
- ***Strategies*** **-** the usage of strategies in mathematics problem solving; be able to formulate, represent and solve mathematical problems
- ***Metacognition*** **-** using what you know effectively and monitoring one's own thinking when involved in processes of reasoning, communications and making connections

The process of analysis:

- Table 1, column 2 and 3: *The content in italic and underlined* are action verbs that shows what the learner is expected to do.
- Table 1, column 4: *The content in italic* create a direct link to the proficiency standards extracted from the literature.
- Table 1, column 5: *The content in italic* reflects the gaps between the NSSCO standards and the proficiency standards that are directly related to the NSSCO standard in the specific row.
- Tabulate the assessment objectives and discuss to clarify what each one of the mean and describe these in terms of mathematical proficiency to find a link to proficiency strands that I am working with.
- In this process, in some cases, the objectives which seemed to comprise too many competencies were first broken. This is done in the in order to first clarify the objective and see a clear analytical separation and link to proficiency.
- In order for them to remain broad, as broad goal (assessment objectives) they need to be reconstructed back together
- Find a correlation between the broad standards with more specific standards
- The other gaps in Table 2 were identified by comparing the existing NSSCO Process standards and Mathematical proficiency standards highlighted by different authors on mathematical proficiency and problem solving.

APPENDIX F

EXPANDED PROCESS STANDARDS

1	Identify and use basic computational facts (+, -, × and ÷)
2	Present or represent solutions in a tabular form, graphs, diagrams and any other organised form.
3	*Define or describe mathematical concepts and relationships*
4	Use a given strategy to solve a mathematical problem.
5	Use mathematical instruments and other instruments to measure and to draw to an acceptable degree of accuracy;
6	Make use of similarities or symmetry and analogies to solve spatial and other mathematical problems.
7	Estimate and approximate and work to degrees of accuracy appropriate to the context.
8	Describe systems of measurement in everyday use and make use of them in the solution or solving a of problems
9	*Generate a strategy and use it in solving a problem*
10	*Describe a strategy use and justify its use in solving a mathematical problem*
11	*Reflecting on the solution or method used to solve a problem and suggest other ways of solving it.*
12	Demonstrate the understanding of mathematical statements expressed in words or symbols. (through transformation, representation or explanation)
13	*Interpret what the mathematical problem mean and restate it in a simpler way.*
14	Visualise through explanation the mathematics behind a mathematical representation.
15	Identify and use spatial relationships in two and three dimensions, particularly in solving problems.
16	*Recall through definition or representation, apply and interpret mathematical knowledge in the context of everyday situations*
17	*Interpret and explain relationship in the mathematical data.*
18	Recognise and identify patterns and structures in a variety of situations, and form generalisations.
19	Transform, simplify and workout abstract and novice/context arithmetic problems by interpreting it into an appropriately structured form.
20	*Analyse a problem, select a suitable strategy and apply an appropriate technique to obtain its solution.*
21	*Explain how a certain solution was reached and find other ways that can be used to get the same solution.*
22	*Given the working to a solution of a specific problem that might have reached a dead end – reflect on what has been done wrong and suggest ways on how to fix it*
23	*Explain why a certain answer is not a solution to a problem.*
24	Use a combinations of mathematical skills and techniques in problem solving;

25	Present mathematical working; solution to problems, in a logical and clear form using appropriate symbols and terminology
26	*Explaining whether or not a mathematical problem to be solved exists or explain whether or not a mathematical problem is truly a problem.*

APPENDIX G

Complexity scoring rubric, the complexity is observed through the creation of links and simplicity through recall

	Complexity Scale				Length dimension
Proficiency category	None (Score 0)	Level 1 – Score 1 (low) knowledge of atoms	Level 2 - Score 2 (Medium) Creating links between atoms	Level 3 – Score 3 (High) Creating links between links to form a coherent structure	Adding to the strategy score by: • 0 for short • 1 for long
Knowledge Base	Nothing, is required	Low: Recall of atoms (known facts, concepts, definitions, notations and relationships.	involving one or more links between two or more atoms	High: Links between links. Creating links between links to form a coherently organized, integrated structure of linked atoms	
Strategies	No strategy is required	• Implementing an atomic strategy (based on the application of a single memorized principle/process) **Or** • Provide a solution without requiring to show a strategy used **Or** • Provide a solution through a simple representation of a diagram	• Implement a composite strategy showing how it can be used and provide a solution to a problem (showing working). It could be through representations like a complex diagram **Or** • Clearly describe a strategy, with some logical links, used to obtain a given solution • Rectify the used strategy to give the correct solution	• Develop and implement a strategy showing how it can be used and provide a solution to the problem • Generate an argument about a composite strategy and giving reasons to justify why the strategy is used to reach a solution or conclusions **Or** • Provide mathematical arguments for or against a solution or a solution strategy and provide justification/ an alternative strategy to obtain an alternative solution. **Or** • Provide an explanation why a given problem is not a problem using mathematical facts to give a full	**Short (S + 0)** Inherently a one-step calculation eg. $4x + 3 = 27$ $x = (27 - 3)/4$ $x = 6$ **Long (S + 1)** Inherently a one two or more steps calculation eg. $\frac{3x(x-2)}{(x+5)(x-2)}$ $= \frac{3x}{x+5}$

				justification or reasoning **Or** • Provide reasons why the used strategy is wrong and rectify the used strategy to give the correct solution	
Metacognition (efficiency)	• No monitoring and self-regulation	• Analysis of a process: identifying which is the better solution. • Problem solving: Choosing a strategy Giving up and trying a different strategy	• Analysis of a process: identify which is the better solution and given some reasons • Problem solving: choosing between two strategies with a reason	• Analysis of a process: identify which is the better solution and provide a coherent argument. • Problem solving: choosing between two or more strategies with process based reasoning	

Terminologies:

1. Knowledge base atoms = basic mathematical fats and concepts, definition, notation or symbols
2. Atomic strategy involves only one set of complementary basic operations eg. (+ and –) or (× and ÷), any combination thus is described as composite eg. (+ and × and or ÷)
3. Links = basic mathematical relationships

APPENDIX H

Markers Comments Summary – **TRIAL ROUND AND ROUND 1 (2013)**

General

The old NSSCO seem to only demand the correct answer from the learners, whereas the new criteria looks at three different categories (**Ma1**).

Learners are losing marks through the new criteria because the original allocation did not show that they are required to do more (**Ma2**). Weaker learners are the ones who do not show working – indicating the dependence on calculators and perhaps a lack of understanding required to work through the questions.

This lead to the questions that; are the weak because they do not show working? or Do they not show working because they are weak?

Allocation of the metacognition mark is very difficult – being an abstract concept it needs questions designed for it. **Example redoing their work or working backwards to recheck their answers would clearly be an M mark**. **(Ma2)** – **Ma2** suggest a specific number of M marks be available for a paper to be allocated. – The M mark could also be deducted for ridiculous answers like negative lengths or ages or a person being 6.5 metres tall etc.

Ma3: The current question although they cater for all three categories in the new criteria, they do not allow objective assessment of knowledge base and metacognition categories. The whole question paper was silent on metacognition and the knowledge base was awarded on assumption that since the learner has given the solution, she or he has knowledge on that content. The new rubric (criteria) benefit the learners. It was observed that almost every question required (previous knowledge) and therefore it was awarded a KB mark. It would have been better if questions were put in such a way that learners show that knowledge instead of assuming it (from their solution to the strategy problem).

Metacognition mark should not be awarded for just a choice of a method, every question might have two or three methods to solve it, maybe only if the learner uses a strategy that have an advantage of the other (methods) perhaps M can be given.

Metacognition just like knowledge should be asked specifically in a question for it to warrant a mark. Questions must be put in a language will ask (allow) learners to do reflection.

Metacognition must best consist of the post evaluation or reflection on the strategy. Questions like justify, hence or using your results must be used to test metacognition.

Extended (P2 and P4) must test more metacognition since the level is higher. Specific question must be set to test specific categories.

The new rubric is beneficial and the level of benefit increases with the old grades.

However the rubric would have worked better if it was applied on a question that was set on the same framework. To enable a proper comparison between the performance of the learners in relation to each framework. End of **Ma3.**

Ma4: The new developed rubric is completely relevant for paper 4, it helps discourage mechanical calculator work and credit alternative thinking and inventions. However the paper set on the new framework might need more time for the learners to respond to as compared to the original.

The C and E grades seem to have less marks added to them using the new developed criteria as compared to the A grade, which in some cases learners gained up to 20 more marks than the 5 – 8 marks gained at C and E grades. It shows that the new developed criteria discriminates (more amongst the grades) and fairer because stronger learners stands (a chance) to gain more marks than weaker candidates.

Suggestions on adaptations

The current NSSCO papers, (the learners) might need more time to be able to elaborate or expand on what there are doing (when responding to the question). Some questions can be solved using two different methods, which could deserve different amount of marks. He gave an example of question 8 (c) in paper 4 – which can be solved either using the reminder theorem (he calls it synthetic method) or long division. The method should therefore be awarded different marks as part of the recognition of metacognition.

In some questions the knowledge base and the metacognition mark is given by assumption if a candidate gives a correct answer despite them having not shown their strategy they are awarded knowledge 1 and strategy 1. He gave the example of question 8(b) paper 4.

Therefore each question should be designed to clearly be able to award the KB, S and M marks without assumptions.

Ma5: Learners with lower grades in general did not show their working on how they obtained their answer. So they lose their s marks and it is difficult to assign M marks.

Examiners should set up questions that allow learners to show their working. Learners should be asked to show their workings and to reflect on their answers in order to gain the metacognition marks. Learners should also be exposed to different strategies to be able to give their answers using different strategies.

Ma6: The new marks changed with more or less the same value within different grades (C, E and G). All three categories were covered within the question paper, with only learners who obtained a C on NSSCO could score the metacognition marks but those at G, only a few manage to obtain metacognition marks. Most learners did not show their working which made it difficult with the award of the strategy score.

Strict measure must be in place for the learners to show their working. Without showing their working. She suggest that only the Knowledge Base mark must be awarded if the answer is correct and penalise on Strategy marks. For the accuracy of the final answer, learners must be penalised only once because the majority of the learners has a correct knowledge and show by rounding to less than the accuracy demanded in the question paper. To avoid too much

penalising for the same thing learners can be penalised on one question. The marks for the question must be depended on Strategy and Knowledge needed to be shown, example question 5 on standard form, some learners has shown having a knowledge of writing numbers in standard form $A \times 10^n$ but failed to find the correct value of A and n. This question on NSSCO question paper is supposed to have 2 marks.

For question 11(a), it could be better if learners were asked to write ¼ litre to millilitres so that Knowledge Base can be shown by the learners instead of asking them to choose which cup contains more and some learners might guess. Question 15 on comparison, it could be better if the learners were asked to work out first before comparing so that the Strategy mark can be easily awarded. Question 17 was difficult to identify the Knowledge Base since bearing is working with angles. The expected answer was a three digit number and some learners were writing any three digit number. It could be better if the answers could have two digits and learners who add a zero in front of the two digits, could be the ones who clearly show their knowledge of bearing and can then be judged easily. Question 12 for the knowledge mark, learners could be asked to write 1:05 in 24 hours (clock time) notation and then asked to work out how long. Some Knowledge Base questions need to be specific, if the assessor want to assess the knowledge of Pythagoras Theorem, he should ask the learners to state it (Pythagoras Theorem).

APPENDIX I

Markers Comments Summary - **ROUND 2 (2014)**

Ma21: Overall the new developed framework seems to be more rewarding with the three levels of complexity having to be used in judging the learners work again. For example Question 9 (c) previously the learners were only awarded a mark for drawing lines but now the Strategy employed in shading were also awarded.

Again the marks allocated for the metacognition is very low in the current NSSCO question paper setup. This must be increased in order to allow proper assessment through the new developed framework e.g. Question 5(d) where the learners were required to solve $y = 2x^2 + 2x - 6$ and $y = 3x + 4$ simultaneously. In this case a lot of candidates knew that there are supposed to combine the two equations and obtain one trinomial equation and there was no more guidance given to them as to how they were going to solve that equation. It was therefore up to the learner to choose the method that will complete the question and with the new framework of marking the learners were awarded a mark for finding a method of their own that will solve the problem which is not looked at in the current NSSCO marking.

Ma22: Note 2014 was marked by allowing Strategy marks for a correct answer, even if the working was not seen. While as 2013 was marked with only the Knowledge Base mark given if no working was shown. There is therefore a big discrepancy (difference) in the marks between the old and the new framework.

Marking with the new framework highlighted the fact that there is often more to a question than meets the eye. (Prior knowledge is taken for granted by the examiner)

Questions often assess more than one aspect of mathematics which raises the question – is this good or bad? Should we be assessing bit-by-bit knowledge of the subject or the overall understanding and the ability to use knowledge in different areas (even in unfamiliar).

Examiners need to have a clear vision of the purpose of the question – what is being assessed? However, if a learner presents something different from what is expected, credit should be given e.g question 3 (c)(i).

In general

I noticed 2 types of G learners, although they are all weak learners, some did poorly throughout scoring and random marks in each question. However, some did badly in almost all questions but managed to score almost all their marks in only 2 or 3 question – i.e. selective understanding.

Foe answers, if a correct one is given, all is good. However, for a wrong answer the issue is; How wrong is it? This is seldom taken into consideration. Slightly wrong answer (with good correct thinking) is often marked as much wrong as no answer, or one with poor or flawed thinking. It would be nice to be able to differentiate between inaccurate final rounding from calculation error from concept error. Some sort of verbal examination would be great – (but horribly impractical logistically). Where by a learner would justify or explain the choice of strategy for example. Maybe an essay type question. Explain how you would set about planning/doing …'

G learners usually do very poorly with Geometry. Especially anything with angles.

Ma23

While the question in the paper do cater for all categories of proficiency they do not allow effective objective assessment of the other two, i.e knowledge base and metacognition. Assessment thus was more on strategy and less on the other two. The whole paper was silent on metacognition while KB was awarded on assumptions new rubric benefits the learner. The level of benefit increases with each grade. Grade A will be the most beneficiary in terms of marks while E benefited less – general increase in marks inn grades.

It was observed that almost every question requires previous knowledge and thus deserves a knowledge base mark. However it could be better if the questions are put in such a way that the candidate shows that knowledge instead of assuming that he or she knows. Questions should be specific to each category. Questions number 1 could have been (a) Find the square root of $\sqrt{\frac{1}{3}}$, (b) Work out $\frac{1}{3}\sqrt{\frac{1}{3}+\frac{1}{3}}\sqrt{\frac{1}{3}}$. No Metacognition mark should be awarded for different strategy since almost every question may have more than one strategy. Metacognition should be asked specifically.

Questions should be put in a language that will ask learners to do reflection of the process e.g which of the mean, mode and median will be best to use in presenting an argument for? The whole of the question paper was silent on metacognition.

Question on probability

Write down all factors of 15.

If a card is chosen at random find the probability that it is a factor of 15 and odd

Question 19 is a good example of testing knowledge followed by using the knowledge.

Extended paper must test more on metacognition

Metacognition should consist of a post evaluation in reflection on the strategy with reference to questions. Justification, hence look at your result and be able to comment.

At extended level more marks should

Ma24

The same trend of marks for all grades A, C and E were observed to follow the trend of the NSSCO 2013 paper 4. The team adjusted marks such that KB was 52 is now 57, S is 73 becomes 88 and M1 became 4 giving a total of 149 from 126. More credits could be allocated with the new developed framework as compared to the current NSSCO paper 4. The S eight now more with 59% whilst the KB and M percentages dropped.

Question 2 (c) could arrive at the same answer of 17 by other not recognised procedures. Examples: (a) 100 – 6 = 94 whereas $\frac{94}{100}\times 18 = 16.92 \approx 17$ or (b) 0.06 × 18 = 1.08 whereas 18 × 1.08 = $16.92 \approx 17$.

Question 8(b)(i) a possible M1 could have been thought of taking into account the possibilities of cosx or sinx as possible workings.

Did not see anything out of ordinary. Possibly because one has now adapted to marking with the new developed frameworks standards.

One observation that could be scrutinised is the following: we expect that a candidate who has performed well in the NSSCO paper should equally be the best performer in the new developed framework. However, one candidates' mark is 91 and another's previous mark is 86 in the NSSCO paper. However, the one with 91 in previous mark get 99 in the new developed framework whilst the one with an 86 gets 104. This trend could be investigated and found out why? This is not a normal scenario, but could be justified somewhere.

Ma25

My observation with the current question paper is still that learners with high symbols got high marks because they show their workings comparing to the learners with lower symbols.

My suggestion is that in order for us to find a clear picture on assessment through the new developed framework, is when a new rubrics has been set, and marked accordingly. If we look at all the papers which I marked, there was no M-score. Learners were not asked to prove their answers or to analyse them and find out which is the better solution.

Problem solving questions where learners supposed to choose between two or more strategies were not asked. Therefore the solution, is to set question papers that allowing learners to use both three categories (KB, S and M.) Examiners should set questions that encourage the learners to show their workings.

Assessment though the new developed framework allows the learners to score high marks comparing to the one used to mark the question papers before.

Ma26

The implementation of the new developed framework was better compared to 2013. The changed of the marks on the same level was slightly the same. The grade C moved to the range of 67 to 74 while for 2013 the marks of the same grade moved up to 80 marks. The changed of the marked of the knowledge and strategic, metacognition moved in the same correlation.

Almost every question covered the knowledge base but some question is a repetition of one basic competencies e,g Q14 (a) (b), all questions are asking the same basic competencies and question 15 (a), (b), (c) and (d) on probability. Questions can be rephrased in such a way that Q15(a) write a probability as a fraction (b) write a probability as a percentage. Although the questions is asking the probability, the way of writing answers will be different.

For the rounding to the given accuracy, the must be a limitation of using numbers giving the irrational numbers especially to question 4(b), 16(b), 19(b). This will avoid to penalise the learners more than twice on rounding off. But using the number giving the exact answers, this will be avoided unless the question want to assess the rounding off like Q11.

Some questions need strategic to be shown and this must be covered in the allocation of marks e.g Q10a, b before using <, > or = signs, the workings is needed. So to compare the numbers it will be advisable to work out first and arrange in ascending or descending order. For the use of <, > or =, there can be a use of integers or some knowledge base questions e.g

$3^0 \ldots\ldots\ldots\ldots 4^0$

$\frac{3}{5} \ldots\ldots\ldots\ldots \frac{4}{5}$

$\frac{5}{7} \ldots\ldots\ldots\ldots \frac{5}{6}$

$-13 \ldots\ldots\ldots\ldots -6$ etc.

Ma27

The current question paper or the marking does not award marks for recalling formulae but the new developed one does. This new developed framework make it easier for learners to score better especially with the KB marks. The current NSSCO question paper does not concentrate or award marks all the knowledge or steps that a learner takes as compared to the new developed framework.

There are few metacognition questions (marks) in the current NSSCO question paper. More questions that ask for a leaners reasoning's should be set. For example: Explain why in $\sqrt{a-2}$, a must be ≥ 2. Most of our questions at the moment only focus on KB or methods and answers, but we do not really test to see if they really understand why a certain step is done or to explain a mathematics concept.

APPENDIX J

FOCUS GROUP COMMENTS FORM

FOCUS GROUP
MEMBER:……

Your general observation with the current NSSCO question paper in relation to the new developed framework

Suggestion on the adaptation of the current NSSCO question paper to allow proper assessment through the new developed framework (Please give examples of questions where possible)

General Comments

APPENDIX K

CONSENT FORM

GENERAL INFORMATION

I Ms Elizabeth Ndjendja, am conducting a research study that investigates the Mathematical proficiency testing within the Namibian Senior Secondary Certificate Ordinary level (NSSCO) Mathematics examinations, framework. The study is done in fulfilment of the requirement of a degree of Doctor of Philosophy (PhD) in Mathematics education at Rhodes University.

You have been selected to take part in a **focus group discussion** to help me evaluate the NSSCO Mathematics answer scripts against the developed framework in order to identify the strength and the weakness of the new developed framework and of the current NSSCO Mathematics Examination framework. The focus group members will be required to develop the question paper specific rubrics from the defined standards and general rubrics given.

INFORMATION ABOUT STUDY PARTICIPANTS

Taking part in this study is completely voluntary. You do not have to participate if you don't want to. You may also leave the discussion group at any time. If you leave the study before it is finished, there will be no penalty to you and it will not affect any services you receive or are entitled to.

(a) Participants are required to be NSSCO Mathematics Chief Markers or Team Leaders. It requires a Mathematics teacher who is well acquainted with the current NSSCO Mathematics examination framework.

(b) It is expected that at most 8 people will take part in the focus group discussion and material development. The activities of this focus group are expected to take 12 to 20 hours spread over a few days.

(c) Your participation in this study may benefit you professionally.

CONFIDENTIALITY

The results and information obtained from you will remain highly confidential; your name will not be quoted during the reporting. To limit the possible bridge of confidentiality through the other focus group participants, the group will be addressed on the importance of keeping the information confidential.

CONSENT

By signing this consent form, you are indicating that you fully understand the above information and agree to participate in this focus group.

Participant's signature: _____

Printed name: _____

Date: _____

If you have any questions or concerns about this study, please contact the researcher;
Mrs Elizabeth Ndjendja at +264-81-2846494 or etndjendja@gmail.com.

APPENDIX L

REPUBLIC OF NAMIBIA

MINISTRY OF EDUCATION

Government Office Park
Ground Floor
Luther Street
Private Bag 12026
WINDHOEK

Enquiries: CM. Nyambe
Telephone: (061) 293 3432
Fax: (061) 293 3431

23 February 2015

Mrs Elizabeth Mutuku
P.O.Box 20639,
Windhoek.

Dear Mrs Mutuku,

Re: PERMISSION TO CONDUCT RESEARCH WITH NSSCO MATHEMATICS CANDIDATES ANSWER SCRIPTS

Your request regarding the above mentioned refers.

Approval is herewith granted for you to conduct the research with a number of Namibian Senior Secondary Certificate Ordinary level (NSSCO) Mathematics candidates answer scripts for the 2014 examinations, in fulfilment of your PhD course. Please ensure that the identity of the candidates whose scripts are going to be analysed is not revealed through the research outcome.

You are wished well with your research study and I look forward to the outcome of the research.

Yours faithfully,

MINISTRY OF EDUCATION
Office of the Permanent Secretary
25 FEB 2015

Mr A Iukena
Permanent Secretary: Ministry of Education